DENSITY BOUNDS FOR THE 3x + 1 PROBLEM.
I. TREE-SEARCH METHOD

DAVID APPLEGATE AND JEFFREY C. LAGARIAS

Dedicated to the memory of D. H. Lehmer

Abstract. The 3x + 1 function T(x) takes the values (3x + 1)/2 if x is odd and x/2 if x is even. Let a be any integer with a £ 0 (mod 3). If n_k(a) counts the number of n with T^k(n) = a, then for all sufficiently large k, (1.302)^k ≤ n_k(a) ≤ (1.359)^k. If π_a(x) counts the number of n with |n| ≤ x which eventually reach a under iteration by T, then for sufficiently large x, π_a(x) ≥ x^{0.5}. The proofs are computer-intensive.

1. Introduction

The 3x + 1 problem concerns the iteration of the function T: Z → Z defined by

(1.1) T(x) = \begin{cases} \frac{3x+1}{2} & \text{if } x \equiv 1 \pmod{2}, \\ \frac{x}{2} & \text{if } x \equiv 0 \pmod{2}. \end{cases}

The 3x + 1 conjecture asserts that, for all n ≥ 1, some iterate T^{(k)}(n) = 1. More generally, it is conjectured that T has finitely many cycles under iteration and that every n ∈ Z eventually enters a cycle (cf. Lagarias [6]). The 3x + 1 conjecture has been verified for all n < 5.6 × 10^{13} by Leavens and Vermeulen [8].

One approach to these conjectures is to study how many integers n below a given bound x have some T^{(k)}(n) = 1. More generally, for any a ∈ Z, set

(1.2) π_a(x) = \#\{n: |n| ≤ x and some T^{(k)}(n) = a, k ≥ 0\}.

It is well known that the growth of π_a(x) depends on the residue class of a (mod 3). If a ≡ 0 (mod 3), then the preimages of a under iterates of T are exactly \{2^k a: k ≥ 1\}; hence π_a(x) grows logarithmically with x. The other cases are covered by the following conjecture.

Conjecture A. For each a ≠ 0 (mod 3), there is a positive constant c_a such that

π_a(x) ≥ c_a x \text{ for all } x ≥ |a|.

In any case, one has, for a ≠ 0 (mod 3),

(1.3) π_a(x) ≥ x^\gamma \text{ for } x ≥ x_0(a),

Received by the editor February 9, 1993 and, in revised form, December 17, 1993.

1991 Mathematics Subject Classification. Primary 11B37, 26A18; Secondary 39B12, 58F08.

©1995 American Mathematical Society
0025-5718/95 $1.00 + $.25 per page

License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use
for some constant \( y > 0 \). This was first shown by Crandall [3], with \( y = .05 \). Crandall's approach directly studies the tree of preimages of \( a \) under \( T \). Sander [9] strengthened Crandall's approach to obtain \( y = .30 \). Krasikov [5] introduced a different method which derives a system of difference inequalities with variables associated to congruence classes \((\mod 3^k)\). Using these inequalities for \( k = 2 \), he obtained \( y = .43 \). Wirsching [10] used Krasikov's inequalities with \( k = 3 \) to obtain \( y = .48 \).

In studying \( \pi_a(x) \), a related problem concerns the size of the tree of preimages of \( a \) under \( T \). Let

\[
n_k(a) := \#\{n: T^{(k)}(n) = a\}.
\]

Lagarias and Weiss [7] proved a result implying that, for \( a \neq 0 \pmod{3} \), the average size of \( n_k(a) \) as \( a \) varies is \( \frac{3}{2} \left(\frac{4}{3}\right)^k \). They conjectured

**Conjecture B.** For each \( a \neq 0 \pmod{3} \),

\[
n_k(a) = \left(\frac{4}{3}\right)^{k(1+o(1))} \quad \text{as } k \to \infty.
\]

For \( a \) not in a cycle, they showed that

\[
\frac{1}{2} (\sqrt{2})^k \leq n_k(a) \leq 2(\sqrt{3})^k,
\]

by studying all possible trees of backward iterates of depth 4.

The object of this paper and its sequel is to obtain improved bounds for \( \pi_a(x) \) and \( n_k(a) \), using computer-assisted proofs. This paper obtains bounds based on the tree-search approach started by Crandall, while the sequel obtains bounds for \( \pi_a(x) \) derived from Krasikov's difference inequalities.

In §2 we study trees \( \mathcal{T}_k^*(a) \) containing all \( n \neq 0 \pmod{3} \) with \( T^{(j)}(n) = a \) for some \( j \leq k \). The structure of this tree depends only on \( a \pmod{3^{k+1}} \). Each leaf \( n \) of the tree is assigned a weight which counts the number of iterates \( T^{(i)}(n) \equiv 1 \pmod{2} \), for \( 0 \leq i \leq k - 1 \). By computer search we find, for all \( k \leq 30 \), upper and lower bound statistics concerning the number of leaves of such trees having a fixed weight. An immediate consequence is

**Theorem 1.1.** For any \( a \neq 0 \pmod{3} \), and for all sufficiently large \( k \),

\[
(1.302053)^k \leq n_k(a) \leq (1.358386)^k.
\]

The proof of Theorem 1.1 is unavoidably computer-intensive; in effect it searches all trees of depth 30.

The upper bound and lower bound statistics for number of leaves lie within a small constant factor of \( \left(\frac{4}{3}\right)^k \). They appear to have a narrower distribution than that predicted by the branching process models for \( 3x + 1 \) trees studied in [7], as we show in detail elsewhere [2].

In §3 we use Chernoff bounds to obtain lower bounds for the number of leaves in such trees having a large weight. Considering trees of depth \( k \), we obtain a bound \( \gamma_k^* \) for the exponent \( \gamma \) in (1.3) by optimizing a "large deviations" bound for the number of heavily weighted leaves in a "worst-case" tree of depth \( k \). Taking \( k = 30 \), we obtain
Theorem 1.2. For each \( a \neq 0 \pmod{3} \), there is a positive constant \( c_a \) such that
\[
\pi_a(x) \geq c_a x^{0.65} \quad \text{for all } x \geq |a|.
\]

This exponent improves on previous bounds; however, in part II we will show that Krasikov's inequalities give still better exponents.

In §3 we also obtain upper bounds for the number of leaves in any tree \( S_k(a) \) that have a large weight. Korec [4] showed for all \( \beta > \beta_c := \frac{\log 3}{\log 4} \) that the set \( \{ n : \text{some } |T^{(k)}(n)| = |n|^{\beta} \} \) has density one. We describe an approach to lower the bound \( \beta_c \) using such upper bound estimates. This approach becomes effective, however, only if a certain threshold is exceeded, and it is not reached by tree depth \( k = 30 \).

