Journal of Statistical Physics, Vol. 76, Nos. 1/2, 1994

Statistical Properties of an Iterated Arithmetic Mapping

M. R. Feix,[1] A. Muriel,[2] and J. L. Rouet[3]

Received October 4, 1993; final April 6, 1994

We study the $(3x+1)/2$ problem from a probabilistic viewpoint and show a forgetting mechanism for the last k binary digits of the seed after k iterations. The problem is subsequently generalized to a trifurcation process, the $(lx+m)/3$ problem. Finally the sequence of a set of seeds is empirically shown to be equivalent to a random walk of the variable $\log_2 x$ (or $\log_3 x$) though computer simulations.

KEY WORDS: $3x+1$ problem; trifurcation; mixing process; random walk; numerical simulation.

1. INTRODUCTION

More than 60 years ago a conjecture concerning a very simple recurrence was introduced which is usually known under the name of the $(3x+1)$ problem,[7] but is called in the following the $(3x+1)/2$ problem.[4] It can be stated as follows. We consider $x_n \in \mathbb{N}^+$ (positive integers). The next term x_{n+1} in the recurrence is defined by the relations

$$\begin{aligned} &\text{if } x_n \text{ is even:} \qquad x_{n+1} = x_n/2 \\ &\text{if } x_n \text{ is odd:} \qquad x_{n+1} = (3x_n+1)/2 \end{aligned} \tag{1}$$

For example, starting from the seed $x_0 = 7$, we get 7, 11, 17, 26, 13, 20, 10, 5, 8, 4, 2, and the sequence ends in the cycle 2, 1, 2, 1,... . The following conjecture was put forward.

---

[1] PMMS/CNRS, 45071 Orléans Cedex 2, France.

[2] University of the Philippines, Manilla, Philippines.

[3] UFR Faculté des Sciences, 45071 Orléans Cedex 2, France.

[4] The historical name is the $(3x+1)$ problem. For logical reasons and especially since we generalize to the trifurcation case with the $(lx+m)/3$ problem, we feel that it is better to introduce into the name the denominator of the recurrence relation.

0022-4715/94/0700-0725$07.00/0

Conjecture. For any seed $x_0 \in \mathbb{N}^+$ the sequence ends on the cycle 2, 1, 2, 1,... .

After 60 years of effort no proof has been given. On the other hand, no counterexample has been found and numerical investigations have proved that the conjecture is true for all seeds x_0 up to 5.6×10^{13}.[1]

Recently, in attempting to understand this conjecture some authors have studied statistical properties of the mapping[2-6] and more precisely the relation between this sequence and a random process (basically a coin-tossing game). For a good review on the problem the reader is referred to ref. 7.

Let us be more precise on the relation between the sequence and the coin-tossing game. We first introduce $u_0 = \log_2 x_0$ and consider a great number of integer seeds x_0 (with all $x_0 \gg 1$) randomly drawn in such a way that u_0 has a density distribution. On each seed x_0 we perform k iterations according to the rules given by (1) (k must not be arbitrarily large and we will come back to this point in Section 4). We obtain a density distribution for the set of $u_k = \log_2 x_k$. Now, starting from the same distribution of u_0, we play the following chance game. We toss a coin. If a head is obtained, we take $u_{n+1} = u_n - 1$, which corresponds to the "even" transformation $x_{n+1} = x_n/2$. If tail is obtained, we take $u_{n+1} = u_n + \text{Log}_2 3 - 1$, which corresponds to the "odd" $x_{n+1} = (3x_n + 1)/2$ transformation, where, since we supposed x_n large enough, we can neglect $\text{Log}_2[1 + (1/3x_n)] \approx (1/3x_n)(\log 2)^{-1}$.

The density distribution of the u_n obtained in the two processes, namely the deterministic one and the random one, are empirically found to agree; see Section 4 for further details. Preliminary results have been given in ref. 2, while related results can be found in refs. 3 and 6.

Of course this result is not very surprising for the statistical physicist. The deterministic process corresponds to a microscopic model where the exact "trajectory" can be computed (here in a rather simple way) provided we know with great precision the value of the initial number. As in a real microscopic trajectory, two seeds differing by one unit (a small quantity, since x_n is supposed very large) have completely different trajectories. But here, as in many problems, we want to compute, not one single trajectory, but an ensemble. Consequently, the precise dynamic can be forgotten and a probabilistic approach can be introduced: here the equiprobability of the steps of size -1 and $\log_2 3 - 1$ corresponding to the "even" and "odd" iterations.

The purpose of this paper is twofold. First, we want to exhibit a mechanism which shows how the last k bits (for a seed written in base 2) are forgotten in exactly k iterations. Second, we generalize these results to more complex mapping involving three different cases for the iteration according to the value of x_n modulo 3. Descriptions of computer experiments follow.

2. MIXING MECHANISM

We first write the number in base 2. The last bit is 0 or 1. If it is zero, the number is even and can be written $2j$. The first iterate is j and since j is odd or even with equal probability, *a priori*, the property that the last bit is zero is forgotten in one step. If now the last bit is 1, the number can be written $2j+1$, its first iterate is $3j+2$, and the property that the last bit is one is also forgotten in one step. In both cases the only needed hypothesis is that the second bit j (second starting from the last one) has the same probability to be 0 or 1.

