LOWER BOUNDS FOR THE TOTAL STOPPING TIME OF $3x + 1$ ITERATES

DAVID APPLEGATE AND JEFFREY C. LAGARIAS

ABSTRACT. The total stopping time $\sigma_\infty(n)$ of a positive integer n is the minimal number of iterates of the $3x + 1$ function needed to reach the value 1, and is $+\infty$ if no iterate of n reaches 1. It is shown that there are infinitely many positive integers n having a finite total stopping time $\sigma_\infty(n)$, such that $\sigma_\infty(n) > 6.14316 \log n$. The proof involves a search of $3x+1$ trees to depth 60, A heuristic argument suggests that for any constant $\gamma < \gamma_{BP} \approx 41.677647$, a search of all $3x + 1$ trees to sufficient depth could produce a proof that there are infinitely many n such that $\sigma_\infty(n) > \gamma \log n$. It would require a very large computation to search $3x + 1$ trees to a sufficient depth to produce a proof that the expected behavior of a "random" $3x + 1$ iterate, which is $\gamma = \frac{2}{\log 4/3} \approx 6.95212$, occurs infinitely often.

AMS Subject Classification (2000): Primary 11B83, Secondary 11Y16, 26A18, 37A45

1. INTRODUCTION

The $3x + 1$ problem concerns the behavior under iteration of the $3x + 1$ function $T : \mathbb{Z} \mapsto \mathbb{Z}$ given by

$$T(n) = \begin{cases} \frac{n}{2} & \text{if } n \equiv 0 \pmod 2 \\ \frac{3n+1}{2} & \text{if } n \equiv 1 \pmod 2 \end{cases} \tag{1}$$

The unsolved $3x+1$ Conjecture asserts that for each $n \geq 1$, some iterate $T^{(k)}(n) = 1$. It has been verified for all $n < 2.702 \times 10^{16}$ by Oliveira e Silva [9], whose subsequent computations have extended this bound to 1.125×10^{17}. The dynamical behavior of the $3x + 1$ function has been extensively studied. It is a deterministic process

Date: April 22, 2002.

which nevertheless appears to exhibit a kind of pseudorandom behavior. Surveys on the $3x+1$ problem can be found in Lagarias [6], Müller [8] and Wirsching [13].

To study the behavior of iterates of the $3x+1$ function on the positive integers quantitatively, we make the following definitions. For each n we call the minimal k such that $T^{(k)}(n) = 1$ the *total stopping time* of n and denote it $\sigma_\infty(n)$, letting $\sigma_\infty(n) = \infty$ if it is otherwise undefined. A rescaled version of the total stopping time, which we call the *stopping time ratio* $\gamma(n)$, is defined by

$$\gamma(n) := \frac{\sigma_\infty(n)}{\log n}. \tag{2}$$

A trajectory *converges* if it has a finite total stopping time, and we consider it to consist of n and all its iterates up to and including the first iterate k with $T^{(k)}(n) = 1$. Its *parity sequence* $\mathbf{v}(n)$ is the zero-one vector of length $\sigma_\infty(n)$ giving the residue (mod 2) of n and its iterates up to $T^{(k-1)}(n)$. The *ones-ratio* $\rho(\mathbf{v})$ of a zero-one vector $\mathbf{v}$ is the ratio of the number of ones in the vector to its length. By extension the *ones-ratio* $\rho(n)$ of a convergent trajectory is defined to be the fraction of odd integers appearing in its parity sequence, i.e. the ones-ratio $\rho(\mathbf{v}(n))$. For example, the trajectory of $n = 3$ is $(3, 5, 8, 4, 2, 1)$, with total stopping time $\sigma_\infty(3) = 5$, parity sequence $\mathbf{v}(3) = (1, 1, 0, 0, 0)$ and ones-ratio $\rho(3) = 2/5 = 0.4$. In this case $\gamma(3) = 5/\log 3 \approx 4.5512$.

This paper is concerned with rigorous results about the behavior of the stopping time ratio $\gamma(n)$. It is easy to see that $\gamma(n) \geq 1/\log 2$ for all n, and that equality holds exactly for $n = 2^k$ for $k \geq 1$. In this paper our object is to get lower bounds for the stopping time ratio that hold for infinitely many n, of the form $\gamma(n) \geq \gamma$, for some constant $\gamma > 1/\log 2$.

Studying the size of the stopping time ratio is essentially the same as studying the size of the ones-ratio, as given by the following bounds. For any convergent trajectory one has

$$\gamma(n) \geq \frac{1}{\log 2 - \rho(n)\log 3}, \tag{3}$$

while if $\rho(n) \le 0.61$ then, for any positive ϵ,

$$\gamma(n) \le \frac{1}{\log 2 - \rho(n) \log 3} + \epsilon,$$

holds provided n is sufficiently large, with $n > n_0(\epsilon)$. (It is believed that the condition $\rho(n) \le 0.61$ should hold for all sufficiently large integers (see below), while to get a uniform bound in terms of ϵ as above, one needs only assume $\rho(n) \le c_0 < \log 2/\log 3 \approx 0.63092$.) Our actual analysis will be based on study of the ones-ratio.

As background, we review the current conjectures about the size of the stopping time ratio. These conjectures were suggested by theoretical results proved for stochastic models. The predictions of these models for the total stopping time function successfully match experimental evidence for iterates of the $3x+1$ function. Simple models for average-case behavior of the stopping time function appear in Crandall [5], Lagarias [6], Rawsthorne [10], Lagarias and Weiss [7] and Borovkov and Pfeifer [4]. These models predict that the average value of $\gamma(n)$ should be

$$\frac{2}{\log 4/3} \approx 6.95212,$$

which corresponds to a ones-ratio of $1/2$ and that nearly all values of $\gamma(n)$ should be within $O((\frac{1}{\sqrt{\log n}})^{1+\epsilon})$ of this value. This is well supported by experimental evidence.