2. 3x + 1 TREES

Throughout this section we suppose that \( a \neq 0 \pmod{3} \). The preimages under \( T^{-1} \) of \( a \) form an infinite labelled by tree \( S(a) \), whose root node is labelled \( a \) and whose nodes at the \( k \)th level are labelled by \( \{ n : T^{(k)}(n) = a \} \).

Note that if \( a \) is not in a cycle, then no two nodes of \( S(a) \) have the same label, while if \( a \) is in a cycle then labels will be repeated. The tree \( S(a) \) is constructed recursively using the multivalued operator
\[
T^{-1}(n) = \begin{cases} \{2n\} & \text{if } n \equiv 0, 1 \pmod{3}, \\ \{2n, 2n^2\} & \text{if } n \equiv 2 \pmod{3}. \end{cases}
\]

Each node \( n \) at level \( k \) of the tree is connected to one or two nodes at depth \( k + 1 \) of the tree, which are labelled using the labels in \( T^{-1}(n) \).

In studying asymptotic properties of \( n_k(a) \), it proves convenient to throw out all preimages \( n \equiv 0 \pmod{3} \), and to estimate instead the quantity
\[
n_k^*(a) := \#\{ n : T^{(k)}(n) = a \text{ and } n \equiv 0 \pmod{3} \}.
\]

It is easy to show that
\[
n_k^*(a) \leq n_k(a) \leq k n_k^*(a)
\]
(see Lemma 3.1 of [7]); hence \( n_k^*(a) \) and \( n_k(a) \) have similar exponential growth rates in \( k \) as \( k \to \infty \).

Thus, following [7], we study the smaller tree \( S^*(a) \) resulting by deleting all nodes \( n \equiv 0 \pmod{3} \) from \( S(a) \). The inverse operator \( (T^*)^{-1} \) of \( T \) on the restricted domain \( \{ n : n \neq 0 \pmod{3} \} \) is
\[
(T^*)^{-1}(n) = \begin{cases} \{2n\} & \text{if } n \equiv 1, 4, 5, \text{ or } 7 \pmod{9}, \\ \{2n, 2n^2\} & \text{if } n \equiv 2 \text{ or } 8 \pmod{9}. \end{cases}
\]

Let \( S_k^*(a) \) denote the depth- \( k \) subtree of \( S(a) \); see Figure 2.1 (next page) for \( S_4(4) \) and \( S^*_4(4) \).

We next assign weights to nodes and edges of the tree which keep track of \( 3x + 1 \) iterates (mod 2). An edge connecting \( 2n \) and \( n \) is assigned weight 0, while one connecting \( \frac{2n-1}{3} \) and \( n \) is assigned weight 1. Each node of a tree (except the root) is then assigned weight equal to the sum of the weights of the edges connecting it to the root node. Thus a leaf \( l \) of \( S_k^*(a) \) is assigned
\[
\text{weight}(l) := \#\{ i : T^{(i)}(l) \equiv 1 \pmod{2}, \ 0 \leq i \leq k - 1 \}.
\]
The weight approximately measures the size of the node label, namely,

$$l \leq 3^{-\text{weight}(l)}2^ka.$$  

In addition it can be shown that

$$l = (1 + o(1))3^{-\text{weight}(l)}2^ka \text{ as } k \to \infty,$$

for all those $l$ having weight $(l) \leq \frac{6}{10}k$.

The branching structure of the tree $\mathcal{T}_k(a)$, together with the weights of all its nodes and edges, is completely determined by the congruence class of $a$ (mod $3^{k+1}$); thus the number of distinct tree structures $\mathcal{T}_k^*(a)$ is at most $2 \cdot 3^k$.

We study various statistics concerning the leaves of the trees $\mathcal{T}_k^*(a)$. Let $w_j^k(a)$ count the number of leaves of $\mathcal{T}_k^*(a)$ having weight $j$. This gives the vector of weights

$$w^k(a) := (w_0^k(a), w_1^k(a), \ldots, w_k^k(a)).$$

Now let $N_k^*(a)$ count the number of leaves of $\mathcal{T}_k^*(a)$, and we have

$$N_k^*(a) = w_0^k(a) + w_1^k(a) + \cdots + w_k^k(a).$$

It is obvious that $n_k^*(a) \leq N_k^*(a)$, and equality holds whenever $a$ is not in a cycle of $T$. Theorem 3.1 of [7] showed that the expected size $E[N_k^*(a)]$ averaged over residue classes $a$ (mod $3^{k+1}$) with $a \not\equiv 0$ (mod 3) is

$$E[N_k^*(a)] = \left( \frac{4}{3} \right)^k.$$

We study the quantities

$$N^+(k) := \max\{N_k^*(a): a \pmod{3^{k+1}} \text{ with } a \not\equiv 0 \pmod{3}\},$$

$$N^-(k) := \min\{N_k^*(a): a \pmod{3^{k+1}} \text{ with } a \not\equiv 0 \pmod{3}\}.$$

We also study the majorant vectors $w^+(k)$ and the minorant vectors $w^-(k)$ which we now define. We say that a vector $w = (w_0, \ldots, w_k)$ majorizes a vector $w' = (w'_0, \ldots, w'_k)$ if

$$\sum_{j=0}^i w_{k-j} \geq \sum_{j=0}^i w'_{k-j}, \quad 0 \leq i \leq k,$$
while \( w \) minorizes \( w' \) if
\[
\sum_{j=0}^{i} w_{k-j} \leq \sum_{j=0}^{i} w'_{k-j}, \quad 0 \leq i \leq k.
\]

The majorant vector
\[
w^+(k) := (w_0^+(k), w_1^+(k), \ldots, w_k^+(k))
\]
is the smallest vector majorizing all the \( w^*_k(a) \) and is determined by the conditions
\[
(2.9) \quad \sum_{j=0}^{i} w_{k-j}^+(k) = \max \left\{ \sum_{j=0}^{i} w_{k-j}^k(a) : a \pmod{3^{k+1}} \text{ with } a \neq 0 \pmod{3} \right\},
\]
\[
0 \leq i \leq k.
\]

Similarly, the minorant vector
\[
w^-(k) := (w_0^-(k), w_1^-(k), \ldots, w_k^-(k))
\]
is determined by the conditions
\[
(2.10) \quad \sum_{j=0}^{i} w_{k-j}^-(k) = \min \left\{ \sum_{j=0}^{i} w_{k-j}^k(a) : a \pmod{3^{k+1}} \text{ with } a \neq 0 \pmod{3} \right\},
\]
\[
0 \leq i \leq k.
\]

It is easy to see that these definitions imply that
\[
(2.11a) \quad N^+(k) = \sum_{j=0}^{k} w_j^+(k),
\]
\[
(2.11b) \quad N^-(k) = \sum_{j=0}^{k} w_j^-(k).
\]

In view of (2.8), we have
\[
(2.12) \quad N^-(k) \leq \left( \frac{4}{3} \right)^k \leq N^+(k), \quad k \geq 1.
\]

We computed the vectors \( w^+(k) \) and \( w^-(k) \) for \( 1 \leq k \leq 30 \); the data for \( w^-(k) \) and \( N^-(k) \) appear in Table 2.1 (next page), and that for \( w^+(k) \) and \( N^+(k) \) in Table 2.2 (see p. 417). Details on the computational method are given at the end of the section.