What about the mechanism for the two last bits? Then the initial number can be written $4j$; $4j+1$; $4j+2$; $4j+3$. Let us call these numbers respectively "0," "1," "2," "3" numbers (the name being the value of the number modulo 4). After one iterate a "0" number gives $2j$, which is a "0" number if j is even and a "2" number if j is odd. We write

$$\text{“0”} \rightarrow \text{“0” or “2”} \qquad \text{with equal probability}$$

A "1" number written $4j+1$ gives, as a first iterate, $6j+2$, i.e., again a "0" number if j is odd, and a "2" number if j is even:

$$\text{“1”} \rightarrow \text{“0” or “2”} \qquad \text{with equal probability}$$

The first iterate of a "2" number of the form $4j+2$ being $2j+1$, we find

$$\text{“2”} \rightarrow \text{“1” or “3”} \qquad \text{with equal probability}$$

In the same way a "3" number $4j+3$ gives for the first iterate $6j+5$

$$\text{“3”} \rightarrow \text{“1” or “3”} \qquad \text{with equal probability}$$

The situation can be summed up in the matrix $M_4(I, J)$ giving the probability than an "I" produces a "J" number after one iteration:

$M_4(I, J) =$

	$J=0$	$J=1$	$J=2$	$J=3$
$I=0$	1/2		1/2	
$I=1$	1/2		1/2	
$I=2$		1/2		1/2
$I=3$		1/2		1/2

One iteration is not enough to erase the two last bits of the seed. Let us consequently iterate a second time. To find the matrix "I" → "J" (after two iterations) we have simply to square M:

$$M_4^2 = \frac{1}{4} \begin{array}{cccc} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{array}$$

All the elements of the matrix are 1/4, indicating an equiprobability of the last two bits, irrespective of their initial values.

If we want to treat the three last bits we must consider the following matrix $M_8(I, J)$, when I and J run from 0 to 7:

$M_8(I, J) =$

	$J=0$	$J=1$	$J=2$	$J=3$	$J=4$	$J=5$	$J=6$	$J=7$
$I=0$	1/2				1/2			
$I=1$			1/2				1/2	
$I=2$		1/2				1/2		
$I=3$		1/2				1/2		
$I=4$			1/2				1/2	
$I=5$	1/2				1/2			
$I=6$				1/2				1/2
$I=7$				1/2				1/2

This time we check that $8M_8^3$ has all its elements equal to one.

We can give a more precise meaning to these results. For example, consider all the numbers which, written in base 2, have the same three last bits, for example 101. They can be written as $8j+5$. We know (see the bijection theorem in ref. 2) that they all have the same parity for the first three iterations. Indeed $8j+5$ is odd and gives $12j+8$, which is even; this last number gives $6j+4$ also even and finally $3j+2$. If now j takes all the values between 0 and 7, we check that $3j+2$ modulo 8 takes the respective values 2, 5, 0, 3, 6, 1, 4, 7.

To obtain the equiprobability for the eight numbers 0, 1,..., 7 formed by the three last bits, as pointed out by the form of the M_8^3 matrix, we must suppose that besides the three last bits, the three next most significant ones are taken randomly (i.e., j in $8j+5$ takes with equal probability the values 0, 1, 2,..., 7). Now we check that the same property holds for any number $8j+\alpha$, where, instead of 5, α takes all the values between 0 and 7.

The generalization to k bits of these properties can be precisely stated in the following way. For α in $A_k = \{0, 1,..., 2^k - 1\}$ we consider the ensemble S_α consisting of 2^k integers of $2k$ bits given by

$$2^k j + \alpha$$

where j runs over all elements of A_k. Consequently, each element of S_α has the same k last bits.

Theorem. (i) For each $\alpha \in A_k$, after k iterations applied to each element of S_α the last k bits of the resulting 2^k integers form a complete set of values in A_k.

(ii) The $2^k \times 2^k$ matrix M_{2^k} is such that

$$(M_{2^k})^k = 2^{-k} J_{2^k}$$

where J_{2^k} is the $2^k \times 2^k$ matrix with all entries equal to one.

To prove this, since the last k bits of the initial numbers are the same, the 2^k numbers experience the same sequence of even–odd iterations. We suppose that among these k iterations, I are odd. Moreover, we call β the kth iterate of α. The kth iterate of $2^k j + \alpha$ is $3^I j + \beta$. To obtain the last k bits of the new numbers we must take this last value modulo 2^k. It is enough to show that $3^I j$ modulo 2^k take all the values $0, 1,..., 2^k - 1$, since the addition of β will simply introduce a cyclic permutation (in fact a rotation if $0, 1, 2,..., 2^k - 1$ are located on regular polygon inscribed in a circle).

