ON THE DISTRIBUTION OF $3x+1$ TREES

DAVID APPLEGATE AND JEFFREY C. LAGARIAS

Abstract. The $3x+1$ iteration when run backwards from any fixed integer a produces a tree of preimages of a. Let $\mathcal{T}_k(a)$ denote this tree grown to depth k, and let $\mathcal{T}_k^*(a)$ denote the pruned tree resulting from the removal of all nodes $n \equiv 0 \pmod{3}$. The maximal and minimal number of leaves in the pruned trees $\mathcal{T}_k^*(a)$ grown from all possible root nodes $a \not\equiv 0 \pmod{3}$ were computed for all tree depths k up to depth 30. We compare this data with predictions made using branching process models designed to imitate the growth of $3x+1$ trees, developed in [10]. Rigorous results are derived for the branching process models. The range of variation exhibited by the $3x+1$ trees is significantly narrower than that of the branching process models. We also study the variation in expected leaf-counts associated to the congruence class of $a \pmod{3^j}$. This variation, when properly normalized, converges almost everywhere as $j \to \infty$ to a limit function on the invertible 3-adic integers.

1. Introduction

The well-known $3x+1$ problem concerns the behavior under iteration of the $3x+1$ function $T : \mathbf{Z} \mapsto \mathbf{Z}$ 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 $3x+1$ Conjecture asserts that for each $n \geq 1$, some iterate $T^{(k)}(n) = 1$; it has now been verified for all $n < 5.6 \times 10^{13}$, see Leavens and Vermeulen [11]. 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.

The $3x+1$ function is a deterministic process that apparently exhibits pseudo-random behavior. It has been extensively studied, see the surveys of Lagarias [9] and Müller [12]. One approach to quantifying its apparent pseudorandomness is

Date: July 21, 1995.

to consider probabilistic models for its behavior on a "random" input, and then to compare model predictions with empirical data. Any systematic discrepancies or similarities uncovered may prove helpful in eventually proving rigorous results.

We now review several probabilistic models for the $3x + 1$ iteration. Consider taking input values n drawn from the uniform distribution U_{2^k} on $[1, 2^k]$, and examine the induced probability distribution on $T^{(j)}(n)$, for $1 \le j \le [\alpha k]$, for a fixed positive α. One can rigorously prove that when $0 < \alpha \le 1$, the successive iterates $(\log \frac{T(n)}{n}, \ldots, \log \frac{T^{([\alpha k])}(n)}{n})$ behave exactly like the trajectory of a random walk which takes i.i.d. steps of size $\log \frac{3}{2}$ or $\log \frac{1}{2}$ with equal probability; see [9], § 2. This result suggests that the evolution of $3x + 1$ function iterates can be modelled by a multiplicative random walk, in which from an initial point X_0 one multiplies by successive i.i.d. random variables X_i which take the values $\frac{3}{2}$ and $\frac{1}{2}$ with probability $\frac{1}{2}$ each, to obtain $Y_j := X_0 X_1 \cdots X_j$. Such a model was first considered in Crandall [6], and in more detail in Rawsthorne [13] and Wagon [14]. The analogue in this model of $\sigma_\infty(X_0)$ is the statistic $\sigma_\infty(X_0, \omega)$ which for a random walk ω starting from X_0 gives the smallest value of J such that $Y_J < 1$. For this model the expected value

$$E[\sigma_\infty(X_0, \omega)] = \left(\frac{1}{2} \log \frac{4}{3}\right)^{-1} \log(X_0) \quad . \tag{2}$$

Recently Borovkov and Pfeifer [5] gave a refined analysis showing that $\sigma_\infty(X_0, \omega)$ obeys a central limit theorem, i.e. the scaled variables

$$\hat{\sigma}_\infty(X_0, \omega) := \frac{\sigma_\infty(X_0, \omega) - c_1 \log X_0}{c_2 (\log X_0)^{1/2}} \quad , \tag{3}$$

in which $c_1 = (\frac{1}{2} \log \frac{4}{3})^{-1}$ and $c_2 = {c_1}^{3/2} (\frac{1}{2} \log 3)$, have distribution converging to the unit normal distribution $N(0, 1)$ as $X_0 \to \infty$. Although this model with n_0 drawn from U_{2^k} is rigorously proved to approximate the distribution of $T^{(\alpha k)}(n_0)$ only for $\alpha \le 1$, empirically it is found that the approximation seems good all the way up to $\alpha = (\frac{1}{2} \log \frac{4}{3})^{-1} \doteq 6.95212$. Furthermore the agreement with the central limit approximation (3) is also reasonably good. Thus this random walk model appears to accurately describe "average" trajectories of $3x + 1$ iterates.

Lagarias and Weiss [10] present two types of probabilistic models intended to simulate "extreme" trajectories of $3x+1$ iterates, i.e. those attaining the largest value of the quantity $\frac{\sigma_\infty(n)}{\log n}$ for all $n \in [1, 2^k]$. The first of these models is a repeated multiplicative random walk model[1], which takes 2^k entirely independent multiplicative random walks as above, with the n-th such walk ω_n starting from $X_0 = n$. An analogous model statistic γ_k to consider is the maximum value of $\frac{\sigma_\infty(n,\omega_n)}{\log n}$ taken over $1 \le n \le 2^k$. For this model, they showed that with probability one the values γ_k tend to a limit γ_{RW} as $k \to \infty$, i.e. with probability one

$$\limsup_{n\to\infty} \frac{\sigma_\infty(n,\omega_n)}{\log n} = \gamma_{\mathrm{RW}} \quad . \tag{4}$$

Here $\gamma_{\mathrm{RW}} \doteq 41.677647$ is the solution of a certain transcendental equation. This model has the deficiency that it assumes independence of trajectories for different starting values n_0 and n_1. This is not true of $3x+1$ trajectories — they must coalesce, since (empirically) all trajectories reach 1.

The second type of stochastic model of [10] is a branching process model that mimics backwards iteration of the $3x+1$ function, and which explicitly includes dependencies among trajectories. Backwards iteration of the $3x+1$ map is multiple-valued, and given an initial value a it produces a tree $\mathcal{T}(a)$ of preimages of a. The branching process models construct "random" trees whose structures imitate the structure of a $3x+1$ tree grown from a "random" starting point a. Lagarias and Weiss present an infinite family $\mathcal{B}[3^j], j = 0, 1, 2, \dots$ of successively more refined branching process models. For these branching process models they proved that an analogue of the asymptotically largest value of $\frac{\sigma_\infty(n)}{\log n}$ as $n \to \infty$ is almost surely a constant γ_{BP}, and $\gamma_{\mathrm{BP}} = \gamma_{\mathrm{RW}} \doteq 41.677647$. Finally they observed that the existing empirical data for extremal trajectories of the $3x+1$ function, computed up to 5.6×10^{13} in [11], is consistent with the predictions made by these two types of models.

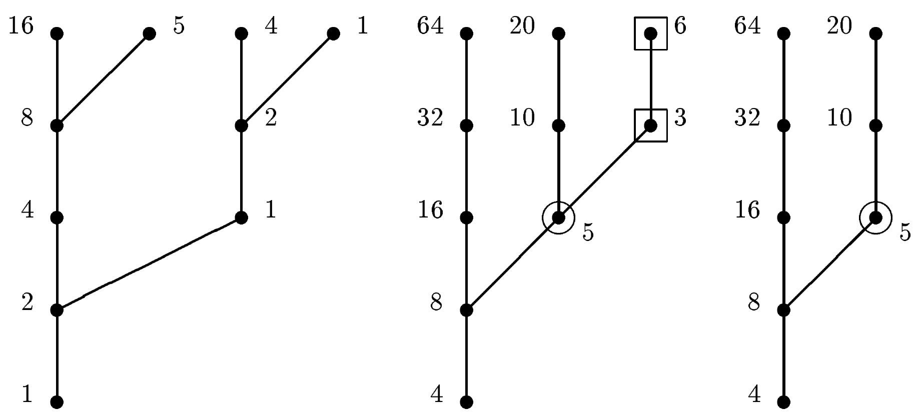
This paper studes extremal properties of ensembles of $3x+1$ trees of depth k. A $3x+1$ tree $\mathcal{T}_k(a)$ is a rooted, labelled tree of depth k, representing the inverse

[1]The model in [10] actually uses additive random walks, and is obtained from the above by taking logarithms.

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 level l of the tree is connected by an edge to a node labelled $T(n)$ at level $l-1$ of the tree[2]. As described in [9], 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 a "pruned" tree $\mathcal{T}_k^*(n)$ consisting of nodes $n \not\equiv 0 \pmod 3$. 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.)



(i) $\mathcal{T}_5(1)$ (ii) $\mathcal{T}_5(4)$ (iii) $\mathcal{T}_5^*(4)$

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

We say 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 node labels modulo 2. Since the node label $n \pmod 2$ determines whether a branch to n comes from a node labelled $\frac{n}{2}$ or $\frac{3n+1}{2}$, the congruence classes $n \pmod 3$ and $T(n) \pmod 9$ suffice to determine $n \pmod 2$. From this, it easily follows that the structure of

[2]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.