The associated growth rates are
\[
(2.13) \quad g^-(k) = N^-(k)^{1/k}, \quad g^+(k) = N^+(k)^{1/k}.
\]

They are tabulated for \( 1 \leq k \leq 30 \) in Tables 2.1 and 2.2.
Table 2.1. Lower bounds for growth rates

<table>
<thead>
<tr>
<th>$k$</th>
<th>$N^-(k)$</th>
<th>$g^-(k)$</th>
<th>minorizing vector $w^-(k)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>$1.000000$</td>
<td>1 0</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>$1.000000$</td>
<td>1 0 0</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>$1.000000$</td>
<td>1 0 0 0</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>$2.189207$</td>
<td>1 1 0 0 0</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>$1.148698$</td>
<td>1 1 0 0 0</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>$2.00937$</td>
<td>1 2 0 0 0 0</td>
</tr>
<tr>
<td>7</td>
<td>4</td>
<td>$1.219014$</td>
<td>1 2 1 0 0 0 0</td>
</tr>
<tr>
<td>8</td>
<td>5</td>
<td>$1.222845$</td>
<td>1 2 2 0 0 0 0 0</td>
</tr>
<tr>
<td>9</td>
<td>6</td>
<td>$1.220285$</td>
<td>1 2 3 0 0 0 0 0 0</td>
</tr>
<tr>
<td>10</td>
<td>9</td>
<td>$1.245731$</td>
<td>1 3 3 2 0 0 0 0 0 0</td>
</tr>
<tr>
<td>11</td>
<td>11</td>
<td>$1.243575$</td>
<td>1 3 4 3 0 0 0 0 0 0 0</td>
</tr>
<tr>
<td>12</td>
<td>16</td>
<td>$1.259921$</td>
<td>1 4 5 5 1 0 0 0 0 0 0 0</td>
</tr>
<tr>
<td>13</td>
<td>20</td>
<td>$1.259155$</td>
<td>1 4 6 6 3 0 0 0 0 0 0 0</td>
</tr>
<tr>
<td>14</td>
<td>27</td>
<td>$1.265436$</td>
<td>1 4 8 8 5 1 0 0 0 0 0 0 0</td>
</tr>
<tr>
<td>15</td>
<td>36</td>
<td>$1.269553$</td>
<td>1 4 10 10 8 3 0 0 0 0 0 0</td>
</tr>
<tr>
<td>16</td>
<td>48</td>
<td>$1.273731$</td>
<td>1 5 11 13 11 7 0 0 0 0 0 0</td>
</tr>
<tr>
<td>17</td>
<td>64</td>
<td>$1.277162$</td>
<td>1 5 12 18 17 8 3 0 0 0 0 0</td>
</tr>
<tr>
<td>18</td>
<td>87</td>
<td>$1.281596$</td>
<td>1 6 14 23 20 16 7 0 0 0 0 0</td>
</tr>
<tr>
<td>19</td>
<td>114</td>
<td>$1.283093$</td>
<td>1 6 16 27 28 23 11 2 0 0 0 0 0</td>
</tr>
<tr>
<td>20</td>
<td>154</td>
<td>$1.286400$</td>
<td>1 6 18 32 39 29 21 8 0 0 0 0 0 0</td>
</tr>
<tr>
<td>21</td>
<td>206</td>
<td>$1.288796$</td>
<td>1 6 20 38 49 45 31 13 3 0 0 0 0 0 0</td>
</tr>
<tr>
<td>22</td>
<td>274</td>
<td>$1.290645$</td>
<td>1 7 22 45 61 61 43 26 8 0 0 0 0 0 0 0</td>
</tr>
<tr>
<td>23</td>
<td>363</td>
<td>$1.292112$</td>
<td>1 7 24 52 77 81 62 40 17 2 0 0 0 0 0 0 0</td>
</tr>
<tr>
<td>24</td>
<td>484</td>
<td>$1.293864$</td>
<td>1 8 26 60 92 106 91 62 29 9 0 0 0 0 0 0 0 0</td>
</tr>
<tr>
<td>25</td>
<td>649</td>
<td>$1.295656$</td>
<td>1 8 29 69 115 135 127 92 54 17 2 0 0 0 0 0 0 0 0</td>
</tr>
<tr>
<td>26</td>
<td>868</td>
<td>$1.297399$</td>
<td>1 8 32 79 139 175 171 134 83 38 8 0 0 0 0 0 0 0 0 0 0</td>
</tr>
<tr>
<td>27</td>
<td>1159</td>
<td>$1.298627$</td>
<td>1 8 35 89 164 223 232 189 131 63 22 2 0 0 0 0 0 0 0 0 0 0</td>
</tr>
<tr>
<td>28</td>
<td>1549</td>
<td>$1.299961$</td>
<td>1 9 38 100 193 276 307 269 194 108 45 9 0 0 0 0 0 0 0 0 0 0</td>
</tr>
<tr>
<td>29</td>
<td>2052</td>
<td>$1.300807$</td>
<td>1 9 40 113 227 339 401 366 275 171 83 25 2 0 0 0 0 0 0 0 0 0 0</td>
</tr>
<tr>
<td>30</td>
<td>2747</td>
<td>$1.302053$</td>
<td>1 10 43 127 265 418 510 506 402 266 134 56 9 0 0 0 0 0 0 0 0 0 0 0</td>
</tr>
</tbody>
</table>

Theorem 2.1. For any $k \geq 1$, and any $a \not\equiv 0 \pmod{3}$,

\begin{equation}
(2.14) \quad g^-(k) \leq \liminf_{j \to \infty} N^*_j(a)^{1/j} \leq \limsup_{j \to \infty} N^*_j(a)^{1/j} \leq g^+(k).
\end{equation}

In addition,

\begin{equation}
(2.15) \quad g^-(k) \leq \liminf_{j \to \infty} n^*_j(a)^{1/j} \leq \limsup_{j \to \infty} n^*_j(a)^{1/j} \leq g^+(k).
\end{equation}

Proof. Since each tree of depth $j k$ splits into trees of depth $k$ attached to each leaf of the tree of depth $j(k - 1)$, we get by an easy induction

$$N^-(k)^j \leq N^*_j(k) \leq N^+(k)^j.$$

For $0 \leq l \leq k$, we obviously have

$$N^-(k)^j \leq N^*_j(k+l) \leq N^+(k)^{j+1}.$$

Taking $j k$th roots and letting $j \to \infty$ yields (2.14).
To prove the upper bound in (2.15), use
\[ n_j(a) \leq jn^*_j(a) \leq jN^*_j(a), \]
and (2.14). The lower bound in (2.15) is immediate if \( a \) is not in a cycle of \( T \), since \( n^*_j(a) = N^*_j(a) \) in this case. If \( a \) is in a cycle, then the tree \( \mathcal{T}^*(a) \) contains some \( a' \) not in a cycle, say at level \( l \). In that case
\[ n^*_j(a) \geq n^*_{j-l}(a') = N^*_{j-l}(a'), \]
and the lower bound (2.15) follows from the lower bound (2.14) for \( N^*_{j-l}(a') \). \( \Box \)
Theorem 1.1 follows immediately from this result, using the \( k = 30 \) entries of Tables 2.1 and 2.2.

