Functiones et Approximatio
56.1 (2017), 7–23
doi: 10.7169/facm/1583

THE $3x + 1$ PROBLEM: A LOWER BOUND HYPOTHESIS

Olivier Rozier

Abstract: Much work has been done attempting to understand the dynamic behaviour of the so-called “$3x + 1$” function. It is known that finite sequences of iterations with a given length and a given number of odd terms have some combinatorial properties modulo powers of two. In this paper, we formulate a new hypothesis asserting that the first terms of those sequences have a lower bound which depends on the binary entropy of the “ones-ratio”. It is in agreement with all computations so far. Furthermore it implies accurate upper bounds for the total stopping time and the maximum excursion of an integer. Theses results are consistent with two previous stochastic models of the $3x + 1$ problem.

Keywords: $3x + 1$ problem, Collatz conjecture, parity vector, ones-ratio, binary entropy function, maximum excursion.

1. Introduction

Let us consider the T function acting on the set of positive integers and defined by

$$T(n) = \begin{cases} \frac{3n+1}{2} & \text{if } n \text{ is odd,} \\ \frac{n}{2} & \text{otherwise.} \end{cases} \tag{1}$$

It is expected but not yet proved that, whatever the initial value of n, the repeated iterations of T reach the value 1 at some point, thus entering the infinite loop $1, 2, 1, 2, \ldots$ called the *trivial cycle.* This question is notoriously intractable, despite its simple statement, and has received various names like the $3x + 1$ *problem*, the *Syracuse problem* or the *Collatz conjecture* [10].

Conjecture 1.1 ($3x + 1$ problem). *For any integer $n > 0$, we have $T^{(j)}(n) = 1$ for some $j \geqslant 0$, where $T^{(j)}$ denotes the j-th iterate of T.*

The $3x+1$ problem may be divided into Conjectures 1.2 and 1.3 below, asserting the absence of any other dynamic than the trivial cycle.

2010 Mathematics Subject Classification: primary: 11B37; secondary: 11B75, 60G50, 94A17

Conjecture 1.2 (Absence of divergent trajectory). *For all positive integer n, the infinite sequence $\left(T^{(k)}(n)\right)_{k=0}^{\infty}$, called the trajectory of n, is bounded.*

Conjecture 1.3 (Absence of non-trivial cycle). *There exist no integers $n > 2$ and $j > 0$ such that $T^{(j)}(n) = n$.*

We propose a heuristic approach, which is greatly inspired by a well-known paper of Lagarias [8], and we mostly follow his notations and denominations throughout the paper.

Combinatorial properties of T iterations are leading us to formulate a new hypothesis (see §2.2) involving the binary entropy function. So far, this function rarely appears in the vast literature on the $3x + 1$ problem with a few notable exceptions (e.g., [15, p. 84]). It has been used by Lagarias [8] to estimate the density of integers whose stopping time is bounded by a given value, thus improving a previous result of Terras [14]. Tao made a similar calculation to give a heuristic estimation of the number of non-trivial cycles, concluding that very likely there is none [13]. Let us also mention the application by Sinai of the notion of entropy of a dynamical system within a statistical modelling of the $3x+1$ problem [7, 12]. Besides, the binary entropy function is widely used in the context of information theory to express the entropy of Bernoulli processes.

In §3, we will see that proving our hypothesis would be more than sufficient to solve the $3x+1$ problem. Unexpectedly, it would further imply accurate upper bounds for the total stopping time and maximum excursion, which constitute the main result of the present paper (cf. Theorem 4.1). A brief comparison will be carried out with the predictions of the random walk model [9]. Then, in §5, we analyze a simple random model that is supporting our hypothesis.

Finally, we investigate a particular case of our hypothesis related to finite sequences of T iterations with only one even term.

2. A lower bound hypothesis

2.1. Combinatorial and heuristic approach

The T function exhibits remarkable combinatorial properties under iterations. Indeed, if we consider for each positive integer n and length j the *parity vector*

$$V_j(n) = \Big(n, T(n), \ldots, T^{(j-1)}(n)\Big) \mod 2, \tag{2}$$

then we have the following result proved independently by Terras [14] and Everett [5]:

Theorem 2.1 (Terras). *Two positive integers n and m have same parity vector of length j if and only if $n \equiv m \pmod{2^j}$.*

An immediate consequence is that every positive integer $n \leqslant 2^j$ is uniquely identified by its parity vector $V_j(n)$. Hereafter, let $I(j, q)$ denote the set of positive integers n for which there are exactly q occurrences of 1 in $V_j(n)$. Then, from $I(j,q)$, we extract the finite subset

$$I_0(j,q) = I(j,q) \cap \{1,2,3,\dots,2^j\}. \tag{3}$$

Conversely, it follows from Theorem 2.1 that $I(j,q)$ is the set of congruence classes modulo 2^j of $I_0(j,q)$ over the positive integers. It is easily seen that, for any fixed j, the set $\{I_0(j,0),\dots,I_0(j,j)\}$ is a partition of $\{1,\dots,2^j\}$ such that

$$\#I_0(j,q) = \binom{j}{q} \quad \text{for } q = 0,\dots,j \tag{4}$$

where # denotes the cardinality. As an example, we exhibit for $j=6$ the partition of $\{1,\dots,64\}$:

$$\begin{aligned}
&I_0(6,0) = \{64\},\\
&I_0(6,1) = \{16,20,21,32,40,42\},\\
&I_0(6,2) = \{4,5,6,8,10,12,13,24,26,34,35,48,49,52,53\},\\
&I_0(6,3) = \{1,2,3,11,17,22,23,25,28,29,36,37,38,44,45,46,50,51,56,58\},\\
&I_0(6,4) = \{7,9,14,15,18,19,30,33,43,54,55,57,59,60,61\},\\
&I_0(6,5) = \{27,31,39,41,47,62\},\\
&I_0(6,6) = \{63\}.
\end{aligned}$$