Now consider the numbers $3^I j$ and $3^I l$. We suppose $j > l$. Can these two numbers gives the same remainder after division by 2^k? We should have

$$3^I j = r + 2^k J; \qquad 3^I l = r + 2^k L$$

and consequently

$$3^I (j - l) = 2^k (J - L) \tag{2}$$

and we see that the first member of (2) must have a factor 2^k. Since 3^I is an odd number, this factor must be present in $j - l$. But the highest value of j is $2^k - 1$ and the smallest for l is zero. We consequently cannot have the two numbers equal modulo 2^k and we recover for the value modulo 2^k of the iterated numbers the elements of A_k. To prove (ii), multiply $(M_{2^k})^k$ on the left by a row vector with entry 1 in column α and 0's in all other columns. The resulting row vector is found by adding up 2^k row vectors, each of which has a single nonzero entry 2^{-k} in some columns, and the columns are given exactly by the last k bits of the kth iterates of (1) on the elements of S_α. Now by (i) all these columns are distinct, so the resulting row vector has all entries equal to 2^{-k}. But this vector is just the αth row of $(M_{2^k})^k$. Thus (ii) follows.

The theorem shows that after k iterations the last k bits are "forgotten," meaning that we have with equal probability the 2^k values of A_k.

We understand now the connection between the coin-tossing game and the deterministic sequence. In fact the connection is already apparent in the bijection theorem. This theorem states that if we consider all seeds with N bits, we have a 2^N-element set. Consider the sequence describing N iterations for each of these elements. The set of all sequences (odd–even,...) has at most 2^N elements (it could be less if two seeds give the same sequence). In fact each seed gives a different sequence and there is a bijection between the 2^N seeds and the 2^N sequences. Proofs of the bijection theorem can be found in refs. 2 and 7–9. This means that the following game can be played.

We select at random an N-bit seed, but give only the information on the parities of the N first iterations, asking if they have been produced by a coin tossing or by the $(3x+1)/2$ sequence. It is impossible to decide. This result was stated for the first time in refs. 8 and 9. Now what about giving more than N iterates of the parity sequence (i.e., a series of more than N odd–even)? The answer is subtle.

If one is provided with a long series, the cycle O, E, O, E becomes apparent if the sequence has been obtained from the deterministic process.

In fact, we do not need such long sequence if we are provided with the information that, possibly, the sequence has been produced by the $(3x+1)/2$ algorithm, with, of course, a finite number of bits for the seed. The inverse algorithm, described in ref. 2, will give the successive bits of the seed starting from the one with the lowest weight. If the seed has N bits, then after N steps the algorithm will indicate that we have exhausted all the significant bits and, correctly, will indicate zero for the next ones. This is a good example to exhibit the importance of an *a priori* information in Bayes' theorem. Not only must we be warned of the possible use of a deterministic algorithm, but we must know the rules of this algorithm to be able to decipher the subtle information contained in the parity sequence.

On the other hand, computer experiments suggest that, as long as the iterate is bigger than one, the deterministic sequence cannot be distinguished from a random one (obtained by coin tossing) provided, of course, we are not alerted to the possibility of a production by the $(3x+1)/2$ game (see Section 4). For example, the number of iterates needed to reach 1 (in the deterministic game) and of random steps (in the random walk of $U_n = \mathrm{Log}_2\, x_n$) are equal within a good precision. This means that the odd–

even sequence keeps its random appearance, although a subtle N-point correlation can distinguish the deterministic or random nature of the process.

3. GENERALIZATION TO A TRIFURCATION PROCESS

We consider the following recurrence: for $x_n \in \mathbb{N}^+$ let

$$\begin{aligned} x_{n+1} &= x_n/3 && \text{if } x_n \text{ modulo } 3 = 0 \\ x_{n+1} &= (l_1 x_n + m_1)/3 && \text{if } x_n \text{ modulo } 3 = 1 \\ x_{n+1} &= (l_2 x_n + m_2)/3 && \text{if } x_n \text{ modulo } 3 = 2 \end{aligned} \tag{3}$$

where

$$(l_1 + m_1) \text{ modulo } 3 = 0$$
$$(2l_2 + m_2) \text{ modulo } 3 = 0$$

A last condition we require is that l_1 and l_2 are each not multiples of 3. To see why, let us write the number x_n in ternary digits (using only 0, 1, and 2). The last digit determines the type of iteration—labeled 0, 1, or 2. We want to be able to use probabilistic arguments. For that to be possible, if the second digit, starting from the digit of smallest weight, is taken at random among 0, 1, 2, we need that the second iteration must be a random event with the three possibilities equally likely.

Now if the first digit is a zero, the number can be written $3j$. The first iteration is of type 0, the first iterate is j, and the next iteration is indeed a random event (if j is taken at random) with the three probabilities equal to 1/3. Notice that we do not consider digits of weight 3^2, 3^3, 3^4, since they do not play any role in the two first iterations.

Consider now a number $3j+1$ with a first iteration of type 1; the first iterate is $l_1 j + (l_1 + m_1)/3$. As we have previously stated, $(l_1 + m_1)/3$ must be an integer, say n_1. What is the last digit of $l_1 j + n_1$? Before giving the answer, we consider a number of type $3j+2$ with a first iteration of type 2 and a first iterate $l_2 j + (2l_2 + m_2)/3$, where $2l_2 + m_2$ is a multiple of 3, let us say n_2. So, our two iterates after type 1 or 2 iterations can be written $lj + n$. If l (i.e., l_1 or l_2) is a multiple of 3, the last digit of $lj + n$ is n modulo 3 irrespective of j and the second iteration does not exhibit the three equally likely possibilities. So we exclude l_1 and l_2 multiple of 3. We must then study $lj + n$ modulo 3 and consequently we can begin by taking l modulo 3 equal to 1 or 2 and we consider the values 0, 1, 2 for j and n. We build now Tables I and II.