$\mathcal{T}_k^*(a)$ is completely determined by $a \pmod{3^{k+1}}$. Consequently there are at most $2 \cdot 3^k$ distinct pruned tree structures $\mathcal{T}_k^*(a)$. The actual number $R(k)$ of distinct tree structures is smaller but still grows exponentially.

We study the extreme (maximum and minimum) leaf counts $N^+(k)$ and $N^-(k)$ for the ensemble of all such trees of depth k. In § 2 we present empirical data for all $k \le 30$, which was computed in [1]. This data suggests two conjectures concerning the asymptotic behavior of the extreme leaf counts as $k \to \infty$, which we call Conjecture C and the (stronger) Conjecture $\mathrm{C}^{\#}$.

We now ask: to what degree do repeated trials of the branching process models of [10] reproduce this empirical $3x+1$ data? We first note that only the models $\mathcal{B}[3^j]$ for $j \ge 2$ can be reasonable models. The models $\mathcal{B}[1]$ and $\mathcal{B}[3]$ were already shown in § 6 of [10] to fail to assign the correct distribution of residue classes $\pmod 3$ to the node labels. Besides this, and more important, the models $\mathcal{B}[1]$ and $\mathcal{B}[3]$ do not possess the following "strict branching" property of pruned $3x+1$ trees: every pruned $3x+1$ tree branches after at most 4 steps from any node. The models $\mathcal{B}[1]$ and $\mathcal{B}[3]$ can produce trees having arbitrarily long chains of nodes with no branching.

As far as one can tell, all the branching process models $\mathcal{B}[3^j]$ for $j \ge 2$ provide reasonable imitations of the $3x+1$ trees. Therefore in § 3 we study the simplest of these models, which is $\mathcal{B}[9]$. We present data for $k \le 30$ on the expected value of extreme leaf counts for a "repeated branching process" model which takes $R(k)$ independent trials using the branching process $B[9]$. (Recall that $R(k)$ is the number of distinct tree structures of depth k.) These expected values for $k \le 30$ appear consistent with Conjecture C, but exhibit larger variability than that empirically observed for the $3x+1$ data up to $k \le 30$.

In § 4 and § 5 we present theoretical results about branching process models. First, in § 4 we prove a result establishing for a large class of branching processes that there is a double-exponential dropoff of tail probabilities for values of $\log N(k)$, where $N(k)$ counts the number of leaves at depth k of the process. Such results are "folklore," and we are indebted to Roben Pemantle for suggesting the method used to prove Theorem 4.1. Then, in § 5 we prove that the analogue of Conjecture C is true for a repeated branching process model using $\mathcal{B}[9]$. We finally prove that the analogue of Conjecture $C^{\#}$ is false for this repeated branching process model.

Thus we have uncovered a difference between the $3x+1$ empirical data and the branching process model: the extreme leaf count statistics for the actual $3x+1$ problem appear to have a significantly narrower range than that given by the branching process models. This seems to be the first evidence found indicating that the $3x+1$ function iterates do not behave as randomly as possible subject to "obvious" constraints.

In § 6 we return to the study of extremal leaf counts by studying the average number of leaves in pruned trees $\mathcal{T}_k^*(a)$, under the restriction $a \equiv l \pmod{3^j}$, in which $l \not\equiv 0 \pmod 3$. This amounts to specifying the branching structure of the first j levels of the tree $\mathcal{T}_k^*(a)$. We prove that this expected value is asymptotic to $W[l \bmod 3^j](\frac{4}{3})^k$ as $k \to \infty$, where $W[l \bmod 3^j]$ is an explicitly computable value (Theorem 6.1). The variation in $W[l \bmod 3^j]$ appears to account for nearly all of the variation in leaf sizes, and we conjecture that:

$$\begin{aligned} \limsup_{k\to\infty} \left(\frac{4}{3}\right)^{-k} N^+(k) &= \sup_{l,j} W[l \bmod 3^j] \ . \\ \liminf_{k\to\infty} \left(\frac{4}{3}\right)^{-k} N^-(k) &= \inf_{l,j} W[l \bmod 3^j] \ . \end{aligned}$$

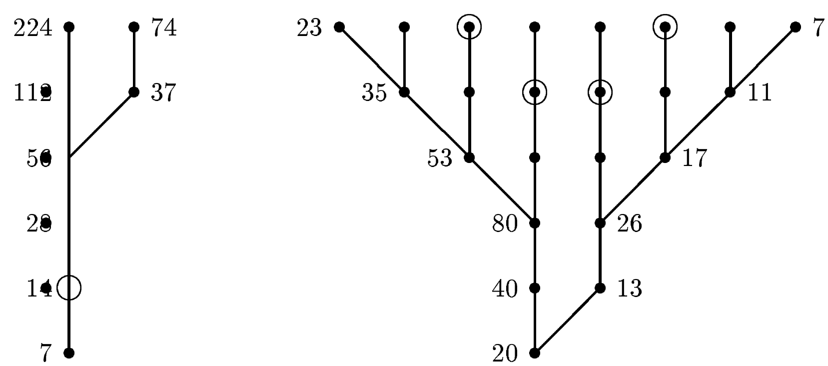
We show that $W[l \bmod 3^j]$ interpolates to a function $W_\infty(l)$ defined almost everywhere on the invertible 3-adic integers $\mathbf{Z}_3^\times = \{l \in \mathbf{Z}_3 : l \equiv 1 \text{ or } 2 \pmod 3\}$ (Theorem 6.2). We conjecture that $W_\infty(l)$ is well-defined on all of $\mathbf{Z}_3^\times$ and is continuous and nonzero. Numerical evidence concerning $W[l \bmod 3^j]$ seems to support the truth of Conjecture $C^{\#}$.

Finally we remark that G. Wirsching [15], [16] has recently introduced other functions on $\mathbf{Z}_3^\times$ associated to backwards iteration of the $3x+1$ mapping. We do not know the relation of these functions, if any, to the function W_∞.

2. $3x+1$ Trees

In studying $3x+1$ trees we follow Applegate and Lagarias [1]. Assign to each $a \not\equiv 0 \pmod 3$ the pruned tree $\mathcal{T}_k^*(a)$ of depth k whose root node is labelled a

and whose other vertices at depth j for $1 \le j \le k$ correspond to labels in the set $\{n : n \not\equiv 0 \pmod 3 \text{ and } T^{(j)}(n) = a\}$. Each node labelled n at level j is connected to that labelled $T(n)$ at level $j-1$, see Figure 2. (The nodes in Figure 2 that have labels $n \equiv 5 \pmod 9$ are circled; such nodes have a preimage $n' = \frac{2n-1}{3} \equiv 0 \pmod 3$ in the unpruned tree $\mathcal{T}_k(a)$.) The branching structure of the pruned tree $\mathcal{T}_k^*(a)$ is completely determined by the value $a \pmod{3^{k+1}}$.



(i) $\mathcal{T}_5^*(7)$ attains $N^-(5) = 2$ (ii) $\mathcal{T}_5^*(20)$ attains $N^+(5) = 8$

FIGURE 2. Pruned $3x+1$ Trees

Let $N_k^*(a)$ denote the number of leaves at depth k of $\mathcal{T}_k^*(a)$ and set

$$N^-(k) := \min\{N_k^*(a) : a \pmod{3^{k+1}} \text{ and } a \not\equiv 0 \pmod 3\}\ , \tag{5}$$

$$N^+(k) := \max\{N_k^*(a) : a \pmod{3^{k+1}} \text{ and } a \not\equiv 0 \pmod 3\}\ . \tag{6}$$

Theorem 3.1 of [10] showed that the expected size of $N_k^*(a)$ averaged over all $a \pmod{3^{k+1}}$ with $a \not\equiv 0 \pmod 3$ is

$$E[N_k^*(a)] = \left(\frac{4}{3}\right)^k\ . \tag{7}$$

In Applegate and Lagarias [1] we proposed:

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

To test such a conjecture it is natural to examine the normalized densities

$$D^+(k) := (\frac{4}{3})^{-k} N^+(k),$$
$$D^-(k) := (\frac{4}{3})^{-k} N^-(k),$$

which must necessarily satisfy $0 < D^-(k) \le 1 \le D^+(k)$ by (7). Table 1 below gives empirical data for $k \le 30$ using the data from Applegate and Lagarias [1], § 2. This data supports Conjecture C, and also appears to support the following stronger conjecture.

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

$$C^- \le D^-(k) < 1 < D^+(k) < C^+$$

for all sufficiently large k.

It even seems conceivable that $D^-(k)$ and $D^+(k)$ have limiting values as $k \to \infty$. In § 6 we give further evidence which seems to support Conjecture C$^\#$and the existence of limiting densities as $k \to \infty$.

3. Branching Process Models for $3x + 1$ Trees

We consider the question: To what extent do the branching process models $\mathcal{B}[3^j]$ for $j \ge 2$ presented in [10] accurately imitate the behavior of $3x+1$ trees? These models are multi-type Galton-Watson processes, for which see [3] and [7]. Recall that such a process describes the evolution of a population of individuals of several types over generations, where each individual lives one generation. Each individual independently gives rise to progeny in the next generation of several types according to a specified probability distribution. The branching process tree describes the descendents of a single individual at generation 0, and level l of the tree includes all individuals in generation l. Edges connect individuals to their progeny in the next generation. Such a process is completely described by the probability distribution of individuals of each type.