Lagarias suggested in [8] that the T function has some mixing properties under iteration, modulo powers of 2. One may further verify that the cumulative distribution function of $I_0(j,q)$ from 1 to 2^j appears fairly linear, for large j and q values. Therefore we may expect that the distribution of $I_0(j,q)$ over $[0,2^j]$ tends to be uniform and, roughly, that

$$\min I_0(j,q) \approx \frac{2^j}{\binom{j}{q}}. \tag{5}$$

Also, one may wonder whether a lower bound of the form

$$\min I_0(j,q) \geqslant j^{-C} \frac{2^j}{\binom{j}{q}} \tag{6}$$

holds for some positive constant C.

2.2. Hypothesis

The previous heuristic approach leads us to formulate the hypothesis to which this paper is dedicated:

Hypothesis 2.2. *For each $j \geqslant 1$ and $0 \leqslant q \leqslant j$, let $I_0(j,q)$ be the set of integers n with $1 \leqslant n \leqslant 2^j$ such that the vector*

$$\Bigl(n, T(n), T(T(n)), \dots, T^{(j-1)}(n)\Bigr)$$

contains exactly q odd terms. Then there is a real constant $C \geqslant 0$ such that, for all $j \geqslant 2$ and all $0 \leqslant q \leqslant j$, the set $I_0(j,q)$ is bounded from below by

$$j^{-C} \cdot 2^{(1-H(r))j} \tag{7}$$

where $r = q/j$ is called the "ones-ratio" and H is the binary entropy function $H(x) = -x\log_2(x) - (1-x)\log_2(1-x)$.

In the literature on the $3x+1$ problem, the "ones-ratio" usually denotes the proportion of odd terms in a sequence leading to the value 1. Hereafter, we extend this notion to all finite sequences of iterations, whatever the value of the last term.

The introduction of the binary entropy function H in Hypothesis 2.2 is due to the first order approximation

$$\log_2 \binom{j}{q} \sim H\left(\frac{q}{j}\right) j \tag{8}$$

for large values of j and q, which can be derived from the Stirling formula. Let us recall that H is a concave function with a single maximum $H(1/2) = 1$ and two minima $H(0) = H(1) = 0$ by continuous extension.

The value $H(r)$ in (7) is a measure of the entropy in the set of parity vectors $V_j(n)$ for $n \in I_0(j,q)$.

Here we provide another formulation of Hypothesis 2.2 with slightly less restrictive conditions, as it may be applied to the infinite sets $I(j,q)$ already introduced in §2.1, and includes the case $j = 1$.

Hypothesis 2.3 (Lower Bound Hypothesis – LBH). *There is a real constant $C \geqslant 0$ such that for all positive integers j and n not both equal to 1, we have*

$$n \geqslant j^{-C} \cdot 2^{\left(1-H\left(\frac{q}{j}\right)\right)j}, \tag{9}$$

where q is the number of odd integers in the vector $\left(n, T(n), \ldots, T^{(j-1)}(n)\right)$.

It is easy to see that Hypotheses 2.2 and 2.3 are equivalent with the same constant C. For convenience purposes, we shall simply refer to both of them by LBH. Nonetheless, the formulation from Hypothesis 2.3 will prove to be more suitable when studying all the implications related to the $3x+1$ problem.

One may first verify that LBH holds and is quite sharp in many cases, in the sense that there is an integer $n \in I_0(j,q)$ for which the lower bound (7) is reached with a small value of C. For example, it is sharp with C near zero for the two extremal values of the ones-ratio $r = 0$ and $r = 1$, since $I_0(j,0) = \{2^j\}$ and $I_0(j,j) = \{2^j - 1\}$ for all positive integer j. Remarkably, it is also sharp with $C = 0$ in the central case $r = 1/2$ arising when $j = 2q$. Indeed the set $I_0(2q,q)$ contains the integer 1 which has parity vector $V_j(1) = (1,0,1,0,\ldots)$.

In fact, LBH is true for any $C \geqslant 0$ in all cases where $r = q/j \leqslant 1/2$. This is a consequence of Lemma 2.4 below together with the inequality

$$H(x) \geqslant 2x, \quad \text{for } x \leqslant \frac{1}{2} \tag{10}$$

which follows from the concavity of the H function.

Lemma 2.4. *For all* $0 \leqslant 2q \leqslant j$, $\min I_0(j,q) = 2^{j-2q}$.

Proof. For all $n \in I_0(j,q)$, we write

$$1 \leqslant T^{(j)}(n) \leqslant 2^{2q-j} n$$

where the second inequality is easily obtained from the fact that $T(m) \leqslant 2m$ for any odd integer m. It follows that

$$\min I_0(j,q) \geqslant 2^{j-2q}.$$

To complete the proof, observe that $2^{j-2q} \in I_0(j,q)$. ■

As a result of Lemma 2.4 along with the strict concavity of the H function, it turns out that the lower bound in LBH is not sharp when the ones-ratio is strictly between 0 and 1/2. However, the numerical results in §2.3 will show that it can be sharp for sequences that tend to grow.