Table I. Values of $lj+n$ for $l=1$ and the Three Possible Values of j and n[a]

	$n=0$	$n=1$	$n=2$
$j=0$	0	1	2
$j=1$	1	2	0
$j=2$	2	0	1

[a] All values are taken modulo 3.

We see that in both cases ($l=1$ and $l=2$) and for all possible values of n, in each column (i.e., for the three possible values of the second digit) we obtain the three possible types for the second iteration.

We see consequently that we can also extend to the "trifurcation" case the bijection theorem obtained in refs. 2, 8, and 9 for the $(3x+1)/2$ bifurcation problem. Our theorem is as follows:

Theorem. There is a bijection between the 3^k numbers with k ternary digits and the 3^k possible sequences of k iterations.

Let us give an example with $k=2$ and $l_1=4$, $m_1=-1$, $l_2=5$, $m_2=2$, and consequently $n_1=1$ and $n_2=4$. Tables I and II allow us to find the types of the two first iterates for all two-digit numbers. For example, what are the two first iterates for the number 21 (7 in decimal system)? The first is obviously 1, and l_1 modulo $3=1$, $n_1=1$, while $j=2$. We read the answer in Table I; for $l=1$, the column for $n=1$, and the third line, for $j=2$: The second iterate is of type 0 and the number 21 is in bijection with 01 (first iterate of type 1, second of type 0).

We easily obtained the following bijection (the first number gives the two last digits, the second the two types of iterations):

$$00 \leftrightarrow 00 \qquad 10 \leftrightarrow 10 \qquad 20 \leftrightarrow 20$$
$$01 \leftrightarrow 11 \qquad 11 \leftrightarrow 21 \qquad 21 \leftrightarrow 01$$
$$02 \leftrightarrow 12 \qquad 12 \leftrightarrow 02 \qquad 22 \leftrightarrow 22$$

Table II. Values of $lj+n$ for $l=2$ and the Three Possible Values of j and n[a]

	$n=0$	$n=1$	$n=2$
$j=0$	0	1	2
$j=1$	2	0	1
$j=2$	1	2	0

[a] All values are taken modulo 3.

and these results can be checked directly. The proof by induction for the generalization to k digits is easy and is omitted here.

In analogy with what has been done in the $(3x+1)/2$ problem, we now look at the forgetting mechanism of the successive ternary digits starting from the least significant one, i.e., the digit of unities. In fact, at the beginning of Section 3 we built the rules of the $(lx+m)/3$ game to exhibit this forgetting mechanism of this last digit. We defined three classes of numbers according their value modulo 3 (labeled respectively "0," "1," and "2" numbers). If the digit next to the last is taken at random and with equal probability among 0, 1, and 2, the first iterate will be a "0," "1," or "2" number (with equal probability for these three issues).

Let us now consider the two last digits. They define nine classes of numbers accordingly the value modulo 9 of the number and we consider "0," "1," "2,"..., "7," "8" numbers. Let us consider, for example, "0" numbers written $9j$. The first iterate is $3j$, which can be a "0" number if $j=0, 3, 6, 9,...$ or a "3" number if $j=1, 4, 7,...$ or a "6" number if $j=2, 5, 8,...$. We write "0" → "0" "3" "6" with equal probability.

The matrix $M_9^*(I, J)$ takes the following form:

$M_9^*(I, J) =$

	$J=0$	$J=1$	$J=2$	$J=3$	$J=4$	$J=5$	$J=6$	$J=7$	$J=8$
$I=0$	1/3			1/3			1/3		
$I=1$		1/3			1/3			1/3	
$I=2$		1/3			1/3			1/3	
$I=3$		1/3			1/3			1/3	
$I=4$			1/3			1/3			1/3
$I=5$	1/3			1/3			1/3		
$I=6$			1/3			1/3			1/3
$I=7$	1/3			1/3			1/3		
$I=8$			1/3			1/3			1/3

Now, to see what happens to an "I" number after two iterates (where $I \in \{0,..., 8\}$), we just have to square M_9^*. From the structure of M_9^* as given above, it is easily found that

$$(M_9^*)^2\,(I, J) = 1/9 \qquad \forall I, J$$

and the two last digits are forgotten after two iterations. The property holds for an arbitrary number of digits. N iterations erase N digits. In strict analogy with what has been done for the $(3x+1)/2$ problem, we state the following theorem. For α in the set $A_k^* = \{0, 1, 2,..., 3^k - 1\}$ (which will play the role of the previous A_k), we consider the set S_α^* consisting of 3^k integers of $2k$ ternary digits given by $3^k j+\alpha$, where j runs over all elements of A_k^*. Consequently all elements of S_α^* have the same k last ternary digits.

Theorem. (i) For each $\alpha \in A_k^*$, after k iterations of the recurrence (3) on the elements of S_α^* the last k ternary digits of the resulting 3^k integers form a complete set of elements A_k^*.

(ii) The $3^k \times 3^k$ matrix $M_{3^k}^*$ has

$$(M_{3^k}^*)^k = 3^{-k} J_{3^k}$$

where J_{3^k} is the $3^k \times 3^k$ matrix with all entries equal to 1.