The multi-type Galton-Watson branching process $\mathcal{B}[9]$ has individuals of six types, labelled with congruence classes 1,2,4,5,7 and 8 $\pmod 9$, and these evolve as pictured in Figure 3. Individuals labelled 1,4,5 and 7 evolve deterministically, having one child of specified type, while individuals of type 2 or 8 always have two children, one of specified type, while the other's type is specified with probability 1/3 each. Figure 3 also indicates edge labels reflecting whether $T^{-1}(n)$ is $2n$ or $(2n-1)/3$, e.g. whether $T^{-1}(n)$ is even or odd. The edge labels are completely

k	# tree types $R(k)$	$N^-(k)$	$N^+(k)$	$\left(\frac{4}{3}\right)^k$	$D^-(k)$	$D^+(k)$
1	4	1	2	1.33	0.750	1.500
2	8	1	3	1.78	0.562	1.688
3	14	1	4	2.37	0.422	1.688
4	24	2	6	3.16	0.633	1.898
5	42	2	8	4.21	0.475	1.898
6	76	3	10	5.62	0.534	1.780
7	138	4	14	7.49	0.534	1.869
8	254	5	18	9.99	0.501	1.802
9	470	6	24	13.32	0.451	1.802
10	876	9	32	17.76	0.507	1.802
11	1638	11	42	23.68	0.465	1.774
12	3070	16	55	31.57	0.507	1.742
13	5766	20	74	42.09	0.475	1.758
14	10850	27	100	56.12	0.481	1.782
15	20436	36	134	74.83	0.481	1.791
16	38550	48	178	99.77	0.481	1.784
17	72806	64	237	133.03	0.481	1.782
18	137670	87	311	177.38	0.490	1.753
19	260612	114	413	236.50	0.482	1.746
20	493824	154	548	315.34	0.488	1.738
21	936690	206	736	420.45	0.490	1.751
22	1778360	274	988	560.60	0.489	1.762
23	3379372	363	1314	747.47	0.486	1.758
24	6427190	484	1744	996.62	0.486	1.750
25	12232928	649	2309	1328.83	0.488	1.738
26	23300652	868	3084	1771.77	0.490	1.741
27	44414366	1159	4130	2362.36	0.491	1.748
28	84713872	1549	5500	3149.81	0.492	1.746
29	161686324	2052	7336	4199.75	0.489	1.747
30	308780220	2747	9788	5599.67	0.491	1.748

TABLE 1. Normalized Extreme Values for $3x+1$ Trees of depth k

determined by the types of the individuals at the two ends of the edge, hence are determined by the Galton-Watson process.

The model $\mathcal{B}[9]$ permits an unambiguous assignment of node labels to all nodes of a branching process tree provided that a root node label is given. If n is a node label at level l, and n' is a node it is connected to at level $l+1$, then we assign $n' = 2n$ or $\frac{2}{3}n$ according as the edge connecting n to n' is labelled even or odd. The Galton-Watson process with the node labels added and interpreted as locations of the individuals on the line $\mathbf{R}$, becomes a branching random walk, which is the term used for these models in [10]. The node labels are needed in [10] to view the branching process as imitating the growth of $3x+1$ iterates, but they play no role in this paper.

Now let X_k be a random variable equal to the number of leaves at depth k of a sample tree drawn from the branching process $\mathcal{B}[9]$, starting from a single individual of type drawn uniformly from $\{1, 2, 4, 5, 7, 8\}$. We consider extreme value statistics for the quantity $(\frac{4}{3})^{-k} X_k$ for a specified number of repeated independent draws of such trees at depth k.

How many independent draws should one allow in such a "repeated branching process" model? The naive model is to take $2 \cdot 3^k$ draws, corresponding to allowing all residue classes $a \pmod{3^{k+1}}$ with $a \not\equiv 0 \pmod 3$. An alternative is to take instead the smaller number $R(k)$ of different distinct $3x+1$ tree structures $\mathcal{T}^*(a)$ of depth k that are possible. The quantities $R(k)$ still grow exponentially in k, and based on the data for $k \le 30$, Applegate and Lagarias [1] estimated (empirically)

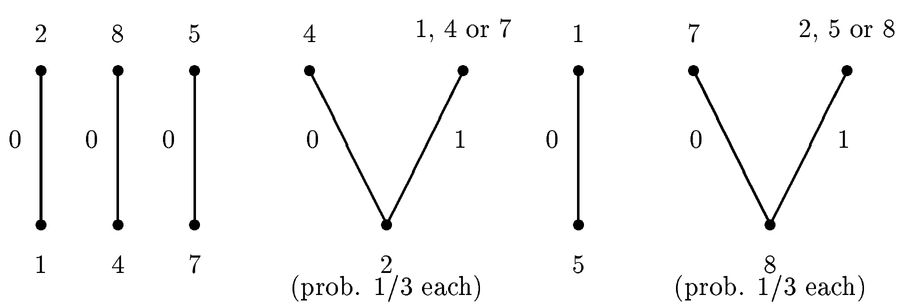


FIGURE 3. Branching Process $\mathcal{B}[9]$ transitions (with edge labels added)

that

$$1.87 < \liminf_{k\to\infty} R(k)^{1/k} < 1.92 \ .$$

How do the data in Figure 2 compare with the predictions from the branching process model $\mathcal{B}[9]$? To obtain as exact a numerical comparison with Table 1 as possible, we computed for $k \le 30$ the quantities

$$\begin{aligned} E[\widetilde{N}^-(k)] &:= E[\min\{X_k : \text{ take } R(k) \text{ i.i.d. draws}\}] \\ E[\widetilde{N}^+(k)] &:= E[\max\{X_k : \text{ take } R(k) \text{ i.i.d. draws}\}] \end{aligned}$$

using the values of $R(k)$ from Table 1, drawing the root node uniformly from 1,2,4,5,7,8. The results appear in Table 2. The method of computating $E[\widetilde{N}^-(k)]$ and $E[\widetilde{N}^+(k)]$ is described in the Appendix. In Table 2 the quantities

$$\begin{aligned} \widetilde{D}^+(k) &= (\frac{4}{3})^{-k} E[\widetilde{N}^+(k)] \ , \\ \widetilde{D}^-(k) &= (\frac{4}{3})^{-k} E[\widetilde{N}^-(k)] \ , \end{aligned}$$

both exhibit some initial fluctuations and then $\widetilde{D}^-(k)$ appears to steadily decrease with k, while $\widetilde{D}^+(k)$ appears to steadily increase with k. This contrasts with the analogous quantities in Table 1, which appear to be roughly constant. If we computed these expected values $E[\widetilde{N}^-(k)]$ and $E[\widetilde{N}^+(k)]$ using $2\cdot 3^k$ draws instead of $R(k)$ draws, the disagreement with Table 1 would be even greater.

In § 5 we prove theoretical results concerning the analogues of Conjectures C and C$^\#$ for the branching process model $\mathcal{B}[9]$. We prove that the analogue of Conjecture C holds for these statistics, using a result on tail probabilities for leaf count distributions for a general class of branching processes, which is proved in § 4. We prove that the analogue of Conjecture C$^\#$ doesn't hold, and $\widetilde{D}^-(k) \to 0$ and $\widetilde{D}^+(k) \to \infty$ as $k \to \infty$.

4. Tail Probabilities for Leaf Count Distributions

We consider multi-type Galton-Watson processes $\mathcal{G}$ which have n types of individuals. In such a process an individual of type i lives for exactly one time period t and gives rise to a set of progeny of various types at time $t+1$. We assume that $\mathcal{G}$ has a finite *mean matrix* $\mathbf{M} = [\mathbf{M}_{i,j}]_{1\le i,j\le n}$, where $\mathbf{M}_{i,j}$ gives the expected

k	# draws $R(k)$	$E[\widetilde{N}^-(k)]$	$E[\widetilde{N}^+(k)]$	$\left(\frac{4}{3}\right)^k$	$\widetilde{D}^-(k)$	$\widetilde{D}^+(k)$
1	4	1.00	2.00	1.33	0.750	1.500
2	8	1.00	2.77	1.78	0.562	1.557
3	14	1.00	3.96	2.37	0.422	1.669
4	24	2.00	5.46	3.16	0.633	1.728
5	42	2.00	7.55	4.21	0.475	1.792
6	76	3.00	9.99	5.62	0.534	1.778
7	138	3.07	14.31	7.49	0.409	1.911
8	254	4.00	19.20	9.99	0.401	1.923
9	470	5.00	26.45	13.32	0.375	1.986
10	876	7.00	35.97	17.76	0.394	2.026
11	1638	8.32	48.63	23.68	0.352	2.054
12	3070	10.81	65.53	31.57	0.342	2.076
13	5766	12.92	89.17	42.09	0.307	2.118
14	10850	17.12	119.58	56.12	0.305	2.131
15	20436	22.49	162.12	74.83	0.300	2.166
16	38550	30.16	218.52	99.77	0.302	2.190
17	72806	38.42	294.11	133.03	0.289	2.211
18	137670	49.91	395.94	177.38	0.281	2.232
19	260612	64.49	533.21	236.50	0.273	2.255
20	493824	85.41	715.96	315.34	0.271	2.270
21	936690	112.45	963.62	420.45	0.268	2.292
22	1778360	148.38	1294.74	560.60	0.265	2.310
23	3379372	193.77	1739.01	747.47	0.259	2.327
24	6427190	254.38	2335.64	996.62	0.255	2.344
25	12232928	334.18	3135.96	1328.83	0.252	2.360
26	23300652	441.25	4207.62	1771.77	0.249	2.375
27	44414366	581.63	5647.11	2362.36	0.246	2.390
28	84713872	766.94	7575.10	3149.81	0.243	2.405
29	161686324	1009.74	10159.40	4199.75	0.240	2.419
30	308780220	1331.40	13623.43	5599.67	0.238	2.433