Lagarias and Weiss [7] developed stochastic models to describe the maximal values attained by the total stopping time. These models are analyzed using the theory of large deviations, and predict that the limit superior of $\gamma(n)$ as $n \to \infty$ should be a certain constant $\gamma_{BP} \approx 41.677647$, which corresponds to a ones-ratio of about 0.609091. (This constant is the solution to a certain functional equation given in [7].) Furthermore, the occurrence of values close to the extremal one are rare events whose probability can be estimated for the stochastic models, see [7, Theorem 2.2 and sec. 6]. Oliveira e Silva's record number below 10^{17} is $n = 1,008,932,249,296,231$ with $\sigma_\infty(n) = 1142$ and $\gamma(n) \approx 33.0558$. Backwards search methods have uncovered some larger integers which have larger values of $\gamma(n)$. E.

Roosendaal [11] found that

$$n = 7,219,136,416,377,236,271,195$$

has total stopping time $\sigma_\infty(n) = 1848$, with $\gamma(n) \approx 36.7169$.

We now turn to the problem of obtaining bounds α such that there are provably infinitely many positive n having $\gamma(n) \geq \alpha$. The current best value of α is that associated to the family of numbers $n = 2^k - 1$. It is easy to show that after k iterations one reaches $T^{(k)}(n) = 3^k - 1$. Since this number cannot get to 1 any faster than a power of 2 (which in fact occurs for $k = 2$), one obtains the rigorous lower bound

$$\gamma(2^k - 1) \geq \frac{\log 2 + \log 3}{(\log 2)^2} \approx 3.729. \tag{4}$$

This corresponds to a ones-ratio about 0.387. It is expected that the numbers $3^k - 1$ for most k have iterates behaving like a "random" integer, in which case the lower bound (4) could be significantly improved for infinitely many k. However, as far as is currently known, it could be the case that for each $\epsilon > 0$ the bound

$$\sigma_\infty(3^k - 1) < (\frac{1}{\log 2} + \epsilon)\log(3^k - 1),$$

holds for all sufficiently large k; if so then (4) could not be improved upon asymptotically as $k \to \infty$.

The object of this paper is to improve this lower bound, as follows.

Theorem 1.1. *There exist infinitely many $n \geq 1$ having a convergent trajectory with a ones-ratio of at least $14/29 \approx 0.4827$. Thus there is an infinite set of positive integers n such that $\gamma(n)$ is finite and*

$$\gamma(n) \geq \frac{29}{29\log 2 - 14\log 3} \approx 6.14316. \tag{5}$$

This result is proved in Section 4 by a study of all $3x + 1$ trees to depth 60, by extensive computation, involving trajectories having up to 29 ones.

In the rest of the paper we describe $3x+1$ trees, the form of proof certificates for lower bounds of the kind in Theorem 1.1, and the results of computations. We note that the results of this paper hold regardless of the truth of the $3x + 1$ conjecture.

The current best results on the number of integers below x that have convergent trajectories are found in Applegate and Lagarias [1], [2].

Notation. We let $\mathbb{N} = \{0, 1, 2, \cdots\}$ denote the natural numbers and $\mathbb{N}^+ = \{1, 2, 3, \cdots\}$ the positive integers.

2. 3x + 1 Trees

The $3x + 1$ iteration when run backwards from any fixed integer a produces a tree of preimages of a.

A $3x+1$ *tree* $\mathcal{T}_k(a)$ is a rooted, labelled tree of depth k, representing the inverse iterates $T^{-j}(a)$ for $0 \le j \le k$. The inverse map $T^{-1}(n)$ is multivalued:

$$T^{(-1)}(n) = \begin{cases} \{2n\} & \text{if } n \equiv 0 \text{ or } 1 \pmod 3, \\ \{2n, \frac{2n-1}{3}\} & \text{if } n \equiv 2 \pmod 3. \end{cases}$$

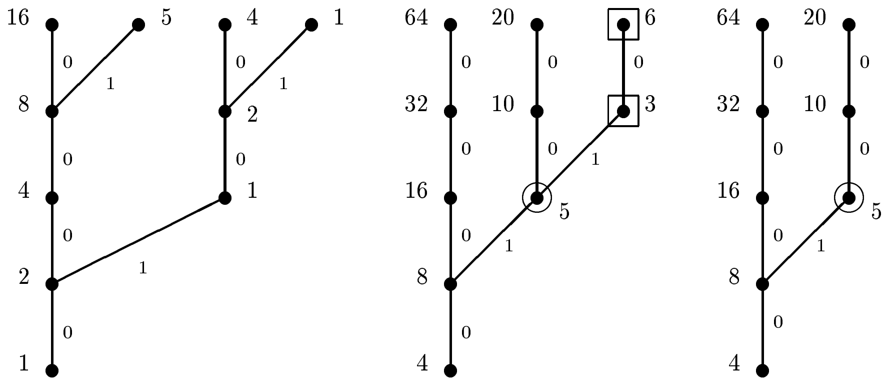
The root node a is at depth 0, and a node labelled n at depth d of the tree is connected by an edge to a node labelled $T(n)$ at depth $d-1$ of the tree[1]. Thus the *depth* of a node is the number of edges in a path from it to the root node. As described in [7], the nodes labelled $n \equiv 0 \pmod 3$ give rise only to a linear chain of nodes labelled $n' \equiv 0 \pmod 3$ at higher levels. It is convenient to remove all such nodes and study the *pruned* $3x + 1$ *tree* $\mathcal{T}_k^*(n)$ that consists of nodes $n \not\equiv 0 \pmod 3$. That is, we study the inverse map on the set of positive integers $n \not\equiv 0 \pmod 3$ given by

$$T^{*(-1)}(n) = \begin{cases} \{2n\} & \text{if } n \equiv 1, 4, 5 \text{ or } 7 \pmod 9, \\ \{2n, \frac{2n-1}{3}\} & \text{if } n \equiv 2 \text{ or } 8 \pmod 9. \end{cases}$$

Figure 1 presents some examples of $\mathcal{T}_k(a)$ and $\mathcal{T}_k^*(a)$. (Nodes $n \equiv 5 \pmod 9$ are circled to indicate that they have some preimage $T^{-1}(n) \equiv 0 \pmod 3$, and nodes $n \equiv 0 \pmod 3$ are indicated with a square.)

We assign to each edge of a $3x+1$ tree between depth d and $d-1$ an *edge label* 0 or 1, which is the value $n \pmod 2$ of the endpoint node at depth d. Traversing

[1] We adopt a convention of "unrolling" any cycles under T, so that the same node label may appear at different levels of the tree if a cycle is present (cf. Figure 1).