The proof is similar to that for the bifurcation problem. Having the same last k digits, the sequence is the same for all numbers. We call respectively U and D the numbers of $(l_1 x+m_1)/3$ and $(l_2 x+m_2)/3$ iterations. Let β be the kth iterate of α, so that iterate for $3^k j+\alpha$ is $\beta + l_1^U l_2^D j$. We can again forget β (which will introduce a simple cyclic permutation) and we show that $l_1^U l_2^D j$ modulo 3^k gives all the elements of A_k^*. Considering the possibility that two numbers $l_1^U l_2^D j$ and $l_1^U l_2^D l$ (with $j>l$) have the same value modulo 3^k, we write

$$l_1^U l_2^D j = r+3^k J; \qquad l_1^U l_2^D l = r+3^k L$$

Consequently

$$l_1^U l_2^D (j-l) = 3^k (J-L)$$

and since l_1 and l_2 are not multiple of 3, the factor 3^k must be present in $j-l$, which is impossible since $l=0$ is the minimum and $j=3^k-1$ the maximum.

4. COMPUTER SIMULATIONS

We present now the simulations showing the possibility of describing by a random walk the evolution of an ensemble of seeds. We proceed as follows. A certain number of seeds (here 10^5) is selected randomly with a uniform probability among the 10^7 numbers ranging from $10^9-(1/2)\,10^7$ to $10^9+(1/2)\,10^7$. For these numbers the deterministic iterations are performed. From these "experiments" we deduce two types of curves:

$$\langle u_n \rangle = \langle \log_2 x_n(I) \rangle \qquad \text{for the bifurcation case}$$

$$\langle u_n \rangle = \langle \log_3 x_n(I) \rangle \qquad \text{for the trifurcation case}$$

$$\sigma_n^2 = \langle [u_n(I) - \langle u_n \rangle]^2 \rangle$$

where I refers to the number of the seed.

These quantities are compared with the "theoretical" curves of the theory of random walks:

$$\langle u_n \rangle = \langle u_0 \rangle - nm$$
$$\sigma_n^2 = n\sigma^2$$

where

$$m = (1/2) - (1/2)(\log_2 3 - 1)$$
$$\sigma^2 = (1/2)(-1 + m)^2 + (1/2)(\log_2 3 - 1 + m)^2$$

for the $(3x + 1)/2$ problem, and

$$m = (1/3) - (1/3)[\log_3 l_1 l_2 - 2]$$
$$\sigma^2 = (1/3)(-1 + m)^2 + (1/3)(\log_3 l_1 - 1 + m)^2 + (1/3)(\log_3 l_2 - 1 + m)^2$$

for the trifurcation problem.

Notice that the average value of the step is $-m$.

Figures 1 and 2 show the results, respectively, for the bifurcation and trifurcation problems. The straight lines are computed with the above values and describe the random walk. The points are the "experimental points" and describe the deterministic game. The agreement is excellent up to values of n (the number of iterations) of the order of 50–60.

An interesting thing to know is how large the number of iterations could be before the random walk model breaks down. Roughly speaking, for initial seeds centered around M the cycle is reached after k steps such that $\mathrm{Log}_2 M - km = 0$, where m has been given above. Consequently for the $(3x + 1)/2$ problem we conjecture for very large M a validity of the random walk picture for $m^{-1} \mathrm{Log}_2 M$ iterations. For the experiments shown Fig. 1, $\mathrm{Log}_2 M = 29.9$ and this calculus give 147 iterations. Nevertheless, the experimental curve show a breaking at a smaller value (typically 100 or even less if we consider the mean deviation). The explanation is indeed in the diffusion process: before the initial mean value of u_k has reached zero a sizable amount of iterated numbers have reached the cycle. With $m = 0.2075$ and $\sigma = 0.7925$, after 100 iterations, $u_0 - nm - \sqrt{n}\,\sigma \simeq 0$. We must consider greater seeds to obtain the asymptotic formula 4.82 $\mathrm{Log}_2 M$ for seeds peaked around the value M.

In the same way a sizable amount of iterated numbers have reached the fixed point 1 and the cycle 23, 53, 123, 41, 95, 221, 515, 1201, 801, 267, 89, 207, 69, 23,... in the trifurcation.

The existence of more than one cycle has been also found in the $(3x+m)/2$ (with $m \neq 1$) bifurcation process.[10,2]

Another, more detailed check consists in comparing the distribution density of $u_n(I)$ with the Gaussian defined by $\langle u_n \rangle$ (mean value) and $n\sigma^2$ (variance), which describes the random walk for large n. Figures 3 and 4 show the comparison for the same experiments as those used to obtain