TABLE 2. Branching Process Expected Values

number of progeny of type j produced by an individual of type i. We assume that $\mathcal{G}$ is *positively regular*, which means that some power $\mathbf{M}^k$ has all entries strictly positive. Under the positive regularity assumption the mean matrix $\mathbf{M}$ has a maximal real eigenvalue ρ of multiplicity one, which we call the *growth rate* of $\mathcal{G}$. Let $N_i(k)$ denote the total number of individuals at time k of a process starting from a single individual of type i at time 0. We say that $\mathcal{G}$ has *finite second moments* if $E[N_i(1)^2] < \infty$ for $1 \le i \le n$.

We prove below a result showing that the upper and lower tails of the logarithm of the leaf count distributions $N_i(k)$ of multi-type Galton-Watson processes have double-exponential decay in k as $k \to \infty$, provided that the processes satisfy some mild extra conditions, which we now introduce. A multi-type Galton-Watson process is *boundedly branching*, if there is an upper bound L on the number of progeny that an individual (of any type) can have in one time period. Such a process is *strictly branching* if an individual always has at least two progeny in each time period.

Theorem 4.1. *Let $\mathcal{G}$ be a multi-type Galton-Watson process with n types which is positively regular, has finite mean matrix $\mathbf{M}$ with maximal real eigenvalue ρ, and which is supercritical, i.e. $\rho > 1$.*

(i) If $\mathcal{G}$ is boundedly branching, then for any $r > \rho$ there are positive constants α, δ depending on r such that for all $k \ge 1$,

$$Prob\{N_i(k) > r^k\} \le \exp(-\alpha(1+\delta)^k)\ ,\ \ 1 \le i \le n\ \ . \tag{8}$$

(ii) If $\mathcal{G}$ is strictly branching and has finite second moments, then for any $r < \rho$ there are positive constants α, δ depending on r such that for all $k \ge 1$,

$$Prob\{N_i(k) < r^k\} \le \exp(-\alpha(1+\delta)^k)\ ,\ \ 1 \le i \le n\ \ . \tag{9}$$

Before giving the proof, we note that the conclusions of both (i) and (ii) certainly require some extra restriction on the Galton-Watson process $\mathcal{G}$ beyond being positively regular and supercritical. Concerning (i), if a single-type Galton-Watson process $\mathcal{G}$ has the probability p_m of m offspring satisfying $p_m = cm^{-4}$ for large m

(so $\mathcal{G}$ has a finite second moment), then for sufficiently large r,

$$\mathrm{Prob}\{N_1(k) > r^k\} \geq \mathrm{Prob}\{N_1(1) > r^k\} \geq p_{r^k} \geq cr^{-4k} \quad ,$$

which violates (8). Concerning (ii), if $\mathcal{G}$ is not strictly branching, and $p_1 > 0$, then

$$\mathrm{Prob}\{N_1(k) < r^k\} \geq (p_1)^k \quad ,$$

which violates (9).

Proof. (i) Suppose that $r > \rho$ is given. By hypothesis there is a finite bound L for the maximum number of progeny that a single individual can have in one time period. The argument we give does not depend on the type of the individual at time 0, so we omit explicit reference to it.

Let $N^{(i)}(k)$ denote the number of individuals of type i at time k, and define the *type vector* $\mathbf{v}(k)$ at period k by

$$\mathbf{v}(k) := (N^{(1)}(k), N^{(2)}(k), \dots, N^{(n)}(k)) \quad .$$

Also let $N^{(i,j)}(k, k+1)$ denote the number of individuals of type j at period $k+1$ that are progeny of an individual of type i at period k.

We claim that there is a constant $\delta > 0$ such that for all sufficiently large k, there is some intermediate time l, with $0 \leq l \leq k-1$ and a pair (i,j) of types with $\mathbf{M}_{i,j} \neq 0$, such that

$$N^{(i)}(l) \geq (1+\delta)^k \quad , \tag{10}$$

and

$$N^{(i,j)}(l, l+1) \geq (1+\delta)\mathbf{M}_{i,j} N^{(i)}(l) \quad . \tag{11}$$

We argue by contradiction, and suppose there were no such time l. Set $\mathbf{e} = (1, 1, \dots, 1)$ and observe that all the type vectors satisfy coordinatewise the inequality

$$\mathbf{v}(l+1) \leq (1+\delta)\mathbf{v}(l)\mathbf{M} + (1+\delta)^k L\mathbf{e} \quad ,$$

because the first term on the right bounds the contribution to $\mathbf{v}(l+1)$ of individuals of type j at time $l+1$ that are progeny of those types i at time l for which (11) doesn't hold, while the second term on the right bounds the contribution from types

i for which (10) doesn't hold. Iterating this inequality for $0 \le l \le k-1$ starting with $\mathbf{v}(0) \le \mathbf{e}$, we have

$$\mathbf{v}(k) \le nL(1+\delta)^k \mathbf{e}(\mathbf{I} + (1+\delta)\mathbf{M} + (1+\delta)^2\mathbf{M}^2 + \ldots + (1+\delta)^k\mathbf{M}^k) \ . \tag{12}$$

By Perron-Frobenius theory the matrix $\mathbf{M}$ has spectral radius ρ and by the positive regularity hypothesis its set of eigenvalues on the circle $|z| = \rho$ consists of a single simple eigenvalue at $z = \rho$, hence there is a constant c_0 with

$$\mathbf{e}\mathbf{M}^k \le c_0 \rho^k \mathbf{e} \ .$$

Thus (12) yields

$$\mathbf{v}(k) \le c_0 nL(1+\delta)^{2k}\rho^k \mathbf{e} \ ,$$

hence

$$N(k) = \mathbf{v}(k)\mathbf{e}^T \le c_0 n^2 L(1+\delta)^{2k}\rho^k \ .$$

If we therefore choose δ so that $1 < 1+\delta < (\frac{r}{\rho})^{1/2}$, then this bound contradicts $N(k) > r^k$ for all $k \ge k_0$, proving the claim.

To bound $\text{Prob}\{N(k) > r^k\}$ it thus suffices to bound the probability of the event (10) and (11) occurring over all triples (i,j,l). Now the random variable $N^{(i,j)}(l,l+1)$ is a sum of $N^{(i)}(l)$ independent draws from an integer-valued probability distribution $\{p_m\}$, where p_m is the probability that an individual of type i on $\mathcal{G}$ has exactly m progeny of type j. By definition the distribution $\{p_m\}$ has expected value $E[p] = \mathbf{M}_{i,j}$, and we also know that $p_m = 0$ for all $m \ge L$. Now we can apply Chernoff's theorem (as quoted in [10], p. 234) with $N^{(i)}(l)$ draws to obtain the bound

$$\text{Prob}\{N^{(i,j)}(l,l+1) \ge (1+\delta)E[p]N^{(i)}(l)\} \le \exp(-\alpha N^{(i)}(l)) \ , \tag{13}$$

where $\alpha = -g((1+\delta)E[p])$ with

$$g(a) := \sup_{\theta \in \mathbf{R}} \left\{ \theta a - \log\left(\sum_{m=0}^{L} p_m e^{m\theta} \right) \right\} \ .$$

We check that $\alpha > 0$. Certainly $g(a) \ge 0$, by taking $\theta = 0$ above, and the strict convexity of $\log\left(\sum_{m=0}^{L} p(m)e^{m\theta}\right)$ allow one to check that for $a > E[p]$ the minimizer on the right side is not at $\theta = 0$, hence $\alpha > 0$.

Now combining (10), (11) and (13) yields

$$\mathrm{Prob}\{N(k) > r^k\} \le n^2 k \exp(-\alpha(1+\delta)^k) \quad ,$$

valid for $k \ge k_0$. Decreasing α and δ towards 0 as necessary, we make (8) valid for all $k \ge 1$.