(i) $\mathcal{T}_4(1)$ (ii) $\mathcal{T}_4(4)$ (iii) $\mathcal{T}_4^*(4)$

FIGURE 1. $3x+1$ trees $\mathcal{T}_k(a)$ and pruned $3x+1$ tree $\mathcal{T}_k^*(a)$

a path from a leaf node labelled n in $\mathcal{T}_k^*(a)$ at depth d to the root node a, the successive edge labels encode the initial part of the parity vector $\mathbf{v}(n)$ of n. We define the *weight vector* $\mathbf{w}(n;a)$ to be the reversal of the vector $\mathbf{v}(n)$ to the root node, which gives the sequence of edge labels from the root node to the node n. The *weight* w of such a path is the sum of the edge labels on the path, and we sometimes say that the node n has *weight* $w(n)$. Note that $w(n)/d$ is the ones-ratio of this vector along the path. The *max-weight* $w_{max}(a)$ of a tree $\mathcal{T}_k^*(a)$ is the largest weight among all paths from a leaf node to the root node. Thus $w_{max}(a)/k$ is the maximal ones-ratio among any path in the tree $\mathcal{T}_k^*(a)$ from a depth k node to the root node.

The *structure* of a pruned $3x+1$ tree $\mathcal{T}_k^*(a)$ of depth k is its isomorphism class as a rooted, edge-labelled tree. That is, two pruned $3x+1$ trees $\mathcal{T}_k^*(a)$ and $\mathcal{T}_k^*(b)$ have the same *structure* if they are isomorphic as rooted trees by an isomorphism that preserves edge labels. The structure of a pruned $3x+1$ tree $\mathcal{T}_k^*(a)$ is completely determined by knowledge of $a \pmod{3^{l+2}}$, where $l = w_{max}(a)$. To demonstrate this, we grow the tree from the root, labelling each new node $n \pmod{3^m}$, where m is determined recursively as follows. The root node is assigned $m = w_{max}(a)+2$. For each node $n \pmod{3^m}$ at depth d, we assert that $m \geq 2$, and if $n \equiv 1, 4, 5$ or $7 \pmod{9}$ there is a single edge labelled 0 to a node at depth $d+1$ labelled

$2n \pmod{3^m}$. If $n \equiv 2$ or $8 \pmod{9}$, there is one node at depth $d+1$ as above plus a second edge labelled 1 to a node at depth $d+1$ labelled $(2n-1)/3 \pmod{3^{m-1}}$. Since we started with $m = w_{max}(a) + 2$ at the root node, and since no path has more than $w_{max}(a)$ edge labels of 1, the value of the exponents m stays at least 2 all the way to, and including, the leaves at depth k.

We can improve on this bound in one circumstance. We say that a tree $\mathcal{T}_k^*(a)$ is *critical* at max-weight (or max-level) l if it has max-weight $w_{max}(a) = l$ and if all nodes at depth $k-1$ have weights $w_{max}(a) - 1$ or less in their path to the root. The structure of a critical tree is completely determined by $a \pmod{3^{l+1}}$ where $l = w_{max}(a)$, rather than $a \pmod{3^{l+2}}$. The proof is the same as before, noting that the exponent m is greater than 1 on every node up to and including depth $k-1$, but m may equal 1 on some leaf nodes, on paths of weight exactly $w_{max}(a)$.

For trees of depth k, one has $l \leq k$; the case $l = k$ is necessarily critical, so by the discussion above all pruned tree structures $\mathcal{T}_k^*(a)$ are determined by $a \pmod{3^{k+1}}$. Thus there are at most $2 \cdot 3^k$ distinct pruned tree structures $\mathcal{T}_k^*(a)$ of depth k. The actual number $R(k)$ of distinct tree structures is smaller but still grows exponentially in k.

We now suppose given as data a congruence class $a \pmod{3^{l+1}}$. From it we grow a pruned $3x+1$ tree with root a to the (unique) depth k at which it becomes critical with max-weight l. We represent $a \pmod{3^{l+1}}$ as a ternary (or 3-adic) vector $\mathbf{c} = \mathbf{c}(a) = (c_0, c_1 ..., c_l)$ of length $l+1$, where

$$a = (c_l c_{l-1} ... c_0)_3 = \sum_{j=0}^{l} c_j 3^j,$$

with $c_j \in \{0, 1, 2\}$. We label this tree $\mathcal{T}_{[l]}^*(\mathbf{c})$, and denote its depth $k = k(\mathbf{c})$. Using the procedure above, we can construct node labels for this tree consisting of such 3-adic vectors of variable length m at each node. Figure 2 illustrates this, depicting the pruned tree $\mathcal{T}_5^*(4)$, which is critical with max-weight 2, and the same tree $\mathcal{T}_{[2]}(011)$ computed using root value $4 \pmod{27}$.

We define the *ones-ratio* $\rho(\mathbf{c})$ for a critical tree $\mathcal{T}_{[2]}^*(\mathbf{c})$ to be

$$\rho(\mathbf{c}) := \frac{l}{k},$$

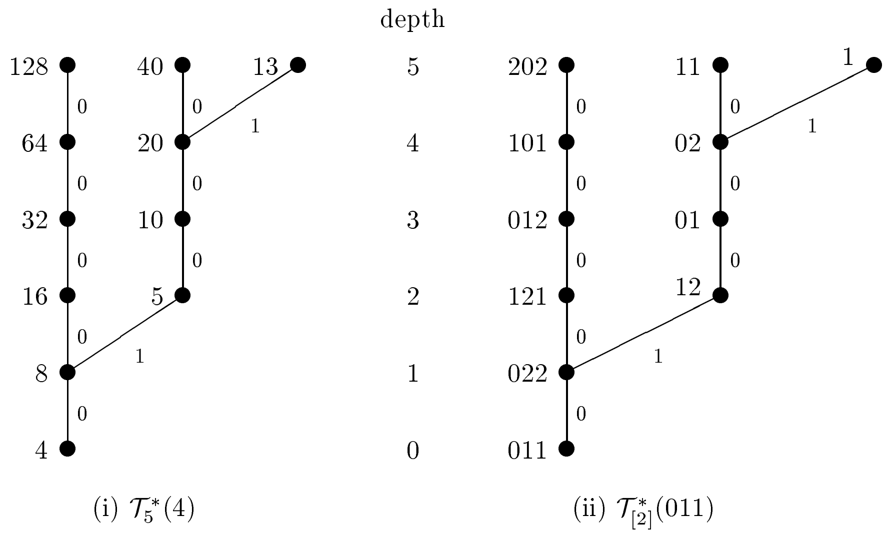


FIGURE 2. Pruned $3x+1$ tree $\mathcal{T}_{[k]}^*(\mathbf{c})$, $\mathbf{c} = 110$, for $a = (011)_3 \equiv 4 \pmod{27}$.

where k is the depth of the tree, and l is its criticality level.