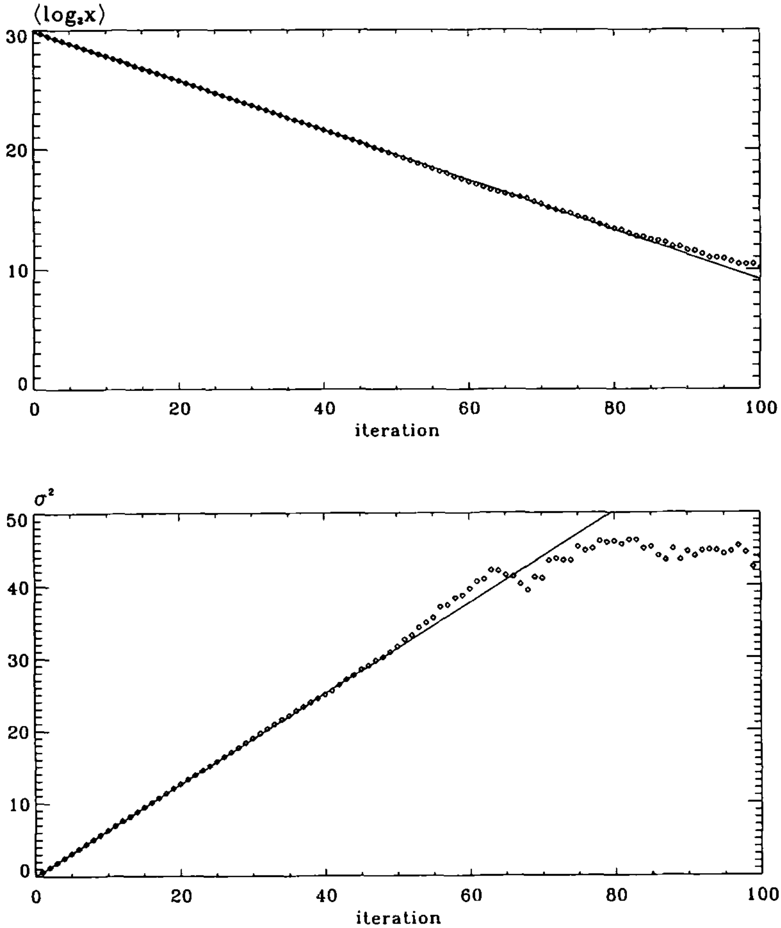

Fig. 1. Bifurcation problem: Average value $\langle u_n \rangle$ and mean squared deviation σn^2 as a function of n. The straight lines are given by the random walk approach. The points are "experimental" as given by the deterministic game. Notice the influence of the cycle for large n (of order 70).

Figs. 1 and 2. Here, up to the value $n = 40$, where we are still far from the cycles, the agreement is excellent, while for $n = 55$ we see that the effect of the cycle $\{1, 2\}$ becomes important.

5. CONCLUSION

First, about the $(3x+1)/2$ conjecture itself this paper does not present any really new material. The generalization of the bifurcation problem to the trifurcation problem simply shows that sometimes we can have many

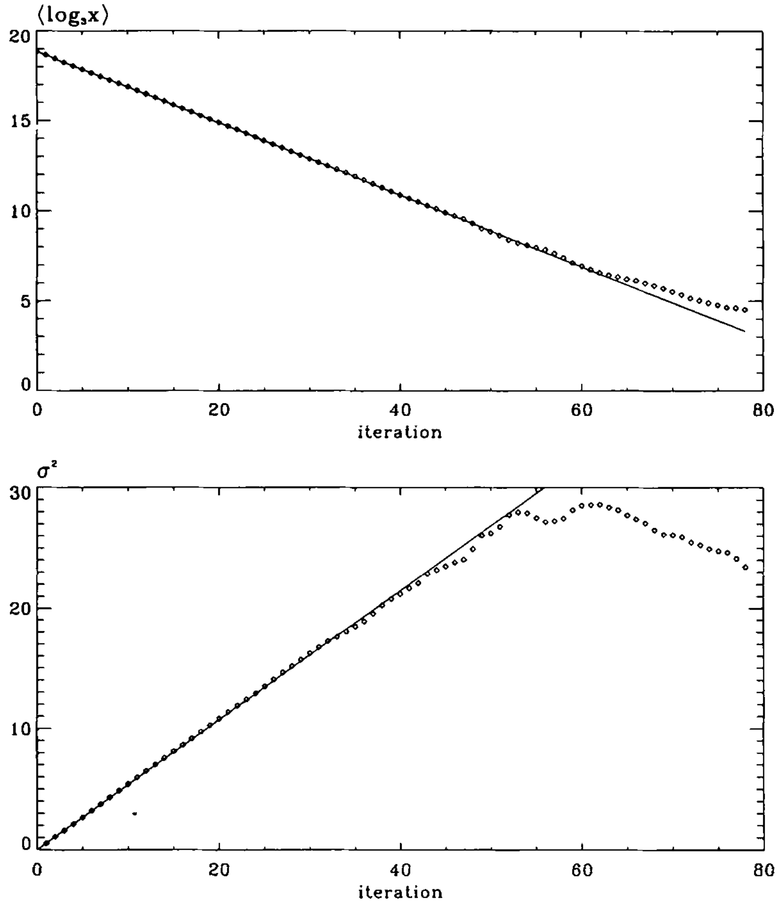

Fig. 2. The trifurcation problem: $l_1 = 2$, $m_1 = 1$, $l_2 = 7$, $m_2 = -2$; curves as for Fig. 1.