The forthcoming Theorem 2.6 states the validity of the inequality (9) for a large part of the sequences with a ones-ratio between 1/2 and $r_H = 0.609\ldots$[1]. It relies on Lemma 2.5, which is a generalization of a formula by Eliahou [4] regarding all cycles of the T function.

Lemma 2.5. *Let* $1 \leqslant q \leqslant j$ *and* $n \in I(j,q)$. *Then*

$$\frac{T^{(j)}(n)}{n} = 2^{-j} \prod_{k=0}^{q-1} \left(3 + m_k^{-1}\right) \tag{11}$$

where $m_0, \ldots, m_{q-1}$ *are the odd terms in the sequence* $\left(T^{(k)}(n)\right)_{k=0}^{j-1}$.

Proof. This result is straightforward to prove by applying the same method as in [4]. Indeed, we have

$$\frac{T^{(j)}(n)}{n} = \prod_{k=0}^{j-1} \frac{T^{(k+1)}(n)}{T^{(k)}(n)} = \frac{1}{2^{j-q}} \prod_{k=0}^{q-1} \left(\frac{3 + m_k^{-1}}{2}\right)$$

since $j-q$ is the number of even terms among $n, T(n), \ldots, T^{(j-1)}(n)$. ■

Theorem 2.6. *Let* $1 \leqslant q < j$ *such that* $r = q/j \leqslant \rho^{-1} = 0.630\ldots$, *where* $\rho = \log_2 3$, *and let* $n \in I(j,q)$ *for which the terms* $n, T(n), \ldots, T^{(j)}(n)$ *are all distinct. Then*

$$n \geqslant j^{-\frac{1}{6}} \cdot 2^{(1-\rho r)j}. \tag{12}$$

Assume further that $r \leqslant r_H = 0.609089767\ldots$, *where* r_H *is the unique non-zero real number such that*

$$H(r_H) \log 2 = r_H \log 3. \tag{13}$$

[1] The value of r_H already appears in various papers of Lagarias and coauthors (e.g., [7, p. 140]) as the ones-ratio upper limit for finite sequences leading to 1 in stochastic models of the $3x+1$ problem.

Then there holds the lower bound

$$n \geqslant j^{-\frac{1}{6}} \cdot 2^{(1-H(r))j}. \tag{14}$$

Proof. From Lemma 2.5, we write

$$\frac{1}{n} \leqslant \frac{T^{(j)}(n)}{n} = 2^{-j} \prod_{k=0}^{q-1} \left(3 + m_k^{-1}\right)$$

where $m_0, \ldots, m_{q-1}$ are the odd terms among $n, T(n), \ldots, T^{(j-1)}(n)$. This gives

$$\log n \geqslant j \log 2 - q \log 3 - \sum_{k=0}^{q-1} \log\left(1 + \frac{1}{3m_k}\right) \tag{15}$$

with

$$\sum_{k=0}^{q-1} \log\left(1 + \frac{1}{3m_k}\right) \leqslant \frac{1}{3} \sum_{k=0}^{q-1} \frac{1}{m_k} \tag{16}$$

by applying $\log(1+x) \leqslant x$. Since $m_0, \ldots, m_{q-1}$ are distinct odd numbers strictly greater than 1, we get

$$\sum_{k=0}^{q-1} \frac{1}{m_k} \leqslant \sum_{k=1}^{q} \frac{1}{2k+1}. \tag{17}$$

Then, from the upper bound

$$\sum_{k=1}^{q} \frac{1}{2k+1} \leqslant \frac{1}{2} \log q + \log 2 + \frac{1}{2}\gamma_{euler} - 1 + \frac{1}{2q}$$

where $\log 2 + \frac{1}{2}\gamma_{euler} = 0.981\ldots$, we infer

$$\sum_{k=1}^{q} \frac{1}{2k+1} \leqslant \frac{1}{2} \log q - \frac{1}{2} \log r = \frac{1}{2} \log j \tag{18}$$

for $q \geqslant 3$, using the fact that $r \leqslant \rho^{-1}$. It is easy to state inequality (18) for $q = 1, 2$. E.g., for $q = 1$, we check that

$$\frac{1}{3} < \frac{1}{2} \log 2 \leqslant \frac{1}{2} \log j.$$

It follows from the inequalities (15) to (18) that

$$\log n \geqslant j \log 2 - q \log 3 - \frac{1}{6} \log j,$$

or equivalently,

$$n \geqslant j^{-\frac{1}{6}} 2^{(1-\rho r)j}.$$

To conclude the proof, observe that the H function is strictly concave. This implies that there is a unique non-zero real r_H for which $H(r_H) = \rho r_H$, and that $H(x) \geqslant \rho x$ if and only if $x \leqslant r_H$. As a result, if we further assume that $r \leqslant r_H$, it yields $\rho r \leqslant H(r)$ and

$$n \geqslant j^{-\frac{1}{6}} 2^{(1-H(r))j}$$

as claimed. ■

The condition that all terms of the sequence are distinct in Theorem 2.6 is mandatory to ensure that no cycle appears. Otherwise, the inequality (12) would be easily falsified, by considering $n = 4 \in I(6,2)$ or $n = 27 \in I(74,43)$ for instance.