(ii) Suppose that $r < \rho$ is given. The strictly branching assumption guarantees that

$$N_i(t) \ge 2^t \ , \quad \text{all } t \ge 1 \quad , \tag{14}$$

holds for $1 \le i \le n$. Now view a tree of depth k as consisting of a rooted tree of depth t which has $N(t)$ subtrees each of depth $l := k - t$ growing from each of its leaves. All of these subtrees grow independently, and each of them can have at most r^k leaves, because the whole tree has r^k leaves by hypothesis. Thus using (14) we obtain the bound

$$\begin{aligned} \mathrm{Prob}\{N_i(k) < r^k\} &\le (\mathrm{Prob}\{\text{depth } l \text{ subtree has } < r^k \text{ leaves}\})^{N_i(t)} \\ &\le (\max_{1 \le j \le n} \{\mathrm{Prob}\{N_j(l) < r^k\}\})^{2^t} \quad . \end{aligned} \tag{15}$$

We choose $t = \alpha k$ for a small α and wish to bound the probability that a tree of depth $l = (1-\alpha)k$ has no more than r^k leaves. Since $\mathbf{M}$ is positively regular and second moments exist, the Kesten-Stigum theorem ([8], Theorem 1) applies to give positive constants u_i such that

$$E[N_i(l)] = (u_i + o(1))\rho^l \quad \text{as } l \to \infty \quad . \tag{16}$$

Furthermore by the finite second moment assumption, there is a finite upper bound on the second moment of $\frac{N_i(l)}{\rho^l}$ valid for all $l \ge 1$ (see Harris [7], Theorem 9.2), hence by Chebyshev's inequality there is a constant $\gamma < 1$ such that

$$\mathrm{Prob}\{N_i(l) < E[N_i(l)]\} \le \gamma \ , \quad \text{all } l \ge 1 \quad , \tag{17}$$

holds for $1 \le i \le n$. To apply this in (15), it suffices to arrange that $E[N_i(l)] > r^k$. Now (16) implies that there is a positive constant c^* such that for $1 \le i \le n$,

$$E[N_i(l)] \ge c^* \rho^l ; \quad \text{all } l \ge 1 \quad . \tag{18}$$

Write $r = \rho^{\bar{c}}$ with $0 < \bar{c} < 1$ and choose

$$l = \bar{c}k - \log_2(c^*) \ ,$$

the point being that with this choice

$$E[N_j(l)] \geq c^* \rho^l \geq r^k \ , \quad 1 \leq j \leq n \ ,$$

the last inequality depending on the fact that $\rho \geq 2$. Thus, for $1 \leq j \leq n$,

$$\text{Prob}\{N_j(l) < r^k\}] \leq \text{Prob}\{N_j(l) < E[N_j(l)]\} \leq \gamma \ , \ \text{all } l \geq 1 \ .$$

Now (15) yields, setting $\gamma = \exp(-\alpha^*)$,

$$\begin{aligned} \text{Prob}\{N(k) < r^k\} &\leq \exp(-\alpha^* 2^{k-l}) = \exp(-\alpha^* c^* c^{(1-\bar{c})k}) \\ &\leq \exp(-\alpha(1+\delta)^k) \ , \end{aligned}$$

with $\alpha, \delta > 0$. ■

5. Application to $3x+1$ Branching Process Models

We consider now a "repeated branching process" model in which the model $\mathcal{B}[9]$ is grown to depth k, making $S(k)$ independent trials. The statistics that we are interested in are the minimum and maximum of the number of leaves over these $S(k)$ trials. We are interested in the case that $S(k)$ grows exponentially in k, so we consider $S(k) = \lfloor \tau^k \rfloor$, where $\tau > 1$ is a fixed constant. The relevant random variables are

$$\widetilde{N}_\tau^-(k) = \min\{X_k : \text{take } \lfloor \tau^k \rfloor \text{ i.i.d. draws from } \mathcal{B}[9]\} \tag{19}$$

$$\widetilde{N}_\tau^+(k) = \max\{X_k : \text{take } \lfloor \tau^k \rfloor \text{ i.i.d. draws from } \mathcal{B}[9]\} \tag{20}$$

The scaled random variables $(\frac{4}{3})^k(\widetilde{N}_\tau^-(k))^{-1}$ and $(\frac{4}{3})^{-k}\widetilde{N}_\tau^+(k)$ are analogous to the quantities $(\frac{4}{3})^k(N^-(k))^{-1}$ and $(\frac{4}{3})^{-k}N^+(k)$ in Table 1.

We first prove that an analogue of Conjecture C holds for this "repeated branching process" model using $\mathcal{B}[9]$.

Theorem 5.1. *For any fixed $\tau > 1$, with probability one, the branching process $\mathcal{B}[9]$ has*

$$\lim_{k\to\infty} (\widetilde{N}_\tau(k))^{1/k} = \lim_{k\to\infty} (\widetilde{N}_\tau^+(k))^{1/k} = \frac{4}{3} \ .$$

Proof. The branching process $\mathcal{B}[9]$ has mean matrix $\mathbf{M}$ given in Table 3, with

$$\mathbf{M} := \begin{array}{c|cccccc|} \text{Type} \diagdown \text{Type} & 1 & 4 & 7 & 2 & 5 & 8 \\ \hline 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 4 & 0 & 0 & 0 & 0 & 0 & 1 \\ 7 & 0 & 0 & 0 & 0 & 1 & 0 \\ 2 & \frac{1}{3} & \frac{4}{3} & \frac{1}{3} & 0 & 0 & 0 \\ 5 & 1 & 0 & 0 & 0 & 0 & 0 \\ 8 & 0 & 0 & 1 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \hline \end{array}$$

TABLE 3. Mean Matrix $\mathbf{M}$ for $\mathcal{B}[9]$

left-eigenvector $\mathbf{v} = (1,1,1,1,1,1)$, and $\mathbf{M}^4$ has positive entries so $\mathcal{B}[9]$ is positively regular, cf. [10], Theorem 3.2. It is certainly *boundedly branching*, so part (i) of Theorem 4.1 applies to give for $r > 4/3$,

$$\text{Prob}\{\widetilde{N}_\tau^+(k) \geq r^k\} \leq \tau^k \exp(-\alpha(1+\delta)^k) \ .$$

Since $\sum_{k=1}^\infty \tau^k \exp(-\alpha(1+\delta)^k)$ converges, we conclude that with probability one

$$\limsup_{k\to\infty} (\widetilde{N}_\tau^+(k))^{1/k} \leq \rho = \frac{4}{3} \ . \tag{21}$$

The key point of the proof concerns the strict branching property. Although $\mathcal{B}[9]$ is not strict branching, repeated application of it for four time periods is. This is easy to check using the branching data in Figure 3. The repeated branching process $\mathcal{G} = (\mathcal{B}[9])^{(*4)}$ has mean matrix $\mathbf{M}^4$, which has growth rate ρ^4, and it has finite second moments since it is boundedly branching. Now part (ii) of Theorem 4.1 applies to $\mathcal{G}$, to yield that for any $r < (\frac{4}{3})^4$, there holds

$$\mathrm{Prob}\{\widetilde{N}_\tau^-(4k) < r^k\} \le \tau^k \exp(-\alpha(1+\delta)^k) \ .$$

As in the argument above, we conclude that with probability one,

$$\liminf_{k\to\infty} (\widetilde{N}_\tau^-(4k))^{1/k} \ge \left(\frac{4}{3}\right)^4 \ .$$

Since $N(k) \le N(k+i) \le 2^i N(k)$ for $0 \le i \le 3$, we conclude that with probability one

$$\liminf_{k\to\infty} (\widetilde{N}_\tau^-(k))^{1/k} \ge \rho = 4/3 \ . \tag{22}$$

Combining (21) and (22) and using the fact that $\widetilde{N}_\tau^-(k) \le \widetilde{N}_\tau^+(k)$ in any sampling of trees, we conclude that $\lim_{k\to\infty}(\widetilde{N}_\tau^+(k))^{1/k}$ and $\lim_{k\to\infty}(\widetilde{N}_\tau^-(k))^{1/k}$ both exist and equal $\frac{4}{3}$ with probability one. ■

Remark. This proof applies to all the branching process models $\mathcal{B}[3^j]$ with $j \ge 2$, because all the processes $(\mathcal{B}[3^j])^{(*4)}$ have the strict branching property for $j \ge 2$. It does not apply to the branching processes $\mathcal{B}[1]$ and $\mathcal{B}[3]$, because they have no iterate possessing the strict branching property. In fact the lower bound (21) is false for $\mathcal{B}[1]$ and $\mathcal{B}[3]$ whenever $\tau > \frac{4}{3}$.

We now show that the analogue of Conjecture $\mathrm{C}^\#$ is false for the "repeated branching process" model using $\mathcal{B}[9]$.