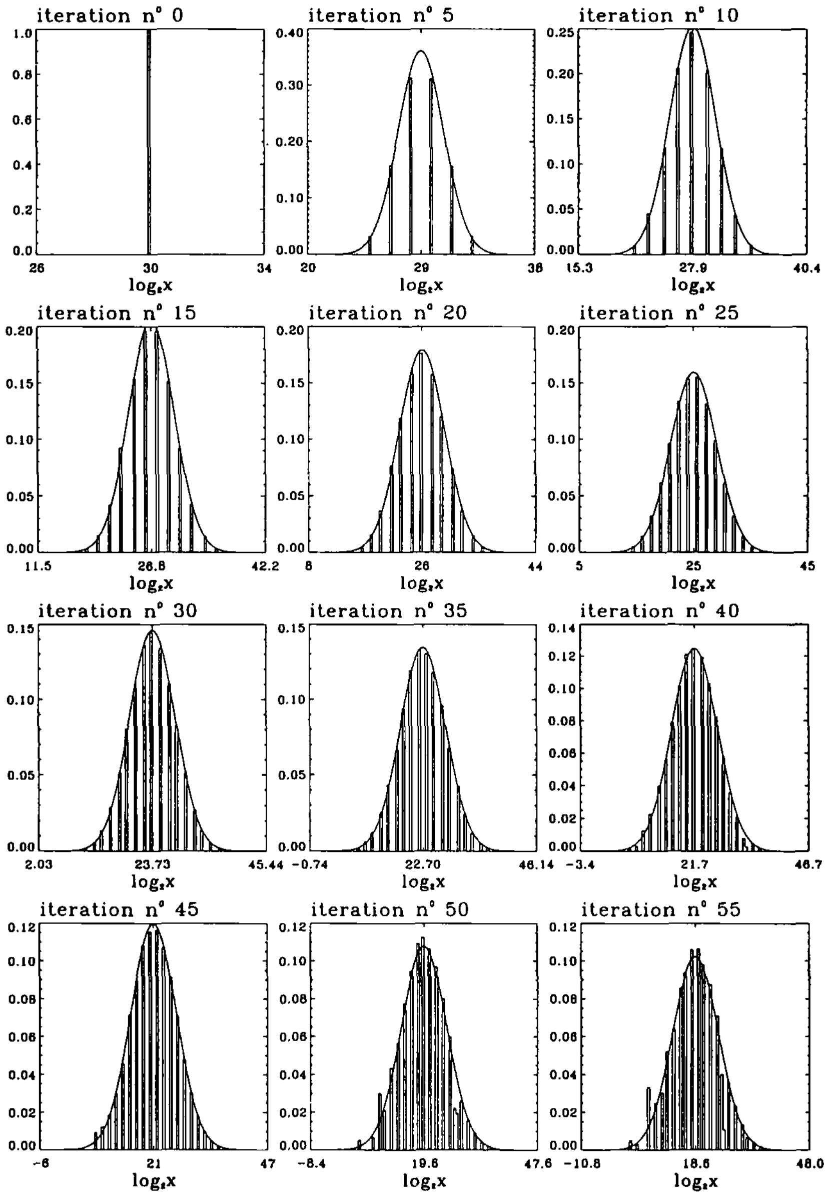

Fig. 3. The density distribution for u_n for different values of n. The Gaussians are obtained by the statistical approach. The "experimental" points are those used in Fig. 1 (case of bifurcation).