2.3. Numerical results

We computed the trajectory of all integers $n \leqslant 10^9$. According to the calculations, LBH holds with $C = 0$ for $n \leqslant 10^9$ in all cases where $j \neq q$, except for a few integers n given in Table 1. The value $c(n)$ in Table 1 denotes the smallest non-negative real such that the lower bound (9) from LBH holds for n provided $C \geqslant c(n)$, whatever the number j of iterations. Note that for $n = 1$, the case $j = 1$ is excluded.

Remark. One may verify that $c(n)$ exists whenever the trajectory of n contains the value 1. This follows from the fact that 1/2 is a critical point for the H function.

We found three successive records 0.472, 0.574 and 0.980 for the values of $c(n)$, corresponding to the integers $n = 27$, 159487 and 319804831, in that order. These numbers are already known as "maximum excursion" record-holders for the $3x+1$ problem [11] (see also §4.1). Table 2 gives the other known record-holders for the maximum excursion that are leading to non-zero values of the c function.

As a result of these calculations, if we assume LBH, then

$$C \geqslant 0.980916600\ldots. \tag{19}$$

However, we previously omitted the case $j = q$, which occurs when $n \equiv -1 \pmod{2^j}$. For $j > 1$, we get the positive lower bound

$$c(2^j - 1) \geqslant \frac{-\log(1 - 2^{-j})}{\log j} \tag{20}$$

by considering exactly j iterations. The highest value of the lower bound (20) is $\log(4/3)/\log(2) = 0.415\ldots$ obtained when $j = 2$. We verified that the equality holds in (20) for all $j \leqslant 1000$ except the cases $j = 5$ and $j = 6$ for which it is necessary to operate more than j iterations of T (see $n = 31$ and 63 in Table 1).

One observes, mainly in Table 1, that the integers n for which $c(n) > 0$ tend to form "clusters" of sequences with very similar lengths and ones-ratios. This is due to a well-known phenomenon of *coalescence* of sequences. For example, the trajectories starting from 27 and 31 are almost identical since $T^{(3)}(27) = 31$.

Table 1. Integers $n \leqslant 10^9$ such that $c(n) > 0$ and j, q values from which $c(n)$ has been derived. All cases where $j = q$ are omitted.

n	j	q	q/j	c(n)
1	3	2	0.666	0.154
27	45	33	0.733	0.472
31	42	31	0.738	0.408
41	44	32	0.727	0.265
47	41	30	0.731	0.195
54	46	33	0.717	0.132
55	46	33	0.717	0.127
62	43	31	0.720	0.058
63	43	31	0.720	0.053
73	48	34	0.708	0.001
159487	35	32	0.914	0.574
212649	37	33	0.891	0.195
239231	34	31	0.911	0.293
358847	33	30	0.909	0.008
5095423	29	28	0.965	0.091
19638399	199	140	0.703	0.034
21916159	40	37	0.925	0.045
319804831	91	77	0.849	0.980
379027947	96	80	0.833	0.774
426406441	93	78	0.838	0.773
479707247	90	76	0.844	0.776
568541921	95	79	0.831	0.575
598957743	103	84	0.815	0.418
639609663	92	77	0.836	0.571
719560871	89	75	0.842	0.571
758055895	97	80	0.824	0.386
852812883	94	78	0.829	0.375
898436615	102	83	0.813	0.226
959414495	91	76	0.835	0.368

3. Back to the $3x + 1$ problem

Hypothesis 2.3 (LBH) asserts that all integers $n \in I(j, q)$ with $j \neq 2q$ are lower bounded by a quasi-exponential function of j whose growth rate depends on the ones-ratio q/j. Moreover it implies that the ones-ratio of any given trajectory always converges to $1/2$ as the number of iterations tends to ∞, thus maximizing the binary entropy function. Proving this property would be sufficient to solve the $3x + 1$ problem, as shown by the next lemma.

Lemma 3.1. *LBH implies that the trajectory of any positive integer leads to the trivial cycle ($3x + 1$ problem).*

Table 2. Known maximum excursion record-holders $n > 10^9$ for which $c(n) > 0$.

n	j	q	q/j	c(n)
1410123943	197	144	0.730	0.145
272025660543	109	91	0.834	0.081
871673828443	107	91	0.850	0.327
3716509988199	201	155	0.771	0.426
9016346070511	202	155	0.767	0.113
1254251874774375	227	175	0.770	0.076
10709980568908647	298	222	0.744	0.077
1980976057694848447	399	292	0.731	0.408

Proof. Suppose that the trajectory of a given positive integer n does not contain the value 1. We may assume, without loss of generality, that n is the smallest term in the trajectory: $3 \leqslant n \leqslant T^{(j)}(n)$ for any j.