3. Lower Bound Certificates

We describe a finite set of information ("certificate") that are sufficient to yield proof that there are infinitely many $n \geq 1$ having a finite total stopping time whose ones-ratio satisfies the lower bound

$$\rho(n) \geq \alpha \ . \tag{6}$$

Definition 3.1. A *prefix code* $\mathcal{C}$ for ternary sequences is a finite set of codewords $\mathbf{c} = (c_0, c_1, \ldots, c_l) \in \{0, 1, 2\}^{l+1}$ of varying lengths $l = l(\mathbf{c})$ having the properties:

(i) If $\mathbf{c}, \mathbf{c}' \in \mathcal{C}$ with $\mathbf{c} \neq \mathbf{c}'$, then $\mathbf{c}$ is not a prefix of $\mathbf{c}'$, and vice versa.

(ii) Any infinite sequence $\mathbf{s} = (s_1, s_2, s_3, \ldots) \in \{0, 1, 2\}^{\mathbb{N}^+}$ has an initial segment $(s_1, s_2, \ldots, s_{l+1})$ that is a codeword in $\mathcal{C}$.

Definition 3.2. A *(lower bound) certificate for ones-ratio* α consists of:

(i) A finite list $\mathcal{C}$ of residue classes $a \pmod{3^{l+1}}$ for various lengths, $l = l(a)$, which are exhaustive, in the sense that their associated vectors $\mathbf{c}(a)$, together with the additional vector $\mathbf{c}_0 = 0$, form a (ternary) prefix code.

$a \pmod{3^{l+1}}$	l	k	critical path
01	1	2	01
11	1	2	01
021	2	6	000101
121	2	6	000101
0221	3	9	000001101
1221	3	9	000001101
02221	4	12	000100000111
12221	4	11	00010001011
22221	4	11	00010001011
02	1	1	1
12	1	3	001
22	1	1	1

TABLE 1. Certificate for $\alpha = \frac{1}{3}$.

(ii) To each $a \in \mathcal{C}$ with $a \pmod{3^{l+1}}$ there is assigned a *critical path vector* $\mathbf{w}(a) \in \{0,1\}^k$, where k satisfies

$$l/k \geq \alpha \ ,$$

such that $\mathbf{w}(a)$ has weight at most l, and if it has weight exactly l, then a 1 occurs in its last coordinate, and $\mathbf{w}(a) = \mathbf{w}(n; a)$ for some leaf node of this $3x+1$ tree, i.e. $\mathbf{w}(a)$ is a labelled path in the $3x+1$ tree of $a \pmod{3^{l+1}}$, from the root node to some node at depth k.

The certificate vector certifies that the associated critical tree $\mathcal{T}^*_{[l]}(\mathbf{c}(a))$ with max-level l has depth[2] k' with $k' \leq k$ and has ones-ratio $l/k' \geq \alpha$.

As an example, we present in Table 1 a certificate for level $\alpha = 1/3$. In this table the numbers $a \pmod{3^{l+1}}$ are represented in their ternary expansion, so form a suffix code, rather than a prefix code. (The vectors $\mathbf{c}(a)$ reverse the order of the digits.) This certificate has max-weight 4 and max-depth 12, and consists of 12 different trees.

Definition 3.3. A *strong lower bound certificate for ones-ratio* α consists of a finite list $\mathcal{C}$ as above satisfying:

[2] A single path in a $3x+1$ tree sometimes can be grown to a depth greater than k, from knowledge of $a \pmod{3^{l+1}}$, as long as it has weight at most l, and if exactly weight l then it has a one in its last coordinate.

(i$'$) Same as (i) above.

(ii$'$) With each $a \pmod{3^{l+1}} \in \mathcal{C}$ are given two critical path vectors $\mathbf{w}_j(a) \in \{0,1\}^{k_j}$, with weights $l_j \le l$, respectively, such that neither is a prefix of the other, and where k_j and l_j satisfy

$$l_j/k_j \ge \alpha \ ,$$

and if $\mathbf{w}_j$ has weight exactly l then it has a 1 in its last coordinate, and each $\mathbf{w}_j(a)$ can be found as a labeled path of length k_j from the root in the $3x+1$ tree of $a \pmod{3^{l+1}}$.

As an example, there exists a strong lower bound certificate for $\alpha = \frac{1}{3}$ which has max-weight 4 and max-depth 12, the same as the lower bound certificate above, but which consists of 42 trees. It appears in Table 2 below.

Theorem 3.1. *(i) The existence of a lower bound certificate for ratio α implies: Given any $a \ge 1$ with $a \not\equiv 0 \pmod 3$, not in a cycle, there exist infinitely many n having*

$$T^{(k)}(n) = a \ , \tag{7}$$

for some $k \ge 1$, and such that the ones-ratio $\rho(n;a) := \rho(\mathbf{w}(n;a))$ of the path from n to a has

$$\rho(n;a) = l/k \ge \alpha \ , \tag{8}$$

(ii) The existence of a strong lower bound certificate for ones-ratio α implies: Given any $a \ge 1$ with $a \not\equiv 0 \pmod 3$ the number of $n \le x$ with $T^{(k)}(n) = a$ for some $k \ge 1$ and with

$$\rho(n;a) = l/k \ge \alpha \ , \tag{9}$$

is at least $c(a)x^{\delta}$ for certain $c(a) > 0$ and $\delta = \delta(a) > 0$.

Proof. (i) Since a is not in a cycle, all the nodes in the (infinite) pruned $3x+1$ tree $\mathcal{T}^*(n)$ with root node a have distinct values.