Theorem 5.2. *For any fixed $\tau > 1$, the branching process $\mathcal{B}[9]$ has*

$$\lim_{k\to\infty} \widetilde{D}^-(k) = 0 \ ,$$

and

$$\lim_{k\to\infty} \widetilde{D}^+(k) = +\infty \ .$$

Proof Let W_k^m for $m \pmod 9$ enumerate the number of leaves of type m of a random tree of depth k drawn from $\mathcal{B}[9]$, with root node drawn uniformly from $\{1, 2, 4, 5, 7, 8\}$. Set

$$\mathbf{W}_k := (W_k^1, W_k^2, W_k^4, W_k^5, W_k^7, W_k^8) \ ,$$

so that $X_k = W_k^1 + W_k^2 + W_k^4 + W_k^5 + W_k^7 + W_k^8$. Now let $\mathbf{w}_k$ denote the probability distribution of the random vector $(\frac{4}{3})^{-k}\mathbf{W}_k$. Now $E[X_1 \log X_1] < \infty$, hence the Kesten-Stigum theorem ([8], Theorem 1) applies to give that the distributions $\mathbf{w}_k$ converge weakly to a limiting distribution $\mathbf{w}_\infty$, which is of the form

$$\mathbf{w}_\infty = w \cdot \mathbf{v} \tag{23}$$

where $\mathbf{v}$ is a constant vector and w is a one-dimensional positive random variable which is absolutely continuous, except for a possible jump at the origin. (The jump at the origin represents the probability of extinction.) Furthermore $\mathbf{v}$ is the (unique) left eigenvector corresponding to the maximal real eigenvalue ρ of the mean matrix $\mathbf{M}$ of the Galton-Watson process; in our case $\mathbf{M}$ is given in Table 3 and $\mathbf{v} = \mathbf{e} = (1, 1, \dots, 1)$.

The conditional distribution $w_i := \{w \mid \text{ initial type } i\}$ has the expectation

$$E[w \mid \text{ initial type } i] = u_i, \tag{24}$$

where bu is a right eigenvector of $\mathbf{M}$, and the jump q_i at the origin depends on the type i. The q_i are just probabilities of extinction, hence in the case of $\mathcal{B}[9]$ there are no jumps (all $q_i = 0$), and *each conditional distribution* $w_i = \{w \mid \textit{ initial type } i\}$ *is strictly positive*[3] *on* $\mathbf{R}^+$, by Theorem 2(iv) of Chapter V.6 of Athreiya and Ney [3]. Now the random variables $(\frac{4}{3})^{-k}\widetilde{N}_\tau^-(k)$ and $(\frac{4}{3})^{k}\widetilde{N}_\tau^+(k)$ sample values in the tails of the distributions $\mathbf{w}_k$, i.e. values that lie outside any fixed region $(\epsilon, 1-\epsilon)$ in the cumulative distribution for large enough k. Since $\mathbf{w}_k$ converge weakly to $\mathbf{w}_\infty$ it follows from the strict positivity of w on $\mathbf{R}^+$ that

$$\widetilde{D}^-(k) = \left(\frac{4}{3}\right)^{-k} \widetilde{N}_\tau^-(k) \to 0 \quad \text{as} \quad k \to \infty$$
$$\widetilde{D}^+(k) = \left(\frac{4}{3}\right)^{-k} \widetilde{N}_\tau^+(k) \to \infty \quad \text{as} \quad k \to \infty \ ,$$

so Theorem 5.2 follows. ■

[3] A detailed proof of the positivity of w for the single-type Galton-Watson process appears as Theorem 2 of Sect. II.5 of Athreiya and Ney [3]. See also Lemma 7 of Kesten and Stigum [8].

6. Average Leaf Counts and Conjecture $C^{\#}$.

We return to the study of $3x+1$ trees, and study fluctuations in the leaf counts of such trees caused by the branching pattern at the base of the tree, in its first j levels. That is, we estimate the expected size of pruned $3x+1$ trees $\mathcal{T}_k^*(a)$ whose root node lies in a fixed congruence clase $l \pmod{3^j}$. This expected value is

$$E_k^*[l \bmod 3^j] := 3^{j-(k+1)} \sum_{\substack{a \pmod{3^{k+1}} \\ a \equiv l \pmod{3^j}}} N_k^*(a) \quad , \tag{25}$$

for $j \geq 1$.

Theorem 6.1. *For each* $j \geq 1$ *and* $l \not\equiv 0 \pmod 3$ *there is a positive constant* $W[l \bmod 3^j]$ *such that*

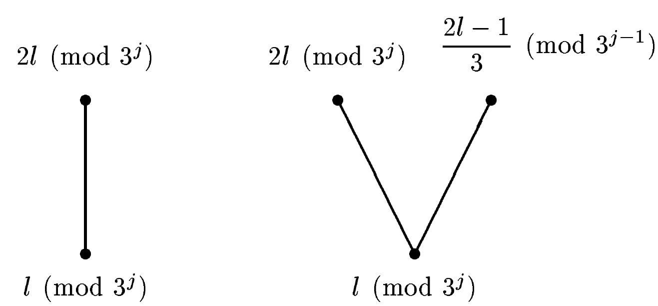
$$E_k^*[l \bmod 3^j] = (W[l \bmod 3^j] + o(1)) \left(\frac{4}{3}\right)^k \quad \text{as } k \to \infty \quad . \tag{26}$$

Proof. We will use the formula (corresponding to $j = 0$)

$$\frac{1}{2 \cdot 3^k} \sum_{\substack{a \pmod{3^{k+1}} \\ a \not\equiv 0 \pmod 3}} N_k^*(a) = \left(\frac{4}{3}\right)^k \quad , \tag{27}$$

which is proved in Theorem 4.1 of [10].

We first establish recursions for the quantities $E_k^*[l \bmod 3^j]$. The recursions are based on the bottom branching of the $3x+1$ tree, which depends on $l \pmod 9$, and which is pictured in Figure 4. This gives the recursion



(i) $l \equiv 1, 4, 5, 7 \pmod 9$ (ii) $l \equiv 2, 8 \pmod 9$

Figure 4. Tree-branching patterns

$$N_k^*(a) = N_{k-1}^*(2a) + \psi(a \bmod 9) N_{k-1}^*\left(\frac{2a-1}{3}\right) , \tag{28}$$

in which

$$\psi(l \bmod 9) := \begin{cases} 0 & \text{if} \quad l \equiv 1,4,5,7 \pmod 9 , \\ 1 & \text{if} \quad l \equiv 2,8 \pmod 9 , \end{cases} \tag{29}$$

is an indicator function for the presence of a branch of the tree with edge label 1. Summing (28) over all $a \pmod{3^{k+1}}$ yields, for $k \geq j \geq 2$,

$$E_k^*[l \bmod 3^j] = E_{k-1}^*[2l \bmod 3^j] + \psi(l \bmod 9) E_{k-1}^*[\frac{2l-1}{3} \bmod 3^{j-1}] . \tag{30}$$

If $j \geq 2$ but $1 \leq k < j$ then

$$E_k^*[l \bmod 3^j] = E_k^*[l \bmod 3^k] . \tag{31}$$

The case $j = 1$ must be treated separately. The recursions become

$$\begin{cases} E_k^*[1 \bmod 3] &= E_{k-1}^*[2 \bmod 3] , \\ E_k^*[2 \bmod 3] &= E_{k-1}^*[1 \bmod 3] + \frac{2}{3}\left(\frac{4}{3}\right)^{k-1} , \end{cases} \tag{32}$$

where (27) was used to obtain the last equation. We have

$$E_1^*[1 \bmod 3] = 1 = \frac{3}{4}\left(\frac{4}{3}\right) \quad \text{and} \quad E_1^*[2 \bmod 3] = \frac{5}{3} = \frac{5}{4}\left(\frac{4}{3}\right) ,$$

from which we deduce

$$E_k^*[l \bmod 3] \;=\; w_l(k)\left(\frac{4}{3}\right)^k , \quad l = 1,2 ,$$

in which $w_1(k)$ and $w_2(k)$ obey the recurrences

$$\begin{aligned} w_1(k) &= \frac{3}{4} w_2(k-1) \\ w_2(k) &= \frac{3}{4} w_1(k-1) + \frac{1}{2} . \end{aligned}$$

This yields $w_2(k) = \frac{9}{16} w_2(k-2) + \frac{1}{2}$, from which one easily deduces

$$\begin{cases} w_1(k) &= \frac{6}{7} + O\left(\left(\frac{3}{4}\right)^k\right) , \\ w_2(k) &= \frac{8}{7} + O\left(\left(\frac{3}{4}\right)^k\right) . \end{cases} \tag{33}$$

Now (26) follows for $j = 1$ with:

$$W[1 \bmod 3] = \frac{6}{7} \quad \text{and} \quad W[2 \bmod 3] = \frac{8}{7} . \tag{34}$$

For $j \geq 2$, let $W[l \bmod 3^j]$ be defined recursively in j as the unique solution to the system of linear equations

$$W[l \bmod 3^j] = \frac{3}{4}\left(W[2l \bmod 3^j] + \psi(l \bmod 9) W[\frac{2l-1}{3} \bmod 3^{j-1}]\right) \quad . \tag{35}$$

Here the quantities $W[\frac{2l-1}{3} \bmod 3^{j-1}]$ are known, and this linear system has matrix $\mathbf{I} - \frac{3}{4}\mathbf{P}$ where $\mathbf{P}$ is a certain permutation matrix, which is clearly invertible since $(\mathbf{I} - \frac{3}{4}\mathbf{P})^{-1} = \mathbf{I} + \frac{3}{4}\mathbf{P} + \left(\frac{3}{4}\mathbf{P}\right)^2 + \dots$.