We obtain by Lemma 2.5

$$n \leqslant T^{(j)}(n) \leqslant 2^{-j}(3 + 3^{-1})^q \cdot n$$

where q is the number of odd terms among $n, T(n), \ldots, T^{(j-1)}(n)$. This gives

$$(3 + 3^{-1})^q \geqslant 2^j.$$

Therefore the ones-ratio $r = q/j$ has lower bound

$$r \geqslant r_0 = \frac{\log 2}{\log(3 + 3^{-1})} = 0.575\ldots$$

and $H(r)$ is upper bounded by $H(r_0) = 0.983\ldots$. It follows that the right-hand side in (9) is unbounded as j tends to infinity, yielding a contradiction with LBH. ■

4. Dynamic behaviour

4.1. Total stopping time and maximum excursion

Since Crandall [3], the dynamic of T is often compared to a multiplicative random walk with a downward drift. Several stochastic models [1, 7, 9] have been proposed in order to explain the empirical observations concerning the *total stopping time* $\sigma_\infty(n)$ and the *maximum excursion* $t(n)$ of a trajectory starting from n. Recall that $\sigma_\infty(n)$ is the number of iterations until the first occurrence of 1, and $t(n)$ is the highest term of the trajectory. Hypothetically, we set $\sigma_\infty(n) = \infty$ if the trajectory of n does not contain the value 1, and $t(n) = \infty$ if it is unbounded. The stochastic models predict that

$$\limsup_{n\to\infty} \frac{\sigma_\infty(n)}{\log n} = \gamma_{RW} \approx 41.677647 \tag{21}$$

and

$$\limsup_{n\to\infty} \frac{\log t(n)}{\log n} = 2, \tag{22}$$

which is consistent with all the empirical data provided independently by Oliveira e Silva and Roosendaal. As reported in [7], the highest known value of the ratio $\sigma_\infty(n)/\log n$ is equal to $36.716\ldots$, due to the finding of a new record-holder $n \approx 7.21 \times 10^{21}$. The accuracy of (22) is also discussed in [7, 11], where all known integers n such that $t(n) > n^2$ are given, starting with $n = 27$. At the time of writing, the verification process covers all integers n up to 5.76×10^{18}, thanks to various optimization techniques like the use of search trees for the preselection of congruence classes [11].

4.2. Main result

According to the next theorem, the expected dynamic behaviour of T under iteration, described as above, may be derived from LBH with some refinement to the second order.

Theorem 4.1. *Assume LBH. Then there hold the upper bounds*

$$\sigma_\infty(n) \leqslant \gamma_H \log n + O(\log\log n) \tag{23}$$

and

$$\log t(n) \leqslant 2\log n + O(\log\log n) \tag{24}$$

for integers $n \geqslant 2$, where

$$\gamma_H = (\log 2 - r_H \log 3)^{-1} = 41.677647655\ldots \tag{25}$$

with $r_H = 0.609089\ldots$ defined as in Theorem 2.6.

Proof of (23)**.** Let $n \geqslant 2$. Assuming LBH, we have $\sigma_\infty(n) < \infty$, by Lemma 3.1. Set $j = \sigma_\infty(n)$, so that $T^{(j)}(n) = 1$.

Let q be the number of odd terms among $n, T(n), \ldots, T^{(j-1)}(n)$.

The case $q = 0$ is simply stated: $n = 2^j$ and $j = (\log 2)^{-1}\log n$.

Consider now the case $0 < r = q/j \leqslant r_H$. Since $n, T(n), \ldots, T^{(j)}(n)$ are distinct, there holds formula (12) in Theorem 2.6, which in turn implies

$$j \leqslant \frac{\log n + \frac{1}{6}\log j}{(1-\rho r)\log 2}$$

with

$$(1-\rho r)\log 2 \geqslant (1-\rho r_H)\log 2 = \gamma_H^{-1}.$$

The case $r > r_H$ gives by assuming LBH

$$j \leqslant \frac{\log n + C\log j}{(1-H(r))\log 2}$$

with $C \geqslant 0$ a constant, and

$$(1-H(r))\log 2 > (1-H(r_H))\log 2 = \gamma_H^{-1}.$$

Thus, in all cases, we obtain the upper bound

$$j \leqslant \gamma_H(\log n + C\log j)$$

since $C > 1/6$, according to (19). It yields $j = O(\log n)$, and we finally get

$$\sigma_\infty(n) \leqslant \gamma_H \log n + O(\log\log n)$$

as claimed. ■

Proof of (24)**.** Let $n \geqslant 3$ be an odd integer. Assuming LBH, we have $t(n) < \infty$, by Lemma 3.1. Let $j \geqslant 1$ such that $T^{(j)}(n) = t(n)$. We may suppose, without loss of generality, that n is the smallest term among $n, T(n), \ldots, T^{(j-1)}(n)$.

Using Lemma 2.5, we get

$$\frac{t(n)}{n} \leqslant \frac{3^q}{2^j}\left(1+\frac{1}{3n}\right)^q$$

where q is the number of odd terms in the iterated sequence going from n to $T^{(j-1)}(n)$. Then we divide by n and apply LBH:

$$\frac{t(n)}{n^2} \leqslant j^C \cdot 2^{(\rho r + H(r) - 2)j} \cdot \left(1+\frac{1}{3n}\right)^q$$

with $\rho = \log_2 3$ and $r = q/j$. Now one verifies that $H(x) + \rho x \leqslant 2$ for any x where the equality holds if and only if $x = 3/4$. It follows that

$$\frac{t(n)}{n^2} \leqslant j^C \cdot \left(1+\frac{1}{3n}\right)^q$$

and, taking the logarithm,

$$\log t(n) \leqslant 2\log n + C\log j + \frac{q}{3n}$$

with the upper bound $\log(1+x) \leqslant x$. Since $q \leqslant j \leqslant \sigma_\infty(n)$, the inequality (23) gives the claimed result

$$\log t(n) \leqslant 2\log n + O(\log\log n).$$

This completes the proof of Theorem 4.1. ■

The above proof suggests that a maximum excursion record is more likely to occur when the ones-ratio of the sequence that goes from n to $t(n)$ is approximately $3/4$. This prediction is in good agreement with the empirical data in Table 2, mostly for long sequences.