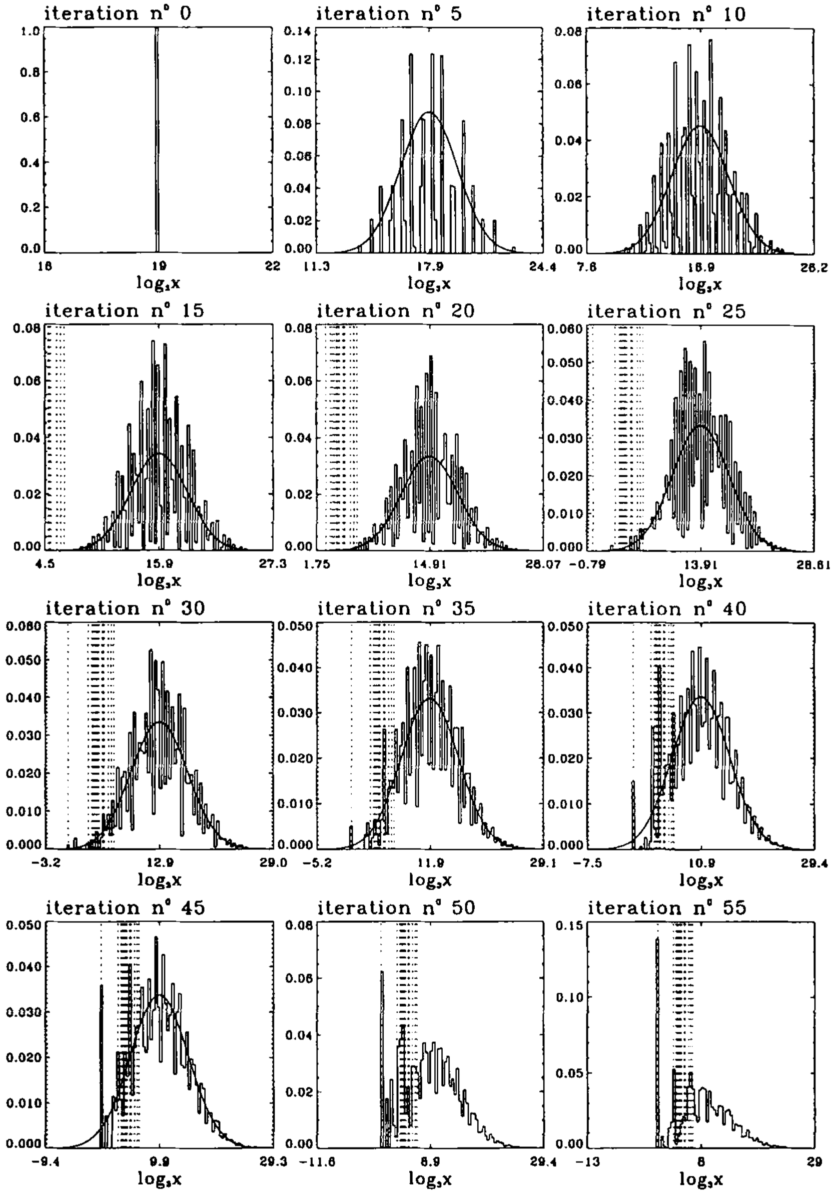

Fig. 4. The density distribution for the same experiments as in Fig. 2. The vertical dotted lines indicate the position of the numbers of the cycle and of the fixed point 1 (at the left). Note the accumulation of the iterated numbers on the cycle and on 1.0.

cycles plus eventually a fixed point. These cycles involve rather small numbers; for example, the cycle exhibited in the trifurcation problem of Figs. 2 and 4 has 23 as smallest and 1201 as largest number. But, of course, nothing precludes the existence of cycles containing large numbers. For the case treated in Figs. 2 and 4, a systematic research of the cycles has been conducted until 10^7 with no other cycle found and very likely we will have to go to much higher numbers to find the next one—if any.

Second, from a statistical point of view we saw, through computer experiments, that the sequence of parities (for a bifurcation problem) or the nature of iteration (0, 1, 2) in the trifurcation problem exhibits all the properties of random sequences—except that if we know that the initial seed has at most N digits and if we know the rules for the iteration, then we can recognize the nature of a sequence longer than N.

Still, from the statistical point of view we can view the sequence of iterations as a succession of "collisions" where the fate of the particle (number) is decided by the last digit. We need to know the number with an extremely great precision in the same way as in a collision we need to know with an extreme precision the impact parameter to know if the particle will be deflected to the right or to the left. The computer simulations justify the use of a random treatment. The diffusion process exhibited in Figs. 3 and 4 gives the behavior of a "particle" (number), the position of which is initially known within an accuracy of $(\log_2 10^9 - \log_2 0.995 \times 10^9)/\log_2 10^9 = 2.4 \times 10^{-4}$, if we consider the interval in which the starting u_0 are scattered.

The analogy with physics can be pushed further. The treatment of very large numbers is equivalent to a classical treatment where the quantum effects (in our game the fact that we must stay with integers) are negligible. But for low numbers cycles can appear stopping the diffusion process and bringing the process to a stationary one. But we should not push this analogy too far.

An interesting matter is to know if probabilistic arguments and computer simulations can help to understand the absence of cycles. *A priori* one is tempted to answer no and that the problem of cycles is strictly an arithmetic problem. But obviously, if we want to consider the possibility of cycles for a seed with, say, 100 bits we cannot proceed to a systematic search even with massively parallel supercomputers. But we can certainly sample some of these large numbers and follow the iterates. It may give some idea of an upper limit of the number of cycles for seeds smaller than a given (very large) number and also on the size of these cycles. Finally, a more sophisticated computer approach, going beyond an exhaustive search for successive integers, is needed.

ACKNOWLEDGMENT

We thank a referee for valuable suggestions.

REFERENCES

1. G. T. Leavens and M. Vermeulen, $3x+1$ search programs, *Comput. Math. Appl.* **24**(11):79–99 (1992).
2. M. R. Feix, A. Muriel, D. Merlini, and R. Tartini, The $(3x+1)/2$ problem: A statistical approach, in *Proceeding of the 3rd International Conference: Stochastic Process in Physics and Geometry* (Locarno, 1991).
3. D. A. Rawsthorne, Imitation of an iteration, *Math. Mag.* **58**(3):172–176 (1985).
4. R. E. Crandall, On the $3x+1$ problem, *Math. Comp.* **32**(144):1281–1292 (1978).
5. H. Müller, Das $3x+1$ problem, *Mitt. Math. Ges. Hamburg* **12**(2):231–251 (1991).
6. J. C. Lagarias and A. Weiss, The $3x+1$ problem: Two stochastic models, *Ann. Appl. Prob.* **2**:229–261 (1992).
7. J. C. Lagarias, The $3x+1$ problem and its generalization, *Am. Math. Monthly* **92**:3–23 (1985).
8. R. Terras, A stopping time problem on the positive integers, *Acta Arith.* **30**(3):241–256 (1976).
9. C. J. Everett, Iteration of the number theoretic function $f(2n)=n, f(2n+1)=3n+2$, *Adv. Math.* **25**(1):42–45 (1977).
10. J. C. Lagarias, The set of rational cycles for the $3x+1$ problem, *Acta Arith.* **56**(1):33–53 (1990).