Next, define the quantities $\Delta_k[l \bmod 3^j]$ by

$$E_k^*[l \bmod 3^j] = \left(W[l \bmod 3^j] + \Delta_k[l \bmod 3^j]\right)\left(\frac{4}{3}\right)^k \quad ,$$

and set

$$\overline{\Delta}_k[3^j] := \max_{l \not\equiv 0 \pmod 3} \left|\Delta_k[l \bmod 3^j]\right| \quad .$$

We claim that there are positive constants c_j such that

$$\overline{\Delta}_k[3^j] \leq c_j \left(\frac{7}{8}\right)^k \quad . \tag{36}$$

If so, then (26) follows.

For $j = 1$ this holds for all $k \geq 1$ by (33), choosing a suitable value for c_1.

We prove (36) for $j \geq 2$ by induction on j, where for each j we verify it for all k by induction on $k \geq 1$. The constants c_j are defined recursively by

$$c_j = \max\left(6c_{j-1}, \max_{1 \leq k \leq j} \overline{\Delta}_k[3^j]\left(\frac{8}{7}\right)^k\right) \quad .$$

Assume (36) is true for $j-1$ and all k. For j and $1 \leq k < j$ (36) holds by definition of c_j.

For $j \geq k$, the recursions (30) give

$$\Delta_k[l \bmod 3^j] = \frac{3}{4}\left(\Delta_{k-1}[2l \bmod 3^j] + \psi(l \bmod 9)\Delta_{k-1}[\frac{2l-1}{3} \bmod 3^{j-1}]\right) \quad .$$

In particular the recursions yield the inequality

$$\overline{\Delta}_k[3^j] \leq \frac{3}{4}\overline{\Delta}_{k-1}[3^j] + \frac{3}{4}\overline{\Delta}_{k-1}[3^{j-1}] \quad . \tag{37}$$

This gives, by the induction hypothesis

$$\begin{aligned}\overline{\Delta}_k[3^j] &\leq \frac{3}{4}c_j\left(\frac{7}{8}\right)^{k-1} + \frac{3}{4}c_{j-1}\left(\frac{7}{8}\right)^{k-1} \\ &\leq c_j\left(\frac{7}{8}\right)^{k} ,\end{aligned}$$

since $c_j \geq 6c_{j-1}$, completing the induction step. (One can prove that $\overline{\Delta}_k[3^j]$ is $O((\frac{3}{4})^{(1+\epsilon)k})$ for any $\epsilon > 0$, by a similar argument.) ■

The densities $W[l \bmod 3^j]$ are determined recursively by solving the linear system (35). For $j = 2$, we obtain

$$(38) \qquad \begin{cases} W[1 \bmod 9] = \frac{3456}{3367} \ , & W[2 \bmod 9] = \frac{4608}{3367} \ , \\ W[4 \bmod 9] = \frac{3258}{3367} \ , & W[8 \bmod 9] = \frac{4344}{3367} \ , \\ W[7 \bmod 9] = \frac{1944}{3367} \ , & W[5 \bmod 9] = \frac{2592}{3367} \ . \end{cases}$$

For later use, we show that the quantities $W[l \bmod 3^j]$ satisfy, for all $j \geq 1$, the mean value formula

$$(39) \qquad \frac{1}{2\cdot 3^{j-1}} \sum_{\substack{l \pmod{3^j} \\ l \not\equiv 0 \pmod 3}} W[l \bmod 3^j] = 1 \ .$$

This holds for $j = 1$ by (34). For $j \geq 2$, summing up (35) over all $l \bmod 3^j$ yields

$$\frac{1}{4} \sum_{\substack{l \pmod{3^j} \\ l \not\equiv 0 \pmod 3}} W[l \bmod 3^j] = \frac{3}{4} \sum_{\substack{l' \pmod{3^{j-1}} \\ l' \not\equiv 0 \pmod 3}} W[l' \bmod 3^{j-1}] \ .$$

Now (39) follows by induction on j, for the above equation yields

$$\frac{1}{2\cdot 3^{j-1}} \sum_{\substack{l \pmod{3^j} \\ l \not\equiv 0 \pmod 3}} W[l \bmod 3^j] = \frac{1}{2\cdot 3^{j-2}} \sum_{\substack{l' \pmod{3^{j-1}} \\ l' \not\equiv 0 \pmod 3}} W[l' \bmod 3^{j-1}] = 1 \ .$$

The quantities $W[l \bmod 3^j]$ yield asymptotic bounds on the number of leaves in extremal trees.

Corollary 6.1. *For each $j \geq 1$, with $D^{\pm}(k) = \left(\frac{4}{3}\right)^{-k} N^{\pm}(k)$,*

$$\limsup_{k\to\infty} D^+(k) \geq W_j^+ := \max_{l \not\equiv 0 \pmod 3} W[l \bmod 3^j] \ ,$$

and

$$\liminf_{k\to\infty} D^-(k) \leq W_j^- := \min_{l \not\equiv 0 \pmod 3} W[l \bmod 3^j] \ .$$

Proof. Extreme values on leaf counts satisfy obvious inequalities in relation to mean values. ■

Letting $j \to \infty$ in Corollary 6.1 yields

$$
\begin{array}{rcccl}
\limsup_{k\to\infty} D^+(k) & \geq & W_\infty^+ & := & \limsup_{j\to\infty} W_j^+ \ , \\
\liminf_{k\to\infty} D^-(k) & \leq & W_\infty^- & := & \liminf_{j\to\infty} W_j^- \ .
\end{array}
\tag{40}
$$

In order for Conjecture C$^\#$ to hold, the quantities W_∞^+ and W_∞^- must satisfy

$$
0 < W_\infty^- < 1 < W_\infty^+ < \infty \ . \tag{41}
$$

This is unproved, but would follow from the Limit Function Conjecture stated below.

Table 4 presents data on the extreme values W_j^- and W_j^+, as well as on the quantities

$$
\eta_j^+ := \max_{l \equiv l' \pmod{3^{j-1}}} |W[l \bmod 3^j] - W[l' \bmod 3^j]|
$$

which bound how fast W_j^- and W_j^+ are changing as $j \to \infty$. It also gives the quantities $l \pmod{3^j}$ attaining W_j^+ and W_j^-, with l expressed in base 3, as well as $l \pmod{3^{j-1}}$ attaining η_j^+.

On comparing the values in Table 4 with the extreme densities $D^+(k)$ and $D^-(k)$ in Table 1, we see that by $k = 9$ the values W^+ and W^- seem to be accounting for nearly all of the observed variation in $D^+(k)$ and $D^-(k)$. (Note that $W_8^- < D^-(8)$; this is not contradictory because W_8^- is an asymptotic limit as $k \to \infty$. However we must have $D^-(8) \leq \min_l E_8^*[l \bmod 3^8](\frac{3}{4})^8$.)

The data in Table 4 suggests that the quantities $W[l \bmod 3^j]$ may explain all the extremal variation in leaf count sizes in an asymptotic sense. We therefore propose:

Extremal Limit Conjecture. *The quantities $D^+(k)$ and $D^-(k)$ satisfy*

$$
\begin{array}{rcl}
\limsup\limits_{k\to\infty} D^+(k) & = & W_\infty^+ \ , \\
\liminf\limits_{k\to\infty} D^-(k) & = & W_\infty^- \ .
\end{array}
$$

	$\min l$		$\max l$		$\max l$	
j	$l \pmod{3^j}$	W_j^-	$l \pmod{3^j}$	W_j^+	$l \pmod{3^{j-1}}$	η_j^+
1	1_3	0.857	2_3	1.143	—	0.286
2	21_3	0.577	02_3	1.369	2_3	0.599
3	221_3	0.528	202_3	1.493	22_3	0.407
4	2221_3	0.517	0202_3	1.561	222_3	0.302
5	02221_3	0.504	12122_3	1.611	2222_3	0.209
6	202221_3	0.503	212122_3	1.649	22222_3	0.166
7	1002221_3	0.498	1212122_3	1.672	222222_3	0.116
8	21002221_3	0.497	21212122_3	1.690	2222222_3	0.085
9	221002221_3	0.494	202020202_3	1.704	22222222_3	0.062
10	1221002221_3	0.493	0202020202_3	1.714	222222222_3	0.045
11	21221002221_3	0.491	20202020202_3	1.721	2222222222_3	0.033

TABLE 4. Extreme densities $W[l \bmod 3^j]$.

Finally we observe that the recursion for $W[l \bmod 3^j]$ has a regular structure. These quantities interpolate to a function defined almost everywhere on the invertible 3-adic integers

$$\mathbf{Z}_3^\times = \{\alpha \in \mathbf{Z}_3 : \alpha \equiv 1 \text{ or } 2 \pmod 3\} \ ,$$

as we now show. We view $\mathbf{Z}_3^\times$ as a measure space with the 3-adic measure μ with $\mu(\mathbf{Z}_3) = 1$, so that $\mu(\mathbf{Z}_3^\times) = \frac{2}{3}$.