Interestingly, the fact that γ_H does not depend on the value of the constant C in LBH further strengthens its relevance.

4.3. From the random walk model to entropy

One may ask whether the constants γ_{RW} and γ_H are identical. Though they are seemingly the same, their respective definitions differ. Recall that γ_{RW} originated in a model described in [9], namely the *random walk model.*

For each integer $n \geqslant 1$, Lagarias & al consider a sequence of i.i.d. random variables $X(n,k)$, $k \geqslant 1$, taking their values in the discrete set $\left\{\log 2, \log \frac{2}{3}\right\}$ with the same probability $\frac{1}{2}$. Starting from $\log n$, each random variable represents a single step towards $-\infty$ within some additive random walk on a logarithmic scale.

Using Chernoff's bound from the theory of large deviations, it was stated that almost surely

$$\limsup_{n\to\infty} \frac{\min_{k\geqslant 1}\left\{k : \log n - \sum_{i=1}^{k} X(n,i) \leqslant 0\right\}}{\log n} = \gamma_{RW}$$

where γ_{RW} is the unique solution with $\gamma_{RW} > \left(\frac{1}{2}\log\frac{4}{3}\right)^{-1}$ of the equation

$$\gamma_{RW} \cdot g\left(\frac{1}{\gamma_{RW}}\right) = 1. \tag{26}$$

The rate function g above is the Legendre transform

$$g(a) = \sup_{\theta\in\mathbb{R}} \left(a\theta - \log M_{RW}(\theta)\right) \tag{27}$$

defined for $\log\frac{2}{3} < a < \log 2$, with the moment generating function

$$M_{RW}(\theta) = \frac{1}{2}\left(2^{\theta} + \left(\frac{2}{3}\right)^{\theta}\right).$$

We give a more simple expression of the rate function in the next lemma[2].

Lemma 4.2. *Let* $0 < r < 1$. *Then*

$$g(\log 2 - r\log 3) = (1 - H(r))\log 2 \tag{28}$$

where g *is defined as in* (27) *and* H *is the binary entropy function.*

Proof. Set $a = \log 2 - r\log 3$ and suppose that θ^* verify

$$g(a) = a\theta^* - \log M_{RW}\left(\theta^*\right).$$

Writing the condition

$$\frac{d}{d\theta}\left(a\theta - \log M_{RW}(\theta)\right)\Big|_{\theta=\theta^*} = 0$$

[2] Lemma 4.2 relates the rate function to the binary entropy function. It can be derived directly by using another form of Chernoff's bound for the simple case of a Bernoulli distribution and based on the notion of relative entropy (see, e.g., [2] for this formulation).

leads to the relation

$$a - \log 2 + \frac{\log 3}{3^{\theta^*} + 1} = 0,$$

which simplifies to

$$r\left(3^{\theta^*} + 1\right) = 1.$$

We obtain after calculation

$$g(a) = r \log r + (1-r)\log(1-r) + \log 2. \qquad \blacksquare$$

Corollary 4.3. *Let γ_{RW} be defined by* (26)*, and let $\gamma_H = (\log 2 - r_H \log 3)^{-1}$ with $r_H = 0.609\ldots$ such that $H(r_H)\log 2 = r_H \log 3$. Then*

$$\gamma_{RW} = \gamma_H. \tag{29}$$

Proof. Setting $\gamma_{RW} = (\log 2 - r_{RW} \log 3)^{-1}$, we get from (26) and Lemma 4.2 the equation

$$H(r_{RW})\log 2 = r_{RW} \log 3$$

with $r_{RW} > \frac{1}{2}$. Thus, $r_{RW} = r_H$, since r_H satisfies the same equation that has a unique positive solution (see Theorem 2.6). The expected result (29) follows. $\blacksquare$

The authors of [9] also consider another stochastic model, namely the *branching process model.* A tree is build from a family of Bernoulli processes that imitate the backward iterations of the T function on a logarithmic scale. Applying a theorem of Biggins based on a Chernoff's bound for Galton-Watson processes, this model predicts that the left-hand side in (21) equals a constant γ_{BP}. Then it is shown that $\gamma_{BP} = \gamma_{RW}$, which is quite satisfying.

The identities $\gamma_{BP} = \gamma_{RW} = \gamma_H$ suggest that LBH is in full agreement with the predictions of both the random walk and the branching process models. As a possible explanation, one might consider that the heuristic reasonings leading to all of them are somehow related.

5. A random model of uniform distribution

The assumptions made in §2.1 give no indication on the value of the constant C in LBH. To this end, let us consider a random model[3] where the elements $n \in I_0(j,q)$ are represented by a set of independent random variables $\left\{X_{j,q,i} : i = 1, 2, \ldots, \binom{j}{q}\right\}$ having a continuous uniform distribution on the interval $[0, 2^j]$.