How fast do \( N^+(k) \) and \( N^-(k) \) grow? In order for Conjecture B to be derivable from Theorem 2.1, it is necessary that

\[
\lim_{k \to \infty} g^+(k) = \lim_{k \to \infty} g^-(k) = \frac{4}{3}.
\]

We restate this as the following conjecture.

**Conjecture C.** Both \( N^+(k) \) and \( N^-(k) \) are \((\frac{4}{3})^k(1+o(1))\) as \( k \to \infty \).

This conjecture is stronger than Conjecture B, because it bounds extreme values over all trees of depth \( k \), while Conjecture B applies to the quantities \( n_k(a) \), which as \( k \to \infty \) should behave like "random" trees. To compare the data with this conjecture, we give in Table 2.3 the quantities \((\frac{4}{3})^k\) and the ratios \((\frac{4}{3})^k(N^-(k))^{-1}\) and \((\frac{4}{3})^{-k}N^+(k)\). Formula (2.8) implies that these ratios must both be at least 1, for all \( k \geq 1 \).

**Table 2.3. Normalized extreme values**

<table>
<thead>
<tr>
<th>( k )</th>
<th>( (\frac{4}{3})^k )</th>
<th>((\frac{4}{3})^k(N^-(k))^{-1})</th>
<th>((\frac{4}{3})^{-k}N^+(k))</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.333333</td>
<td>1.333333</td>
<td>1.500000</td>
</tr>
<tr>
<td>2</td>
<td>1.777778</td>
<td>1.777778</td>
<td>1.687500</td>
</tr>
<tr>
<td>3</td>
<td>2.370370</td>
<td>2.370370</td>
<td>1.687500</td>
</tr>
<tr>
<td>4</td>
<td>3.160494</td>
<td>1.580247</td>
<td>1.898438</td>
</tr>
<tr>
<td>5</td>
<td>4.213992</td>
<td>2.106996</td>
<td>1.898438</td>
</tr>
<tr>
<td>6</td>
<td>5.618656</td>
<td>1.872885</td>
<td>1.779785</td>
</tr>
<tr>
<td>7</td>
<td>7.491541</td>
<td>1.872885</td>
<td>1.868774</td>
</tr>
<tr>
<td>8</td>
<td>9.988721</td>
<td>1.997744</td>
<td>1.802032</td>
</tr>
<tr>
<td>9</td>
<td>13.318295</td>
<td>2.219716</td>
<td>1.802032</td>
</tr>
<tr>
<td>10</td>
<td>17.757727</td>
<td>1.973081</td>
<td>1.802032</td>
</tr>
<tr>
<td>11</td>
<td>23.676969</td>
<td>2.152452</td>
<td>1.773876</td>
</tr>
<tr>
<td>12</td>
<td>31.569292</td>
<td>1.973081</td>
<td>1.742199</td>
</tr>
<tr>
<td>13</td>
<td>42.092389</td>
<td>2.106996</td>
<td>1.758038</td>
</tr>
<tr>
<td>14</td>
<td>56.123185</td>
<td>2.078636</td>
<td>1.781795</td>
</tr>
<tr>
<td>15</td>
<td>74.830914</td>
<td>2.078636</td>
<td>1.790704</td>
</tr>
<tr>
<td>16</td>
<td>99.774552</td>
<td>2.078636</td>
<td>1.784022</td>
</tr>
<tr>
<td>17</td>
<td>133.032736</td>
<td>2.078636</td>
<td>1.781516</td>
</tr>
<tr>
<td>18</td>
<td>177.376981</td>
<td>2.038816</td>
<td>1.753328</td>
</tr>
<tr>
<td>19</td>
<td>236.502641</td>
<td>2.074585</td>
<td>1.746281</td>
</tr>
<tr>
<td>20</td>
<td>315.336855</td>
<td>2.047642</td>
<td>1.737824</td>
</tr>
<tr>
<td>21</td>
<td>420.449140</td>
<td>2.041015</td>
<td>1.735050</td>
</tr>
<tr>
<td>22</td>
<td>560.598854</td>
<td>2.045981</td>
<td>1.762401</td>
</tr>
<tr>
<td>23</td>
<td>747.465138</td>
<td>2.059133</td>
<td>1.757942</td>
</tr>
<tr>
<td>24</td>
<td>996.620184</td>
<td>2.059133</td>
<td>1.749914</td>
</tr>
<tr>
<td>25</td>
<td>1328.826912</td>
<td>2.047499</td>
<td>1.737623</td>
</tr>
<tr>
<td>26</td>
<td>1771.769217</td>
<td>2.041209</td>
<td>1.740633</td>
</tr>
<tr>
<td>27</td>
<td>2362.358955</td>
<td>2.038273</td>
<td>1.748253</td>
</tr>
<tr>
<td>28</td>
<td>3149.811941</td>
<td>2.033449</td>
<td>1.746136</td>
</tr>
<tr>
<td>29</td>
<td>4199.749254</td>
<td>2.046614</td>
<td>1.747961</td>
</tr>
<tr>
<td>30</td>
<td>5599.665672</td>
<td>2.038466</td>
<td>1.747961</td>
</tr>
</tbody>
</table>
The data support Conjecture C, and even suggest the following stronger conjecture.

**Conjecture C'**. There are positive constants $C^+$ and $C^-$ such that

$$C^- \left(\frac{4}{3}\right)^k \leq N^-(k) < N^+(k) \leq C^+ \left(\frac{4}{3}\right)^k$$

for all sufficiently large $k$.

Lagarias and Weiss [7] developed branching process models intended to mimic the behavior of $3x + 1$ trees. For the branching process models $\mathcal{B}[3^j]$ with $j \geq 2$ of [7] we prove elsewhere [2] that the analogue of Conjecture C is true, but the analogue of the stronger Conjecture C' is false. That is, $3x + 1$ trees empirically have a narrower variation of leaf counts than that predicted by such stochastic models. This deviation merits an explanation, and we raise this as an open question.