$a \pmod{3^{l+1}}$	l	k_1	critical path 1	k_2	critical path 2
001	2	4	0101	6	010001
101	2	4	0101	5	00011
201	2	5	00011	6	010001
0011	3	7	0100101	8	01000011
1011	3	7	0100101	9	010010001
2011	3	8	01000011	9	010010001
111	2	3	010	3	011
211	2	3	011	5	01001
0021	3	8	00010101	9	000100011
1021	3	6	000101	9	000100011
02021	4	10	0001010101	12	000101010001
12021	4	10	0001010101	11	00010100011
22021	4	11	00010100011	12	000101010001
0121	3	8	00000111	9	000101001
1121	3	7	0001011	8	00000111
2121	3	7	0001011	9	000101001
00221	4	11	00000100111	12	000001101001
10221	4	10	0000011011	12	000001101001
20221	4	10	0000011011	11	00000100111
01221	4	11	00000110101	12	000001100011
11221	4	11	00000110101	12	000100010101
21221	4	9	000001101	12	000100010101
002221	5	14	00010000011101	15	000100000110011
102221	5	12	000100000111	15	000100010100101
202221	5	14	00010000011101	15	000100010100101
12221	4	11	00010001011	12	000100000111
22221	4	11	00010001011	12	000001001101
02	1	1	1	3	001
012	2	5	00101	6	000011
0112	3	7	0010101	8	00100011
1112	3	7	0010101	9	001010001
2112	3	8	00100011	9	001010001
212	2	3	001	6	000011
022	2	4	1001	6	100001
122	2	2	11	4	1001
222	2	2	11	6	100001

TABLE 2. Strong certificate for $\alpha = \frac{1}{3}$

We can find an infinite path $\mathbf{s} = (s_1, s_2, s_3, \ldots) \in \{0,1\}^{\mathbb{N}^+}$ from the root node a in $\mathcal{T}^*(a)$ which contains infinitely many nodes satisfying (8). View a as a 3-adic integer, encoded as

$$\mathbf{c}(a) = (c_0, c_1, c_2, c_3, \ldots) \in \{0,1,2\}^{\mathbb{N}} \; ;$$

in which $c_0 \neq 0$ and $c_j = 0$ for all sufficiently large j. Then $\mathbf{c}(a)$ has a prefix in the certificate $\mathcal{C}$, of length $l+1$, say. By property (ii) of the certificate, there exists a path in the associated tree $\mathcal{T}^*_{[l_1]}(\mathbf{c}(a))$ from leaf node n to a with

$$T^{(k_1)}(n_1) = a$$

and

$$\rho(n; a) = \frac{l_1}{k_1} \geq \alpha \ , \tag{10}$$

where k_1 is the depth of $\mathcal{T}^*_{[l]}(c(\mathbf{a}))$. Now $n_1 \in \mathbb{N}^+$ is not in a cycle and $n_1 \not\equiv 0 \pmod{3}$, so we can repeat the same construction, and find a prefix of $\mathbf{c}(n_1)$ of length l_2+1, say, with tree $\mathcal{T}^*_{[l_2]}(\mathbf{c}(n_1))$ of depth k_2 having a path from leaf node n_2 to n_1 with

$$T^{(k_2)}(n_2) = n_1$$

and

$$\rho(n_2, n_1) = \frac{l_2}{k_2} \geq \alpha\,.$$

Thus

$$T^{(k_1+k_2)}(n_2) = a$$

and

$$\rho(n_2, a) \geq \min\left[\frac{l_2}{k_2}, \frac{l_1}{k_1}\right] \geq \alpha \ . \tag{11}$$

Continuing inductively, we find an infinite chain $a \leftarrow n_1 \leftarrow n_2 \leftarrow n_3 \leftarrow \ldots$ for which (7) and (8) hold.

(ii) We use a similar argument. Suppose first that a is not in a cycle. Then at the first round we find two different paths in $\mathcal{T}^*_{[l_1]}(\mathbf{c}(a))$ with $n_1 \to a$, $n_1' \to a$ satisfying (8). At the j-th stage we produce 2^j distinct elements $n_j \to a$ satisfying (8). Furthermore, if the largest depth tree $\mathcal{T}^*_{[l]}(\mathbf{c})$ for $\mathbf{c} \in \mathcal{C}$ is of depth d then, all the elements n_j are at depth at most jd from the root node a, hence each such element satisifes $n_j \leq 2^{jd}a$. Let $\pi_a^\alpha(x)$ count the number of elements $\leq x$ which satisfy (7) and (8). It follows that for $2^{jd}a \leq x < 2^{(j+1)d}a$ we have

$$\pi_a^\alpha(x) \geq 2^j > \frac{1}{2}\left(\frac{x}{a}\right)^{1/d} = c(a)x^{1/d} \tag{12}$$

with $c(a) = \frac{1}{2}a^{-1/d}$.

Now suppose that a is in a cycle, i.e. is periodic. Then in the first tree $\mathcal{T}^*_{[l]}(\mathbf{c}(a))$ at least one of the two preimages n_1 and n_1' of a that satisfies (9) cannot be in the periodic cycle containing a. Let it be n_1, and we may then apply the argument above to the tree with root n_1 for which

$$\pi^{\alpha}_{n_1}(x) \geq \frac{1}{2}(n_1)^{-1/d}x^{1/d} \ .$$

But any such element $n \to n_1 \to a$ by adjoining the path from n_1 to a, and the inequality (11) gives

$$\rho(n,a) \geq \min[\rho(n,n_1), \rho(n_1,a)] \geq \alpha \ .$$

Thus

$$\pi^{\alpha}_{a}(x) \geq \pi^{\alpha}_{n_1}(x) \geq \frac{1}{2}(n_1)^{-1/d}x^{1/d} \geq \frac{1}{4}a^{-1/d}x^{1/d} \ , \tag{13}$$

since $n_1 \leq 2^d a$. ☐

Theorem 3.1 reduces the problem of finding lower bounds for $\rho(n)$ for infinitely many n to that of finding suitable certificates (resp., strong certificates). Such certificates can be searched for by computer, as we describe in Section 4. We note that for any fixed $\alpha > 1/2$ the stochastic models studied in Lagarias and Weiss [7, Theorem 2.2] predict that the number of $n \leq x$ that have $\rho(n) \geq \alpha$ is bounded above by $x^{\gamma(\alpha)+o(1)}$ for a certain exponent $\gamma(\alpha) < 1$; the value $\gamma(\alpha)$ is effectively computable. It follows that strong lower bound certificates for such α, if they could be found, would produce a lower bound qualitatively of the same form x^c expected from the stochastic model predications, but (presumably) with a much smaller exponent than $\gamma(\alpha)$.