Let $P(j,q)$ denote the probability that

$$X_{j,q,i} < j^{-C} \cdot 2^{(1-H(r))j} \tag{30}$$

[3] This random model is much more simple than the random walk model [9]. We point out that, in our model, the number of elements in $I_0(j,q)$ is set to $\binom{j}{q}$, whereas, in the random walk model, the number of sequences of length j, starting from $\log n$ with $n \leqslant 2^j$, and having q terms considered as "odd" is a Gaussian random variable with mean $\binom{j}{q}$. Yet we expect that those models lead to similar predictions regarding LBH.

where i, j, q, r are taken such that $0 \leqslant q \leqslant j$ $(j \neq 0)$, $1 \leqslant i \leqslant \binom{j}{q}$ and $r = q/j$. The value of $P(j, q)$ obviously does not depend on i:

$$P(j, q) = j^{-C} \cdot 2^{-H(r)j}.$$

Let us introduce the infinite sum of probabilities

$$S(C) = \sum_{j=1}^{\infty} \sum_{q=0}^{j} \binom{j}{q} P(j, q),$$

which estimates the number of times the inequality (30) is satisfied over all admissible values of i, j, q. By the Stirling formula, we have

$$\binom{j}{q} \sim \frac{2^{H(r)j}}{\sqrt{2\pi r(1-r)j}}. \tag{31}$$

for $\varepsilon \leqslant r \leqslant 1 - \varepsilon$, with $\varepsilon > 0$ fixed. Then we get the approximations

$$\binom{j}{q} P(j, q) \sim \frac{j^{-\frac{1}{2}-C}}{\sqrt{2\pi r(1-r)}}$$

and

$$\sum_{q=0}^{j} \binom{j}{q} P(j, q) \sim \frac{j^{\frac{1}{2}-C}}{\sqrt{2\pi}} \int_0^1 \frac{dx}{\sqrt{x(1-x)}} = \sqrt{\frac{\pi}{2}} \cdot j^{\frac{1}{2}-C}.$$

One may verify (e.g., by developing the Stirling series to the next order) that the latter approximation is still valid when summing on q.

We infer that $S(C) < \infty$ if and only if $C > 3/2$, by considering the conditional convergence of the Riemann Zeta function on the real line. Thus, almost surely, inequality (30) occurs at most finitely many times when $C > 3/2$, as a consequence of the Borel-Cantelli lemma.

This simple model suggests that LBH is likely to be true for any $C > 3/2$ and all j sufficiently large. Nevertheless, there is no randomness in the sets $I_0(j, q)$, and the previous estimation of the plausible values of C may be flawed for various reasons[4]:

(i) The elements of $I_0(j, q)$ have a lower bound given by Lemma 2.4 when $r \leqslant 1/2$, and all constant $C \geqslant 0$ is admissible in that case.

(ii) The values $\min I_0(j, q)$ and $\min I_0(j + 1, q)$ are often equal, and thus, correlated. For example,

$$\min I_0(50, 30) = \min I_0(51, 30) = 103.$$

(iii) The values $\min I_0(j, q)$ and $\min I_0(j, q+1)$ are interdependent when leading to coalescent sequences after a few iterations.

Therefore, it remains plausible that LBH holds for a constant C lower than 3/2. Yet, the exact value of C does not matter in most cases, as the first order of the lower bound (9) is the exponential term.

[4] See also [7, p. 153] for a similar discussion on the random walk model.

6. A particular case

6.1. Effective lower bound

On a theoretical level, we know very little about the smallest element of the set $I_0(j,q)$ in all cases where $\log_3 2 < q/j < 1$, which relates to sequences that tend to grow and have at least one even term. Here we briefly investigate the most simple of those cases, that is $q = j-1$, for which $\#I_0(j,j-1) = j$.

Lemma 6.1. *Let $j \geqslant 2$. Then $I_0(j,j-1) = \{n_{j,k}\}_{k=0}^{j-1}$ with*

$$n_{j,0} = 2^j - 2 \quad \textit{and} \quad n_{j,k} = \left(\frac{2}{3}\right)^k \left(b_k(j-k) \cdot 2^{j-k} - 1\right) - 1$$

for $1 \leqslant k \leqslant j-1$, where $b_k(l) = 2^{-l} \mod 3^k$. Moreover, the set $I_0(j,j-1)$ is bounded from below by

$$2^{j/(1+\rho)} - 2 \tag{32}$$

with $\rho = \log_2 3$.

Proof. It is straightforward to state that the first $j-1$ iterates of $n_{j,k}$ are odd integers, except $T^{(k)}(n_{j,k})$. We have indeed

$$T^{(k)}(n_{j,k}) = b_k(j-k) \cdot 2^{j-k} - 2$$

for $1 \leqslant k \leqslant j-1$. Thus, it suffices to observe that $1 \leqslant n_{j,k} \leqslant 2^j$ for all k to conclude that the j elements of $I_0(j,j-1)$ are the $n_{j,k}$ integers.