The computation of Tables 2.1 and 2.2 was done as follows. For a given $a \pmod{3^{k+1}}$ with $a \not\equiv 0 \pmod{3}$, let $m_{wk}(a)$ denote the maximum weight of a leaf of the tree $\mathcal{T}_k^*(a)$. Then all trees $\mathcal{T}_k^*(a')$ with $a' \equiv a \pmod{3^{m_{wk}(a)+1}}$ have identical branching structure and node weights. In doing the computation we group all these trees together, specifying them by a single congruence class $a \pmod{3^{k+1}}$, which we call a clone. Let $R_k^f$ count the number of distinct clones of depth $k$ having a maximum weight leaf of weight $l$. The values of $R_k^f$ up to $k = 23$ are given in Table 2.4 (next page).

The quantities $R_k^f$ satisfy the identity

$$\sum_{l=0}^{k} R_k^f 3^{k-l} = 2 \cdot 3^k. \quad (2.16)$$

Let $R(k)$ count the total number of clones of depth $k$,

$$R(k) := \sum_{l=0}^{k} R_l^f. \quad (2.17)$$

Then $R(k)$ counts all possible tree structures of depth $k$ that occur using the $3x + 1$ function. Data on $R(k)$ and on $R(k)^{1/k}$ also appear in Table 2.4. By means of

$$w_k^f(a) = \begin{cases} w_{i-1}^k(2a) & \text{if } a \equiv 1, 4, 5, \text{ or } 7 \pmod{9}, \\ w_{i-1}^k(2a) + w_{i-1}^{k-1}(\frac{2a-1}{3}) & \text{if } a \equiv 2, 8 \pmod{9}, \end{cases}$$

for $0 \leq i \leq k$, and

$$m_{wk}(a) = \begin{cases} m_{wk-1}(2a) & \text{if } a \equiv 1, 4, 5, \text{ or } 7 \pmod{9}, \\ \max\{m_{wk-1}(2a), m_{wk-1}(\frac{2a-1}{3}) + 1\} & \text{if } a \equiv 2 \text{ or } 8 \pmod{9}, \end{cases}$$

all clones of depth $k$ can be identified. In addition, $w_k^*(\cdot)$ and $m_{wk}(\cdot)$ are computed in $O(kR(k))$ operations from a hash table containing $w_{k-1}^*(\cdot)$ and
Table 2.4. \( R^*_k \)-values

<table>
<thead>
<tr>
<th>( k )</th>
<th>( \sum_{l=0}^{k} R^*_l )</th>
<th>( \left( \sum_{l=0}^{k} R^*_l \right)^{1/3} )</th>
<th>( R^*_k )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
<td>4.000000</td>
<td>1.3</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>2.828427</td>
<td>0.53</td>
</tr>
<tr>
<td>3</td>
<td>14</td>
<td>2.414214</td>
<td>0.383</td>
</tr>
<tr>
<td>4</td>
<td>24</td>
<td>2.236068</td>
<td>0.283</td>
</tr>
<tr>
<td>5</td>
<td>42</td>
<td>2.111866</td>
<td>0.182</td>
</tr>
<tr>
<td>6</td>
<td>76</td>
<td>2.058112</td>
<td>0.123</td>
</tr>
<tr>
<td>7</td>
<td>138</td>
<td>2.021608</td>
<td>0.097</td>
</tr>
<tr>
<td>8</td>
<td>254</td>
<td>1.999840</td>
<td>0.066</td>
</tr>
<tr>
<td>9</td>
<td>470</td>
<td>1.981070</td>
<td>0.047</td>
</tr>
<tr>
<td>10</td>
<td>876</td>
<td>1.969021</td>
<td>0.035</td>
</tr>
<tr>
<td>11</td>
<td>1503</td>
<td>1.959794</td>
<td>0.028</td>
</tr>
<tr>
<td>12</td>
<td>3070</td>
<td>1.952517</td>
<td>0.023</td>
</tr>
<tr>
<td>13</td>
<td>5766</td>
<td>1.946696</td>
<td>0.018</td>
</tr>
<tr>
<td>14</td>
<td>10850</td>
<td>1.941981</td>
<td>0.015</td>
</tr>
<tr>
<td>15</td>
<td>20436</td>
<td>1.938026</td>
<td>0.013</td>
</tr>
<tr>
<td>16</td>
<td>38550</td>
<td>1.934757</td>
<td>0.012</td>
</tr>
<tr>
<td>17</td>
<td>72806</td>
<td>1.929201</td>
<td>0.011</td>
</tr>
<tr>
<td>18</td>
<td>137670</td>
<td>1.929706</td>
<td>0.010</td>
</tr>
<tr>
<td>19</td>
<td>260612</td>
<td>1.927757</td>
<td>0.009</td>
</tr>
<tr>
<td>20</td>
<td>493824</td>
<td>1.926099</td>
<td>0.008</td>
</tr>
<tr>
<td>21</td>
<td>936690</td>
<td>1.924694</td>
<td>0.007</td>
</tr>
<tr>
<td>22</td>
<td>1778360</td>
<td>1.923498</td>
<td>0.006</td>
</tr>
<tr>
<td>23</td>
<td>3379372</td>
<td>1.922483</td>
<td>0.006</td>
</tr>
</tbody>
</table>

The quantity \( R(k) \) grows at a somewhat slower exponential growth rate than \( 2 \cdot 3^k \), which makes the computation feasible up to \( k = 30 \). By analogy with a branching process model in Lagarias and Weiss [7], one expects that there is a constant \( \theta \) such that \( R(k) = \Theta(k^{(\theta+o(1))}) \) as \( k \to \infty \), and empirically we estimate \( 1.87 < \theta < 1.92 \). Here, the lower bound 1.87 comes from \( R(k)^{1/(k+1)} \), which is monotonically increasing for \( 8 \leq k \leq 28 \). Observe also that \( R^*_k = 0 \) for small values of \( l \), which occurs because branching of the tree is unavoidable. By analogy with branching process models, one expects that there exists a positive constant \( \phi \) such that \( R^*_l = 0 \) for \( l < (\phi+o(1))k \) and \( R^*_l > 0 \) for \( (\phi+o(1))k \leq l \leq k \), as \( k \to \infty \).
3. Large deviation estimates: lower bounds and upper bounds

We use minorant vectors to get lower bounds for \( r \) in (1.3), as follows. For any constant \( \alpha \in (0, 1] \), set
\[
N^*_j(a; \alpha) := \# \{ l : l \text{ is a leaf in } T_j^*(a) \text{ with weight}(l) \geq \alpha j \}.
\]
By (2.4) all such leaves satisfy the bound
\[
l \leq \exp(j(\log 2 - \alpha \log 3))a.
\]
If we set \( x = \exp(j(\log 2 - \alpha \log 3))a \), and let \( j \to \infty \), then we obtain for any \( \varepsilon > 0 \) that
\[
\pi_a(x) \geq x^{r - \varepsilon}, \quad \forall x \geq x_0(\varepsilon),
\]
where
\[
(3.2) \quad r = \frac{1}{\log 2 - \alpha \log 3} \lim_{j \to \infty} \frac{1}{j} (\log N^*_j(a; \alpha)).
\]