Theorem 6.2. *For μ-almost all $\alpha = \sum_{j=0}^\infty a_j 3^j \in \mathbf{Z}_3^\times$, the following limit exists:*

$$W_\infty(\alpha) := \lim_{j\to\infty} W[\alpha \bmod 3^j] \ . \tag{42}$$

Proof. Let $\mu^\times = \frac{3}{2}\mu$ so that $\mu^\times(\mathbf{Z}_3^\times) = 1$ is a probability measure. Define for $j \geq 1$ the functions $W_j : \mathbf{Z}_3^\times \to \mathbf{R}$ by

$$W_j(\alpha) := W[\alpha \bmod 3^j] \ ,$$

and view $\{W_j : j \geq 1\}$ as random variables on $\mathbf{Z}_3^\times$ with respect to $\mu^\times$. We claim that $\{W_j : j \geq 1\}$ is a martingale with respect to the σ-fields $\{\mathcal{F}_j : j \geq 1\}$ with

$\mathcal{F}_j = \{\text{residue classes } \pmod{3^j}\}$. The martingale property is that, for each residue class $\alpha \pmod{3^j}$,

$$E[W_{j+1}(\beta) \mid \beta \equiv \alpha \pmod{3^j}] = W_j(\alpha) \quad ,$$

which is equivalent to

$$\frac{1}{3}\sum_{k=0}^{2} W[\alpha + k\cdot 3^j \bmod 3^{j+1}] = W[\alpha \bmod 3^j] \quad . \tag{43}$$

To establish (43), we define

$$X[l \bmod 3^j] := W[l \bmod 3^j] - W[l \bmod 3^{j-1}] \quad .$$

The recursion (35) for $l \bmod 3^j$ subtracted from that for $l \bmod 3^{j+1}$ gives

$$X[l \bmod 3^{j+1}] = \frac{3}{4}\left(X[2l \bmod 3^{j+1}] + \psi(l \bmod 9)X[\frac{2l-1}{3} \bmod 3^j]\right) \quad . \tag{44}$$

We now prove by induction on $j \geq 1$ that

$$A[l \bmod 3^{j+1}] := \sum_{k=0}^{2} X[l + k\cdot 3^j \bmod 3^{j+1}] = 0 \quad , \quad \text{all } l \pmod{3^{j+1}} \quad . \tag{45}$$

The base case $j=1$ is verified by direct computation, using (38). For the induction step, (44) summed over $l, l+3^j, l+2\cdot 3^j$ gives

$$A[l \bmod 3^{j+1}] = \frac{3}{4}\left(A[2l \bmod 3^{j+1}] + \psi(l \bmod 9)A[\frac{2l-1}{3} \bmod 3^j]\right) \quad . \tag{46}$$

By the induction hypothesis, the last term $A[\frac{2l-1}{3} \bmod 3^j] = 0$. Now (46) becomes the invertible linear system

$$(\mathbf{I} - \frac{3}{4}\mathbf{P})\left(A[l \bmod 3^{j+1}]\right) = \mathbf{0} \quad ,$$

hence (45) follows. Substituting the definition of $X[l \bmod 3^{j+1}]$ in (45) gives (43), hence $\{W_j : j \geq 1\}$ is a martingale.

The mean value formula (39) gives

$$E[|W_j|] = E[W_j] = \int_{\mathbf{Z}_3^\times} W_j(\alpha)d\mu^\times(\alpha) = 1 \quad , \quad j \geq 1 \quad .$$

Now the Martingale Convergence Theorem (see Billingsley [4], Theorem 35.4) applies to $\{W_j : j \geq 1\}$ so (42) follows. ∎

We may define the limit function $W_\infty(\alpha)$ for all $\alpha \in \mathbf{Z}_3^\times$, by

$$W_\infty(\alpha) = \limsup_{j\to\infty} W[\alpha \bmod 3^j] \ .$$

Here $W_\infty(\alpha) \geq 0$ and the value $+\infty$ is allowed. The Martingale Convergence Theorem also gives

$$E[W_\infty] = \int_{\mathbf{Z}_3^\times} W_\infty(\alpha) d\mu^\times(\alpha) = E[|W_\infty|] = E[|W_1|] = 1 \ . \tag{47}$$

The data in Table 4 suggest that $\eta_j^+ \to 0$ rapidly enough that $\sum_{j=0}^\infty \eta_j^+ < \infty$, in which case $W[\alpha \bmod 3^j]$ would converge uniformly to $W_\infty(\alpha)$ for all $\alpha \in \mathbf{Z}_3^\times$. Therefore we propose:

Limit Function Conjecture. *The function $W_\infty : \mathbf{Z}_3^\times \to \mathbf{R}$ is continuous and nonzero, and*

$$W_\infty(\alpha) = \lim_{j\to\infty} W[\alpha \bmod 3^j]$$

holds for all $\alpha \in \mathbf{Z}_3^\times$.

If this conjecture is true, then taking $\lim_{j\to\infty}$ in (35) shows that $W_\infty(\alpha)$ satisfies the functional equation

$$W_\infty(\alpha) = \frac{3}{4}\left(W_\infty(2\alpha) + \psi(\alpha \bmod 9) W_\infty(\frac{2\alpha - 1}{3})\right) \ . \tag{48}$$

Since $\mathbf{Z}_3^\times$ is compact, this conjecture also implies that

$$W_\infty^+ = \sup_{\alpha\in\mathbf{Z}_3^\times} W_\infty(\alpha) < \infty$$

and

$$W_\infty^- = \inf_{\alpha\in\mathbf{Z}_3^\times} W_\infty(\alpha) > 0 \ .$$

Since (47) and (48) imply that $W_\infty(\alpha)$ cannot be the constant function 1, we must have $W_\infty^- < 1 < W_\infty^+$, and (41) follows. Thus the Limit Function Conjecture and the Extremal Limit Conjecture together imply Conjecture $\mathrm{C}^\#$.

Finally, we note some resemblance of the recursion (35) to the Krasikov inequalities studied in [2].

Appendix A. Computation of Expected Values of Extreme Leaf Counts

Although there are a double-exponential number of different trees possible at depth k of such a branching process, the data $E[\widetilde{N}^-(k)]$ and $E[\widetilde{N}^+(k)]$ in Table 2 were computed in single-exponential time as follows: Let X_k^i for $i \pmod 9$ be a random variable counting the number of leaves at depth k of a sample tree drawn from the branching process $\mathcal{B}[9]$, starting from a single individual of type i, and let $P[X_k^i = x] := \text{Prob}\{X_k^i = x\}$. Then, the distributions of X_k^i and X_k were computed from the recursion

$$\begin{aligned}
P[X_0^i = 1] &= 1 \ , \\
P[X_k^i = x] &= P[X_{k-1}^{2i} = x] \text{ if } i = 1, 4, 5, \text{ or } 8 \ , \\
P[X_k^2 = x] &= \sum_{y=0}^{\infty} P[X_{k-1}^{2i} = x - y] \left(\frac{P[X_{k-1}^1 = y] + P[X_{k-1}^4 = y] + P[X_{k-1}^7 = y]}{3} \right) , \\
P[X_k^8 = x] &= \sum_{y=0}^{\infty} P[X_{k-1}^{2i} = x - y] \left(\frac{P[X_{k-1}^2 = y] + P[X_{k-1}^5 = y] + P[X_{k-1}^8 = y]}{3} \right) , \\
&\text{and} \\
P[X_k = x] &= \frac{1}{6} \sum_{i \pmod 9} P[X_k^i = x] \ .
\end{aligned}$$

The cumulative distribution function $f_k(t)$ of the number of leaves was then computed. Finally the cumulative distributions of the minimum and maximum of $R(k)$ draws were computed using $(1 - (1 - f_k(t)))^{R(k)}$ and $f_k(t)^{R(k)}$, respectively. The entire computation took about 15 minutes on 150 Mhz MIPS R4400 processor.

Acknowledgement

We are indebted to R. Pemantle for suggesting the method used to prove Theorem 4.1. We thank J. A. Reeds and 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] K. B. Athreiya and P. E. Ney, *Branching processes*, Springer-Verlag, New York, 1972.

[4] P. Billingsley, *Probability and measure*, John Wiley & Sons, New York, 1979.

[5] K. Borovkov and D. Pfeifer, *Estimates for the Syracuse problem via a probabilistic model*, preprint, 1993.

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

[7] T. E. Harris, *The theory of branching processes*, Springer-Verlag, Berlin, 1963, (Reprint: Dover 1988).

[8] H. Kesten and B. P. Stigum, *A limit theorem for multidimensional Galton-Watson processes*, Ann. Math. Stat. **37** (1966), 1211–1223.

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

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

[11] G. Leavens and M. Vermeulen, *$3x+1$ search programs*, Computers & Mathematics, with Applications **24** (1992), no. 11, 79–99.

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

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

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

[15] G. J. Wirsching, *On a Markov chain underlying the backwards Syracuse algorithm*, Rev. Roum. Math. Pure Appl. **39** (1994), 915–926.

[16] G. J. Wirshing, *The dynamical system on the natural numbers generated by the $3n+1$ function*, Habilitationsschrift, Katholische Universität Eichstätt, 1995.

AT&T Bell Laboratories, Murray Hill, NJ 07974

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

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