Now we are left to prove (32). On the one hand, we can write

$$n_{j,k} \geqslant \left(\frac{2}{3}\right)^k \left(2^{j-k} - 1\right) - 1 \geqslant \frac{2^j}{3^k} - 2 = 2^{j-\rho k} - 2 \tag{33}$$

for $1 \leqslant k \leqslant j-1$, by using $b_k(l) \geqslant 1$. On the other,

$$n_{j,k} = 2^k \left(\frac{b_k(j-k) \cdot 2^{j-k} - 1}{3^k}\right) - 1 \geqslant 2^k - 1 \tag{34}$$

which follows from the fact that $(b_k(l) \cdot 2^l - 1)$ is a non-zero multiple of 3^k for any $l \geqslant 1$. Putting (33) and (34) together yields

$$n_{j,k} \geqslant 2^{\max(k,j-\rho k)} - 2.$$

The lower bound

$$\min_{1 \leqslant k \leqslant j-1} \max(k, j-\rho k) \geqslant \frac{j}{1+\rho}$$

completes the proof. ■

The effective lower bound (32) has an exponential growth, but it is quite weak compared to LBH which asserts that

$$\min I_0(j, j-1) \geqslant j^{-C} \cdot 2^{(1-H(1-1/j))j} \sim e^{-1} \cdot j^{-(C+1)} \cdot 2^j. \tag{35}$$

A simple calculation shows that the assumption (35) on the $n_{j,k}$ integers leads to the roughly equivalent lower bound

$$b_k(l) \geqslant \frac{3^k}{e \cdot (l+k)^D} \quad \text{for all } k, l \text{ positive integers}, \tag{36}$$

where $D \approx C + 1$ is a constant. To our knowledge, proving (36) is a non-trivial problem in number theory.

As the multiplicative group $(\mathbb{Z}/3^k\mathbb{Z})^*$ is cyclic of order $2 \cdot 3^{k-1}$, the b_k functions are periodic with period $2 \cdot 3^{k-1}$, and there holds

$$b_k(2 \cdot 3^{k-1}) = 1.$$

The inverse functions b_k^{-1} are the discrete logarithms in base $1/2$, modulo 3^k. Thus, the hypothetical lower bound (36) may be related to the distribution of discrete logarithms [6], which is expected to be uniform.

6.2. Further numerical results

In order to test numerically LBH in that particular case, we checked the lower bound (9) by setting $q = j - 1$, $n = n_{j,k}$ and $C = 0$ for all $0 \leqslant k < j \leqslant 10000$. When it was falsified, which occurred 3741 times, then we computed $c(n)$, where the c function is defined as in §2.3. Here we only give the three highest values found:

$$\begin{aligned} c\left(n_{85,56}\right) &= 0.865\ldots, \\ c\left(n_{2858,1270}\right) &= 0.817\ldots, \\ c\left(n_{5461,488}\right) &= 0.813\ldots. \end{aligned}$$

The three above $n_{j,k}$ integers have 22, 854 and 1637 decimal digits, in that order. Let us mention that the corresponding $c(n_{j,k})$ values have been obtained after exactly j iterations. These numerical results do not improve the bound (19), supporting the idea that LBH may hold with C near 1.

Acknowledgements. I wish to acknowledge for the fruitful correspondence with Claude Terracol that originated this research. I also thank Christian Boyer, Nik Lygeros and Clément Narteau for their positive influences, and the referee for his comments and suggestions.

References

[1] K.A. Borovkov, D. Pfeifer, *Estimates for the Syracuse problem via a probabilistic model*, Theory Probab. Appl. **45** (2001), 300–310.

[2] T.M. Cover, J.A. Thomas, *Elements of Information Theory*, John Wiley, New York, 1991.

[3] R.E. Crandall, *On the* $3x + 1$ *problem*, Math. Comp. **32** (1978), 1281–1292.

[4] S. Eliahou, *The* $3x+1$ *problem: new lower bounds on nontrivial cycle lengths*, Discrete Math. **118** (1993), 45–56.

[5] C.J. Everett, *Iteration of the number theoretic function* $f(2n) = n, f(2n+1) = 3n + 2$, Adv. Math. **25** (1977), 42–45.

[6] D.J. Gibson, *Discrete logarithms and their equidistribution*, Unif. Distrib. Theory **7** (2012), 147–154.

[7] A.V. Kontorovich, J.C. Lagarias, *Stochastic models for the* $3x + 1$ *and* $5x + 1$ *problems and related problems*, published in [10], 131–188.

[8] J.C. Lagarias, *The* $3x + 1$ *problem and its generalizations*, Amer. Math. Monthly **92** (1985), 3–23.

[9] J.C. Lagarias, A. Weiss, *The* $3x + 1$ *problem: two stochastic models*, Ann. Appl. Probab. **2** (1992), 229–261.

[10] J.C. Lagarias, editor, *The ultimate challenge: The 3x+ 1 problem*, Amer. Math. Soc., 2010.

[11] T. Oliveira e Silva, *Empirical verification of the* $3x+1$ *and related conjectures*, published in [10], 189–207.

[12] Y.G. Sinai, *Statistical* $(3x+1)$ *problem*, Commun. Pure Appl. Math. **56** (2003), 1016–1028.

[13] T. Tao, *The Collatz conjecture, Littlewood-Offord theory, and powers of 2 and 3*, blog post published August 25, 2011 at http://terrytao.wordpress.com/2011/08/25 .

[14] R. Terras, *A stopping time problem on the positive integers*, Acta Arith. **30** (1976), 241–252.

[15] G.J. Wirsching, *The Dynamical System Generated by the* $3n + 1$ *Function*, Lecture Notes in Math. 1681, Springer, 1998.

Address: Olivier Rozier: 237, rue de Romainville, 93230 Romainville, France.

E-mail: olivier.rozier@gmail.com

Received: 29 September 2015; **revised:** 28 September 2016