4. Computational Results

One may search for lower bound certificates (resp., strong lower bound certificates) by a "greedy" algorithm, as follows.

Certificate Search Algorithm.

(0) *Input Data*: α with $0 < \alpha < 1$.

(1) *Initialization*: $\mathcal{C} = \{\mathbf{c}_{ij} : i \in \{1,2\}, j \in \{0,1,2\}\}$ with $\mathbf{c}_{ij} = (i,j) \in \{0,1,2\}^2$. Declare all vectors $\mathbf{c}_{ij} \in \mathcal{C}$ *open*.

(2) *Tree Test Step*: Select an open vector $\mathbf{c} \in \mathcal{C}$ of minimal length, and let $l+1$ be the length of $\mathbf{c}$. Compute the critical tree$\mathcal{T}^*_{[l]}(\mathbf{c})$ and determine its depth $k = k(l, \mathbf{c})$. If $l/k \geq \alpha$ declare the vector $\mathbf{c}$ *closed*. Otherwise split $\mathbf{c}$ into three vectors of length $l+2$, $(\mathbf{c}, 0)$, $(\mathbf{c}, 1)$ and $(\mathbf{c}, 2)$. Update $\mathcal{C}$ by deleting $\mathbf{c}$ and adding the three new vectors. Declare the three new vectors *open*.

(3) *Termination Test*: If all $\mathbf{c} \in \mathcal{C}$ are closed, halt. Otherwise, repeat the tree test step.

In order to keep the set of open prefixes $\mathcal{C}$ as small as possible during this algorithm, it is convenient to choose at each test step to examine an open prefix vector that is one of those currently of greatest level.

This algorithm is not guaranteed to halt, but if it does, then it produces a lower bound certificate for ratio α. The correctness of the algorithm is based on the observation that $\mathcal{C}$ is a prefix code at all times: the "splitting" step replacing $\mathbf{c}$ by $(\mathbf{c}, 0)$, $(\mathbf{c}, 1)$ and $(\mathbf{c}, 2)$ preserves the prefix code property. It is also easy to see that the algorithm will find a certificate $\mathcal{C}$ if any certificate exists for ones-ratio α, and it will be one of minimal depth.

The algorithm above is easily modified to give a *Strong Certificate Search Algorithm*. It is exactly the same, except that the rule for closing a node is modified: A vector $\mathbf{c} \in \mathcal{C}$ is declared *closed* (only) if $l/k \geq \alpha$ and at least two paths of weight l exist from the root to a leaf of $\mathcal{T}^*_{[l]}(\mathbf{c})$. If the algorithm halts, it produces a strong lower bound certificate for ones-ratio α.

We searched by computer for certificates and strong certificates, for different values of α. The certificate and strong certificate for level $\alpha = 1/3$ presented in Table 1 and Table 2 were found in this way. The size of certificates grows rapidly with increasing α. Table 3 presents data giving the maximal value of α for which

ones-ratio α	size $\lvert\mathcal{C}\rvert$	max-weight l	max-depth k
$1/4 = 0.250$	6	1	4
$2/7 \approx 0.286$	8	2	7
$3/10 = 0.300$	10	3	10
$1/3 \approx 0.333$	12	4	12
$5/14 \approx 0.357$	34	5	14
$5/14 \approx 0.357$	34	6	14
$7/19 \approx 0.368$	68	7	19
$8/21 \approx 0.381$	120	8	21
$9/23 \approx 0.391$	268	9	23
$2/5 = 0.400$	276	10	25
$11/27 \approx 0.408$	704	11	27
$12/29 \approx 0.414$	1522	12	29
$13/31 \approx 0.419$	2404	13	31
$14/33 \approx 0.424$	4758	14	33
$3/7 \approx 0.429$	4782	15	35
$16/37 \approx 0.432$	10646	16	37
$15/34 \approx 0.441$	42336	17	38
$4/9 \approx 0.444$	48718	18	40
$19/42 \approx 0.452$	282326	19	42
$5/11 \approx 0.455$	285098	20	44
$21/46 \approx 0.457$	519802	21	46
$11/24 \approx 0.458$	829044	22	48
$23/50 = 0.460$	1413986	23	50
$7/15 \approx 0.467$	4303530	24	51
$8/17 \approx 0.471$	8035246	25	53
$26/55 \approx 0.473$	16669294	26	55
$9/19 \approx 0.474$	16671812	27	57
$28/59 \approx 0.475$	26948336	28	59
$14/29 \approx 0.483$	350688758	29	60

TABLE 3. Size and depth of certificates

a certificate exists in which all trees have a critical path containing at most l ones, for $1 \le l \le 29$.

Theorem 4.1. *A lower bound certificate $\mathcal{C}$ exists for ones-ratio $\alpha = 14/29 \approx 0.4827$.*

Proof. The certificate was found by computer search, of size indicated in Table 3. The computer search required examining trees to depth 60, with some paths having 29 ones, although the path with the largest ones-ratio occured in a tree of depth 58, having 28 ones. It is an interesting feature of this certificate that the "worst" tree is not one of maximal depth in the search. □

Proof of Theorem 1.1. Since $a = 1$ is in a cycle, we cannot apply Theorem 3.1(i) directly. Instead we consider $n = 41$ and note that $\rho(n) > 55/100$ and $\gamma(n) > 20$. Now we apply Theorem 3.1(i) to $n = 41$ using $\alpha = 14/29$ and the certificate of Theorem 4.1. Then the elements $n \to 41$ produced in Theorem 3.1(i) have $n \to 41 \to 1$ hence

$$\rho(n) \geq \min[\rho(n; 41), \rho(41)] \geq \frac{14}{29},$$

as required. □

The corresponding results for strong certificates are given in Table 4, giving the maximal value of α for which a strong certificate exists having trees with all critical paths having at most l ones. This table is slightly less extensive because larger searches were required. The breakpoint values of α are not the same as for certificates. In many cases they do not involve a larger depth for the largest tree, but sometimes do.