We next use the minorant vector \( w^-(k) \) to obtain an asymptotic lower bound for \( N^*_j(a; \alpha) \). Form a minorizing tree \( T^-_k \) consisting of \( N^-(k) \) leaves of depth one, with exactly \( w^-_i(k) \) of these leaves having edges assigned the weight \( i \), for \( 0 \leq i \leq k \). Now, for all \( j \geq 1 \), recursively construct the concatenated minorizing tree\(^1\) \( T^-_k(j) \) by setting \( T^-_k(1) = T^-_k \) with root node labelled 1 and then forming \( T^-_k(j) \) from \( T^-_k(j - 1) \) by attaching copies of the tree \( T^-_k \) to each leaf of \( T^-_k(j - 1) \). Each leaf of \( T^-_k(j) \) is assigned a weight consisting of the sum of edge weights from it to the root node. Let
\[
(3.3) \quad w^-(k)(^*j) := (x_0^-(j), \ldots, x^k_j(j))
\]
be a vector counting the number of leaves of \( T^-_k(j) \) of weight \( i \), for \( 0 \leq i \leq jk \). (The notation \( w^-(k)(^*j) \) is intended to indicate repeated convolution of \( w^-(k) \), as explained below.) Note also that the number of leaves of \( T^-_k(j) \) is \( N^-(k)^j \). We claim that
\[
(3.4) \quad w^-(k)(^*j) \text{ minorizes } w^-(jk).
\]
To prove the claim, it suffices to show that \( w^-(k)(^*j) \) minorizes each \( w^*_j(a) \). We proceed by induction on \( j \), it being obviously true for \( j = 1 \). Take any tree \( T^j_k(a) \) and view it as a tree \( T^j_{k(j-1)}(a) \) with various trees \( T^j_k(b) \) attached to its leaves. By the induction hypothesis (3.4), the tree \( T^j_k(j - 1) \) can have its leaves paired with those of \( T^j_{k(j-1)}(a) \) in such a way that each leaf of \( T^j_{k(j-1)}(j - 1) \) has a weight no larger than the corresponding leaf of \( T^j_{k(j-1)}(a) \), and \( T^j_{k(j-1)}(a) \) may have some unpaired leaves left over. Then replace \( T^j_{k(j-1)}(a) \) with \( T^j_{k(j-1)}(j - 1) \), and throw away all trees \( T^j_k(b) \) attached to the unpaired nodes; the weight vector of the resulting new tree minorizes that of the old tree \( T^j_k(a) \). Next, in the resulting tree, replace each tree \( T^j_k(b) \) with the tree \( T^-_k \), and the weight vector of the resulting tree minorizes the one before. This final tree is \( T^-_k(j) \), hence we have shown that \( w^-(k)(^*j) \) minorizes \( w^*_j(a) \), and the induction step follows.

\(^1\)The tree \( T^-_k(j) \) has depth \( j \), but we will show that its leaf counts minorize those of any \( 3x + 1 \) tree of depth \( jk \), see (3.4).
Now (3.4) yields the lower bound

\[ N_{jk}(a; \alpha) \geq P_{j,k}^{-}(\alpha) := \sum_{i > jk} x_{i}^{k}(j). \]  

The right side of (3.5) depends only on \( w^{-}(k) \) and can be estimated in a standard fashion (see Lemma 3.1 below). We can then interpolate estimates for \( N_{jk+1}^{*}(a; \alpha) \) using

\[ N_{jk+1}^{*}(a; \alpha) \geq N_{(j+1)k}^{*} \left( a; \alpha + \frac{1}{jk} \right), \quad 0 \leq l \leq k. \]

We estimate the right side of (3.5) using a “large deviations” bound in probability theory. To do this, we assign a value to each node of the tree \( \mathcal{T}_{k}^{-} \). Each leaf of weight \( i \) has value \( \beta = 2^i 3^{j} \). (These values actually approximate the ratio of a leaf label to the root label.) We can use this scheme to recursively assign values to all the nodes of the trees \( \mathcal{T}_{k}^{-}(j) \), starting by assigning the root node the value 1. Next, let \( Z_{k}^{-} \) be a random variable which draws a leaf \( l \) of \( \mathcal{T}_{k}^{-}(1) \) uniformly, and has

\[ Z_{k}^{-} := \log l = k \log 2 - i \log 3. \]

The convolved random variable \( (Z_{k}^{-})^{(*)} \) then equals \( \log l \), where \( l \) is the value of a leaf of \( \mathcal{T}_{k}^{-}(j) \) drawn uniformly. Now, the right side of (3.5) counts exactly those leaves of \( \mathcal{T}_{k}^{-}(j) \) with \( l = 2^j 3^{-i} \leq 2^j 3^{-j \alpha} \); hence

\[ P_{j,k}^{-}(\alpha) = (N^{-}(k))^{j} \text{Prob}[(Z_{k}^{-})^{(*)} < jk(\log 2 - \alpha \log 3)]. \]

The estimation of (3.7) is a standard “large deviations” result.

**Lemma 3.1.** The random variable \( Z = Z_{k}^{-} \) has moment generating function

\[ M_{k}^{-}(\theta) = E[e^{\theta Z}] = \sum_{i=0}^{k} \frac{w_{i}^{-}(k)}{N^{-}(k)} 2^{k \theta} 3^{-i \theta}, \]

whose Legendre transform is

\[ g_{k}^{-}(\beta) := \sup_{\theta \in \mathbb{R}} [\beta \theta - \log M_{k}^{-}(\theta)]. \]

If \( 0 < \log 2 - \alpha \log 3 < \frac{1}{k} E[Z_{k}^{-}] \), then

\[ \lim_{j \to \infty} \frac{1}{jk} (\log P_{j,k}^{-}(\alpha)) = \frac{1}{k} (\log N^{-}(k) - g_{k}^{-}(k(\log 2 - \alpha \log 3))). \]

**Proof.** This is just an application of Chernoff’s theorem (see [7, Lemma 2.1].)

Combining (3.2), (3.5), (3.7), and (3.8) yields the bound

\[ \gamma \geq \frac{\log N^{-}(k) - g_{k}^{-}(k(\log 2 - \alpha \log 3))}{k(\log 2 - \alpha \log 3)}, \]

provided

\[ 0 < \log 2 - \alpha \log 3 \leq \frac{1}{k} E[Z_{k}^{-}] = \frac{1}{k} \sum_{i=0}^{k} \frac{i w_{i}^{-}(k)}{N^{-}(k)} \cdot \]

License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use
For each value of $k$ it remains to optimize the bound (3.9) by choosing the optimal $\alpha = \alpha_k^*$. Data on the expected value $\frac{1}{k}E[Z_k^-]$, the optimal cutoff value $\alpha_k^*$, and the resulting lower bound $\gamma_k^*$ are given in Table 3.1. The quantity $\frac{1}{k}E[Z_k^-]$ is always greater than the expected growth rate of labels on a random branch of a “random” tree $\mathcal{T}_k(a)$, which is $\log 2 - \frac{1}{4}\log 3 = .418494$ (cf. [7, Theorem 3.3]). The data shows that $\frac{1}{k}E[Z_k^-]$ is not a monotonically decreasing function of $k$. Consequently, the estimates $\gamma_k^*$ are also not monotonically increasing, but tend to increase. The largest value we found was $\gamma_{30}^* = .654717$; this proves Theorem 1.2. It is natural to conjecture that $\frac{1}{k}E[Z_k^-] \to \log 2 - \frac{1}{4}\log 3$ and that $\gamma_k^* \to 1$ as $k \to \infty$.

We similarly use majorant vectors $w+(k)$ to get upper bounds on $N^*(a; a)$.

### Table 3.1. Lower bound exponent $\gamma_k^*$ and exceptional set exponent $\chi_k^*$ for upper bound

<table>
<thead>
<tr>
<th>$k$</th>
<th>$\frac{1}{k}E[Z_k^-]$</th>
<th>$\alpha_k^*$</th>
<th>$\gamma_k^*$</th>
<th>$\frac{1}{k}E[Z_k^+]$</th>
<th>$\chi_k^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.693147</td>
<td>0.000000</td>
<td>0.000000</td>
<td>0.143841</td>
<td>0.693147</td>
</tr>
<tr>
<td>2</td>
<td>0.693147</td>
<td>0.000000</td>
<td>0.000000</td>
<td>0.143841</td>
<td>0.549306</td>
</tr>
<tr>
<td>3</td>
<td>0.693147</td>
<td>0.000000</td>
<td>0.000000</td>
<td>0.143841</td>
<td>0.462098</td>
</tr>
<tr>
<td>4</td>
<td>0.555821</td>
<td>0.146657</td>
<td>0.318622</td>
<td>0.189671</td>
<td>0.446038</td>
</tr>
<tr>
<td>5</td>
<td>0.583286</td>
<td>0.113143</td>
<td>0.240657</td>
<td>0.198772</td>
<td>0.413368</td>
</tr>
<tr>
<td>6</td>
<td>0.571079</td>
<td>0.123454</td>
<td>0.324572</td>
<td>0.198772</td>
<td>0.381467</td>
</tr>
<tr>
<td>7</td>
<td>0.536203</td>
<td>0.172320</td>
<td>0.380922</td>
<td>0.233524</td>
<td>0.371315</td>
</tr>
<tr>
<td>8</td>
<td>0.528355</td>
<td>0.177998</td>
<td>0.392520</td>
<td>0.243021</td>
<td>0.354660</td>
</tr>
<tr>
<td>9</td>
<td>0.530390</td>
<td>0.171384</td>
<td>0.385017</td>
<td>0.255737</td>
<td>0.344757</td>
</tr>
<tr>
<td>10</td>
<td>0.510045</td>
<td>0.208448</td>
<td>0.451594</td>
<td>0.274301</td>
<td>0.335695</td>
</tr>
<tr>
<td>11</td>
<td>0.511558</td>
<td>0.200535</td>
<td>0.443546</td>
<td>0.284140</td>
<td>0.327324</td>
</tr>
<tr>
<td>12</td>
<td>0.504323</td>
<td>0.215607</td>
<td>0.481701</td>
<td>0.291987</td>
<td>0.320073</td>
</tr>
<tr>
<td>13</td>
<td>0.498777</td>
<td>0.222555</td>
<td>0.487401</td>
<td>0.307148</td>
<td>0.314558</td>
</tr>
<tr>
<td>14</td>
<td>0.492607</td>
<td>0.233806</td>
<td>0.507598</td>
<td>0.316480</td>
<td>0.310239</td>
</tr>
<tr>
<td>15</td>
<td>0.487666</td>
<td>0.239987</td>
<td>0.521993</td>
<td>0.322570</td>
<td>0.306111</td>
</tr>
<tr>
<td>16</td>
<td>0.485727</td>
<td>0.241974</td>
<td>0.531764</td>
<td>0.327843</td>
<td>0.302016</td>
</tr>
<tr>
<td>17</td>
<td>0.482109</td>
<td>0.250439</td>
<td>0.544585</td>
<td>0.329942</td>
<td>0.298374</td>
</tr>
<tr>
<td>18</td>
<td>0.479177</td>
<td>0.254813</td>
<td>0.558130</td>
<td>0.336952</td>
<td>0.294522</td>
</tr>
<tr>
<td>19</td>
<td>0.477077</td>
<td>0.258955</td>
<td>0.564602</td>
<td>0.342297</td>
<td>0.291267</td>
</tr>
<tr>
<td>20</td>
<td>0.471641</td>
<td>0.268750</td>
<td>0.581699</td>
<td>0.347626</td>
<td>0.288169</td>
</tr>
<tr>
<td>21</td>
<td>0.469158</td>
<td>0.274193</td>
<td>0.591224</td>
<td>0.352461</td>
<td>0.285683</td>
</tr>
<tr>
<td>22</td>
<td>0.467338</td>
<td>0.277591</td>
<td>0.599240</td>
<td>0.356174</td>
<td>0.283427</td>
</tr>
<tr>
<td>23</td>
<td>0.465503</td>
<td>0.281793</td>
<td>0.606344</td>
<td>0.357878</td>
<td>0.281210</td>
</tr>
<tr>
<td>24</td>
<td>0.463229</td>
<td>0.286125</td>
<td>0.614848</td>
<td>0.360225</td>
<td>0.279030</td>
</tr>
<tr>
<td>25</td>
<td>0.461643</td>
<td>0.288560</td>
<td>0.621639</td>
<td>0.362450</td>
<td>0.276928</td>
</tr>
<tr>
<td>26</td>
<td>0.459670</td>
<td>0.292749</td>
<td>0.629568</td>
<td>0.365854</td>
<td>0.275026</td>
</tr>
<tr>
<td>27</td>
<td>0.458350</td>
<td>0.296715</td>
<td>0.636608</td>
<td>0.368401</td>
<td>0.273360</td>
</tr>
<tr>
<td>28</td>
<td>0.456692</td>
<td>0.299908</td>
<td>0.643504</td>
<td>0.369927</td>
<td>0.271726</td>
</tr>
<tr>
<td>29</td>
<td>0.455602</td>
<td>0.303753</td>
<td>0.649338</td>
<td>0.371744</td>
<td>0.270171</td>
</tr>
<tr>
<td>30</td>
<td>0.454481</td>
<td>0.305943</td>
<td>0.655417</td>
<td>0.372635</td>
<td>0.268692</td>
</tr>
</tbody>
</table>
We construct trees \( \mathcal{F}_k^+ \) and \( \mathcal{F}_k^+(j) \) analogously to the lower bound case, using \( w^+(k) \) instead of \( w^-(k) \). The vector
\[
*^+(k)^{(*)} := (y_0^+(j), \ldots, y_j^k(j))
\]
enumerates the number of leaves in the tree \( \mathcal{F}_k^+(j) \) of different weights. We then show, analogously to the lower bound case, that
\[
(3.10) \quad w^+(k)^{(*)} \text{ majorizes } w^+(jk),
\]
from which we conclude
\[
(3.11) \quad N^*_k(a; \alpha) \leq P^*_{j,k}(\alpha) := \sum_{i > jk \alpha} y_i^k(j).
\]
The right side of (3.11) is estimated by a Chernoff inequality argument. Let \( Z_k^+ \) be a random variable which draws a leaf \( l \) from \( \mathcal{F}_k^+(1) \) uniformly and assigns the value \( \log(l) \), similarly to (3.6). The convolution \( (Z_k^+)^{(*)} \) then describes the value \( \log(l) \) for a random leaf of \( \mathcal{F}_k^+(j) \), and we have
\[
P^*_{j,k}(\alpha) = (N^+(k))^j \text{Prob}[(Z_k^+)^{(*)} < jk(\log 2 - \alpha \log 3)].
\]
The Chernoff bound formula is analogous to Lemma 3.1.