Let $\pi_\alpha(x)$ count the number of n with $1 < n \leq x$ that have finite total stopping time and ones-ratio $\rho(n) \geq \alpha$.

Theorem 4.2. *A strong certificate $\mathcal{C}$ exists for ones-ratio $\alpha = 28/59 \approx 0.4745$. Furthermore there is a positive constant c such that*

$$\pi_\alpha(x) > cx^{1/59}.$$

Proof. The proof is similar to that for Theorem 4.1. The bound for the exponent in $x^{1/59}$ comes from (13). □

The certificates, although very large, can be checked by a much smaller computation than the computation needed to find them. The certificates can be verified without computing the entire tree $\mathcal{T}^*_{[l]}(\mathbf{c})$, by backtracking along the given test path(s). However to find a minimal-depth certificate in the first place it appears necessary to calculate much of the tree $\mathcal{T}^*_{[l]}(\mathbf{c})$. Note that Lagarias and Weiss [7, Theorem 3.1] showed that the expected number of leaves in a pruned $3x+1$ tree of depth k is $\left(\frac{4}{3}\right)^k$, while searching a single path in it involves examining at most k nodes.

ones-ratio α	size $\lvert\mathcal{C}\rvert$	max-weight l	max-depth k
$1/6 \approx 0.167$	6	1	6
$1/4 = 0.250$	14	2	8
$3/10 = 0.300$	26	3	10
$4/13 \approx 0.308$	34	4	13
$1/3 \approx 0.333$	36	5	15
$6/17 \approx 0.353$	98	6	17
$7/19 \approx 0.368$	204	7	19
$8/21 \approx 0.381$	390	8	21
$9/23 \approx 0.391$	848	9	23
$2/5 = 0.400$	914	10	25
$11/27 \approx 0.407$	2242	11	27
$12/29 \approx 0.414$	4720	12	29
$13/31 \approx 0.419$	8020	13	31
$14/33 \approx 0.424$	16044	14	33
$3/7 \approx 0.429$	16182	15	35
$16/37 \approx 0.432$	34264	16	37
$17/39 \approx 0.436$	64960	17	39
$18/41 \approx 0.439$	91170	18	41
$4/9 \approx 0.444$	158182	19	42
$5/11 \approx 0.455$	838262	20	44
$21/46 \approx 0.457$	1475962	21	46
$11/24 \approx 0.458$	2374052	22	48
$23/50 = 0.460$	4114846	23	50
$6/13 \approx 0.462$	4114922	24	52
$23/49 \approx 0.469$	25328092	25	53
$26/55 \approx 0.473$	47636512	26	55
$9/19 \approx 0.474$	47658126	27	57
$28/59 \approx 0.475$	72183824	28	59

TABLE 4. Size and depth of strong certificates

We briefly describe the search method to find the maximal value of α for a given level l, the largest number of ones allowed in a critical path in a certificate. We initialize the search with α taken to be the maximal value at level $l-1$, and with the given certificate. Given a current test value of α, a rational number, we proceed with the algorithm above, looking for a certificate at level l. If one exists, we determine the critical path in the certificate having the largest ones-ratio, which gave a value α', which is our current champion. Then we take as a new test value $\alpha = \alpha' + 0.0001$ and search it to level l for a certificate. If we get a new certificate by level l, we continue. Otherwise, if some tree remains unclosed at level l, we halt the computation, and the current value of α' is maximal at level l. This search procedure works because the maximal value α must be a rational number with denominator no larger than the maximal depth k, which in our search is always less than 100. All such values fall inside the Farey sequence of order 100, and the members of this Farey sequence differ by more than 0.0001.

The straightforward method of computation to determine the critical path for a given $3x + 1$ tree is to compute the entire tree. If the maximal level l to be searched is known in advance (as it is in the incremental search algorithm above) then we can get a speedup by pruning all branches of the tree at the point where they cannot contribute any critical path, i.e. even if extended to the full allowed depth $k = \lfloor l/\alpha \rfloor$ with all edge labels 1 they never have ones-ratio exceeding α.

To conclude the computational results, we give statistics on the structure of the certificates found, as described by the number of open vectors remaining at a given level $l' \leq l$ of the certificate search; this value counts the number of prefixes in the certificate having length at least l'. We present data on these values for different values of α. This data is plotted on a logarithmic scale in Figure 3. When

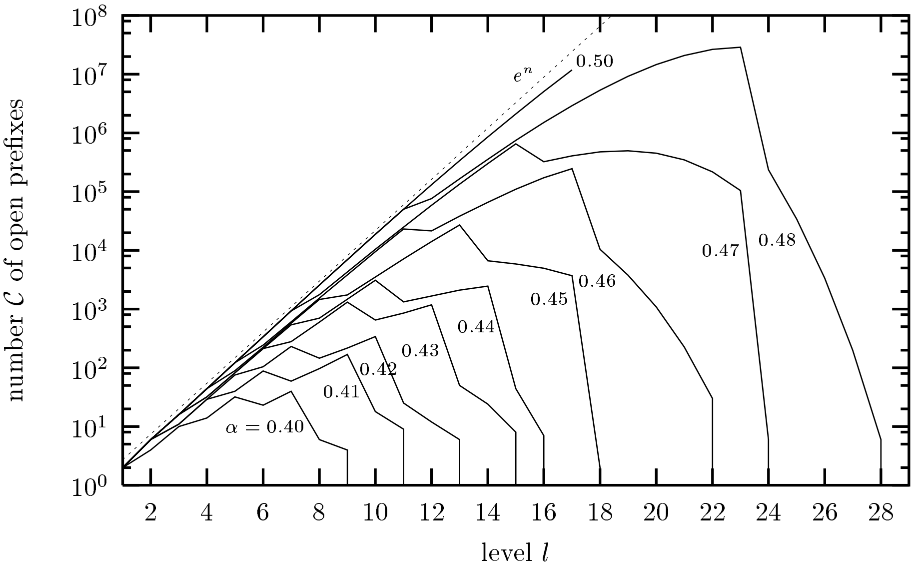


FIGURE 3. Number of open vectors in $\mathcal{C}$ by depth, for various values of α

extrapolated, this data gives some indication of what depth of tree one is likely to have to search to get a certificate for larger values of α than we obtained. The rate of increase of α must slow down as l increases, so we view this extrapolation as indicating a *lower bound* on the depth to which one will have to search. We note, however, that the data in Figure 3 exhibit some irregular jumps in the value of α, including a strange non-increase in the value of α from level $l = 5$ to $l = 6$.