**Lemma 3.2.** The random variable \( Z = Z_k^+ \) has moment generating function
\[
M_k^+(\theta) = \sum_{i=0}^{k} \frac{w_i^+(k)}{N^+(k)} 2^{k\theta} 3^{-i\theta},
\]
whose Legendre transform is
\[
g_k^+(\beta) := \sup_{\theta \in \mathbb{R}} [\beta \theta - \log M_k^+(\theta)].
\]
If \( \log 2 - \beta \log 3 > \frac{1}{k} E[Z_k^+] \), then
\[
(3.12) \quad \lim_{j \to \infty} \frac{1}{jk} (\log P^*_{j,k}(\alpha)) = \frac{1}{k} (\log N^+(k) - g_k^+(k(\log 2 - \alpha \log 3))).
\]

Table 3.1 presents data on \( \frac{1}{k} E[Z_k^+] \). It is always less than the expected growth rate \( \log 2 - \frac{1}{4} \log 3 \approx .418494 \) of labels on a random branch of a “random” tree \( \mathcal{F}_k(a) \). Empirically, it appears to be a monotone function of \( k \), unlike the lower bound case. It is natural to conjecture that \( \frac{1}{k} E[Z_k^+] \to \log 2 - \frac{1}{4} \log 3 \) as \( k \to \infty \).

Upper bound estimates for \( N^*_k(a; \alpha) \) are also relevant to proving results saying that “almost all” integers decrease under iteration by \( T \). Currently the best quantitative result of this kind is that of Korec [4].

**Theorem 3.1** (Korec). For any \( \beta > \beta_c := \frac{\log 3}{\log 4} \approx .7925 \) the set
\[
S(\beta) := \{n: \text{some } |T^{(k)}(n)| < |n|^\beta\}
\]
has density one.

Korec’s method actually shows that almost all \( \{n: |n| \leq x\} \) satisfy
\[
(3.13) \quad |T^{(k)}(n)| \leq x^\beta, \quad \text{for } k = \left[ \frac{\log x}{\log 2} \right],
\]
as \( x \to \infty \), for any fixed \( \beta > \beta_c \).
We show below that one can get improved bounds for $\beta_c$ in Theorem 3.1 provided that the quantity

$$\chi_k := \frac{1}{k} \log \frac{N^+(k) - g_k^+(k(\log 2 - 1/2(\log 3)))}{\log x}$$

is sufficiently small. This quantity is the upper bound (3.12) with $\alpha = 1/2$, and its values are given in Table 3.1.

Consider the set of "bad elements"

$$R_\delta(x) := \left\{ n: |n| < x \text{ and no } T^{(j)}(x) < x^{1-\delta} \text{ for } 1 \leq j < \left\lceil \frac{\log x}{\log 2} \right\rceil \right\}.$$ 

The cardinality of $R_\delta(x)$ decreases as $\delta \to 0$ and

$$\lim_{\delta \to 0} \frac{\log |R_\delta(x)|}{\log x} = H \left( \log \frac{\log 2}{\log 3} \right) \approx .94995,$$

where $H(t) = -t \log_2 t - (1-t) \log_2 (1-t)$ is the binary entropy function (cf. [6, Theorem D]). Almost all $\{n: |n| \leq x\}$ satisfy (3.13). We can get an improvement if almost all such $T^{(k)}(n)$ with $k = \left\lceil \frac{\log x}{\log 2} \right\rceil$ do not lie in a "bad element" set $R_\delta(x^\beta)$, for some fixed $\delta > 0$. How many such $n$ can hit a particular "bad" element $y$? They must lie in the tree of preimages of $y$, at height $j = \left\lceil \frac{\log x}{\log 2} \right\rceil$, so we need an upper bound for the number of leaves $l$ in such a tree, at this height, having $y \approx x^\beta$ and $l \leq x$. Such leaves correspond to paths having $\alpha \geq \frac{1}{2}$ as explained in [7, §2], and we can apply the upper bounds (3.11)–(3.13) to bound the number of such leaves by $\exp(\chi_k \log \frac{\log x}{\log 4})$. Now the number of such "bad elements" as $\beta \to \beta_c$ and $\delta \to 0$ satisfies

$$\log |R_\delta(x^\beta)| = \left( .94995 \frac{\log 3}{\log 4} + o(1) \right) \log x;$$

hence the number of preimages $n \leq x$ which these generate is at most

$$\exp \left( .94995 \frac{\log 3}{\log 4} + \frac{\chi_k}{\log 2} \right) \log x.$$ 

This bound will be $O(x^{1-\epsilon'})$ for some $\epsilon' > 0$, if and only if

$$\chi_k < \log 2 - \frac{1}{2} H \left( \log \frac{\log 2}{\log 3} \right) \log 3 \approx .171331.$$

As the data of Table 3.1 show, however, for $k \leq 30$ we are a long way from attaining the bound (3.16).

The assumption that $3x + 1$ trees behave like the branching process models of [7] leads to the heuristic prediction that $\chi_k \to 0$ as $k \to \infty$. If so, this approach to lowering $\beta_c$ should eventually work, for $k$ large enough. The data of Table 3.1 indicate that the smallest $k$ for which (3.16) holds will be so large that it will be impossible to compute by an exhaustive tree search.

ACKNOWLEDGMENT

We are indebted to T. H. Foregger for a critical reading and to a referee for helpful comments.

---

2To get a rigorous bound, one must also count a few extra leaves having $\alpha < \frac{1}{2}$, which creep in because $T^{-1}$ has $\frac{2x-1}{x}$ instead of $\frac{2x}{x}$. However a rigorous variant of (2.5) can be used to show that these leaves make an asymptotically negligible contribution.
BIBLIOGRAPHY


E-mail address, D. Applegate: david@research.att.com
E-mail address, J. C. Lagarias: jcl@research.att.com