5. Discussion

(1) How large a lower bound can the method of this paper produce?

The stochastic models for $3x+1$ trees (branching random walk models) studied in Lagarias and Weiss [7] predict that for any $\epsilon > 0$ with probability one a random tree grown to a large depth $k \to \infty$ will contain a path $\mathcal{P}$ having $\gamma(\mathcal{P}) > \gamma_{BP} - \epsilon$. However, the method of proof given above requires the lower bound γ_0 to be achieved for *all* trees at some given depth [3], and not just for "almost all" trees of that depth. In terms of the stochastic models of [7] and [3], we are sampling an exponential number of trees (namely, independently drawing 3^k trees at depth k) out of the double-exponential number of possible trees in the branching random walk model at this depth, and demanding that they all have $\gamma(\mathcal{P}) \geq \gamma_0$. In our deterministic situation, if this event ever occurs at some depth k , it yields a rigorous proof of the lower bound. In the analogous stochastic model, we should expect this cutoff event to occur *at some finite depth* k with probability one whenever the set of "exceptional trees" with $\gamma < \gamma_0$ has an exponentially small probability, less than 3^{-k}, for all sufficiently large k. In [7] a large deviations analysis based on exactly this idea produced the constant γ_{BP}. We conclude that the stochastic model supports the heuristic and that this proof method should be expected to work for any $\gamma_0 < \gamma_{BP}$, provided that we can search trees to a sufficient depth, depending on γ_0.

To gain more confidence in this heuristic, it would be interesting to make the stochastic model analysis of [7] more precise. One would like for each γ_0 an estimation of the minimal depth k such that, with probability at least $1/2$, in an

[3]We actually require only that each such tree contain some subtree containing the root vertex, of possibly smaller depth, having a path which achieves the given bound $\gamma \geq \gamma_0$.

independently generated sample of 3^k trees of depth k (for the branching random walk), every tree contains a maximal path with $\gamma(\mathcal{P}) \geq \gamma_0$.

(2) Can a lower bound on the ones-ratio of 1/2 be attained?

It remains surprising (to us) that a search to depth 60 of all trees was insufficient to produce a rigorous proof of the bound $\gamma(n) > \gamma_0$ for infinitely many n, for some

$$\gamma_0 \geq \frac{2}{\log 4/3} \approx 6.95212.$$

Such a bound is believed to apply to the vast majority of initial values n. The statistical evidence provided by the cutoff values for increasing values of the ones-ratio from 0.41 through 0.48, when extrapolated, seem to indicate that a search to max-weight l of at least 38, which corresponds to trees of depth 76, will be needed. We experimented on growing a few trees starting from the certificate with ones-ratio 0.48 to see how deep they had to become before a path with ones-ratio of 0.50 or greater appeared; these seemed to require depth around 68. If this represents the "knee" of the curve for 0.50 in Figure 3, then perhaps the depth necessary for a certificate might be smaller, say 72. In any case this would require a very large computation.

With regard to extending the computations further, we note that the certificate search algorithm of §4 is well suited to parallel and distributed computation, since each open vector can be analyzed separately. Thus it should be possible to search to considerably greater depth using a network of machines. We leave a rigorous proof of the lower bound 1/2 as a challenge to future researchers.

(3) What do these computations suggest about the truth of the $3x+1$ conjecture?

These results are consistent with the nonexistence of divergent trajectories and the truth of the $3x+1$ conjecture. They suggest that all trajectories reach 1 at least as fast as predicted by the stochastic models in Lagarias and Weiss [7], and perhaps even faster.

Acknowledgement. The authors thank the referee for helpful comments.

References

[1] D. Applegate and J. C. Lagarias, *Density bounds for the* $3x+1$ *problem I. Tree-search method*, Math. Comp. **64** (1995), 411–426.

[2] ———, *Density bounds for the* $3x+1$ *problem II. Krasikov inequalities*, Math. Comp. **64** (1995), 427–438.

[3] ———, *On the distribution of* $3x+1$ *trees,* Experimental Math. **4** (1995), 101–117.

[4] K. Borovkov and D. Pfeifer, *Estimates for the Syracuse problem via a probabilistic model*, Theor. Prob. Appl. **45** (2000), 300–310.

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

[6] J. C. Lagarias, *The* $3x+1$ *problem and its generalizations*, Amer. Math. Monthly **92** (1985), 3–21.

[7] J. C. Lagarias and A. Weiss, *The* $3x+1$ *problem: Two stochastic models*, Ann. Applied Prob. **2** (1992), 229–261.

[8] H. Müller, *Das '*$3n+1$*' Problem*, Mitteilungen der Math. Ges. Hamburg **12** (1991), 231–251.

[9] T. Oliveira e Silva, *Maximum excursion and stopping time record-holders for the* $3x+1$ *problem: computational results,* Math. Comp. **68**, No. 1 (1999), 371–384.

[10] D. W. Rawsthorne, *Imitation of an iteration*, Math. Mag. **58** (1985), 172–176.

[11] E. Roosendaal, private communication. See also: On the $3x+1$ problem, electronic manuscript, available at `http://personal.computrain.nl/eric/wondrous`

[12] S. Wagon, *The Collatz problem*, Math. Intelligencer **7** (1985), 72–76.

[13] G. J. Wirsching, *The dynamical system on the natural numbers generated by the* $3n+1$ *function*, Lecture Notes in Math. No. 1681, Springer-Verlag: Berlin 1998.

AT&T Laboratories, Florham Park, NJ 07932-0971

E-mail address, D. Applegate: `david@research.att.com`

E-mail address, J. C. Lagarias: `jcl@research.att.com`