A probabilistic model for the 5k+1 problem and related maps

Volkov, Stanislav

Published in:
Stochastic Processes and their Applications

2006

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Stochastic Processes and their Applications 116 (2006) 662–674
stochastic processes and their applications
www.elsevier.com/locate/spa

A probabilistic model for the $5x + 1$ problem and related maps

Stanislav Volkov*

Department of Mathematics, University of Bristol, BS8 1TW, UK

Received 20 December 2004; received in revised form 13 September 2005; accepted 1 November 2005
Available online 15 December 2005

Abstract

We construct a probabilistic model which "mimics" the behaviour of a certain number-theoretical algorithm. This model involves study of a binary tree with randomly labelled edges, such that the labels have different distributions, depending on their directions. A number of properties of this tree are rigorously studied. As an application, this study could suggest what one could expect in the original algorithm.

MSC: primary 60D05, 60K37; secondary 11B73

Keywords: Random environment on trees; Large deviations; First-passage percolation; $3x + 1$ Problem

1. Motivation

Take a positive integer x. If x is even, divide it by 2; otherwise multiply it by 5 and add 1. This will be the new x. Call this mapping $M(x)$. Formally,

$$M(x) = \begin{cases} 5x + 1, & \text{if } x \text{ is odd,} \\ x/2, & \text{if } x \text{ is even.} \end{cases} \tag{1.1}$$

This mapping closely resembles the famous still unresolved $3x + 1$ problem, which states that if we replace 5 by 3 in (1.1), then for any $k \in \mathbb{Z}_+$, $M^{(n)}(k) = 1$ for some n. See [5] and [4] for ample references.

* Tel.: +44 117 928 9146; fax: +44 117 928 7999.
E-mail address: S.Volkov@bristol.ac.uk.

0304-4149/$ - see front matter
doi:10.1016/j.spa.2005.11.007

In our case, however, from numerical computations it seems intuitively to follow that any $k \in \mathbb{Z}_+$ (say below 10^8) after consecutive applications of M either converges to one of the three cycles

- $1 \to 6 \to 3 \to 1$,
- $13 \to 66 \to 33 \to 166 \to 83 \to 416 \to 208 \to 104 \to 52 \to 26 \to 13$ or
- $17 \to 86 \to 43 \to 216 \to 108 \to 54 \to 27 \to 136 \to 68 \to 34 \to 17$,

or "diverges" (in our computations becomes larger than 10^{12}).

An interesting question is how many numbers between 1 and K converge to a finite cycle (we conjecture that there are only three cycles mentioned above) as K becomes large. Namely, what is the asymptotic behavior of $\tilde{Q}(K)$ where

$$\tilde{Q}(K) = \mathsf{card}\,\{k \in \mathbb{Z}_+,\ k \le K : M^{(n)}(k) \not\to \infty\}?$$

Numerical estimations consistently support the conjecture that $\tilde{Q}(K) \sim K^{0.68}$, that is $\tilde{Q}(K)$ grows asymptotically as some power of K. Our primary purpose is to "explain" (but not to prove!) why this could be the case, by analyzing the corresponding probabilistic model.

First, let us show how to construct all k's for which $M^{(n)}(k) = 1$ for some n. We call this set B_1, and we will show that it has a tree structure. The root of B_1 will be the vertex denoted as '1'. Next vertices are defined recursively. If vertex 'k' $\in B_1$, then vertex '$2k$' $\in B_1$ as well. Also, if 'k' is even and can be represented as $k = 5y + 1$ where y is a positive integer, then 'y' $\in B_1$ as well. Now construct a tree whose vertices are numbers of B_1 and such that 'k' is a parent of either one or two vertices: '$2k$' (the northwest child) and perhaps 'y' when applicable (the northeast child). Finally, to preserve the tree structure, we artificially remove the edge connecting '6' to '1'.

Observe that 'k' "branches to the northeast" exactly on every fourth step as $16 \bmod 5 = 1$. Also every fifth time this happens (since $16^5 \bmod 25 = 1$), the northeast child 'y' will be "infertile" — divisible by 5 — and hence will **never** produce a "northeast" child again. Starting from an infertile child, the subtree will be then isomorphic to $\mathbb{Z}_+$. See Fig. 1.

Let trees B_{13} and B_{17} be constructed in a similar fashion, starting from 13 and 17 respectively, and prohibiting the links '66' $\to$ '13' and '86' $\to$ '17' respectively. Suppose that the only possible cycles are the three described above; then under this assumption the union $B_1 \bigcup B_{13} \bigcup B_{17}$ gives the set of *all* $k \in \mathbb{Z}_+$ for which the sequence $M^{(n)}(k)$ does not diverge to infinity.

Consider the tree B_1, and throw away all its infertile branches[1] as they contribute to a negligible fraction of vertices of B_1; take two consecutive nodes of the remaining tree (=vertices adjacent to three other vertices) and denote by 'k' the one which is closer to the root '1'. The number at the other vertex is either '$16k$' or '$256k$' (the latter happens with "probability" $1/5$, because we have removed the infertile branches), whenever the other vertex is the northwest child. If the other vertex is the northeast child, then the number there is approximately $2^p k/5$ where $p = 1, 2, 3, 4$ with equal "probability" $1/4 \times (1-1/5)$ or $256 \times 2^p k/5$ where $p = 1, 2, 3, 4$ with equal "probability" $1/4 \times 1/5$. The fraction $1/5$ here reminds us again that we throw away infertile branches. Then the nodes of B_1 (after throwing away all infertile branches) are isomorphic to a binary tree with a number assigned to each vertex. See Fig. 2.

In the next section we consider a stochastic model, related to the binary tree we have just constructed. In this model, we will consider a binary tree with the numbers assigned to its

[1] In [6] this would be called a *pruned tree*.

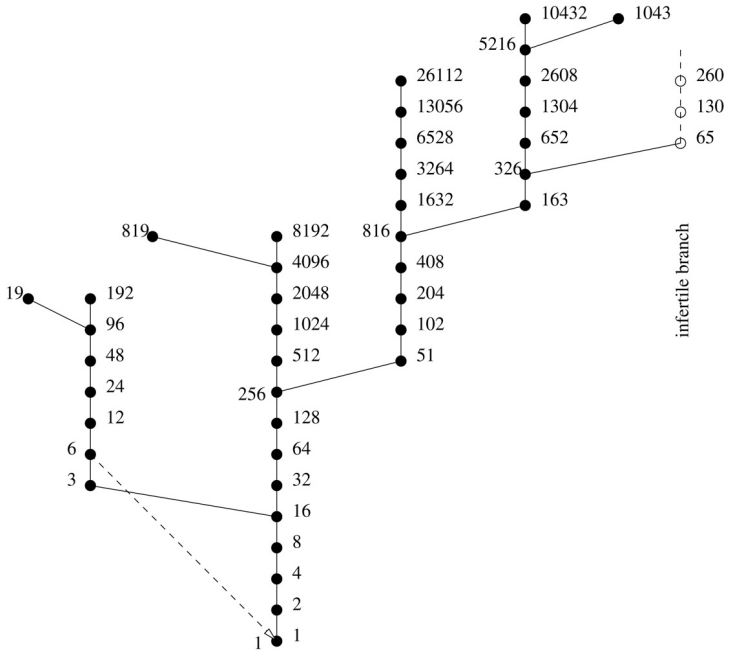

Fig. 1. Tree B_1. To preserve the tree structure we "prohibit" the link from '6' to '1'.

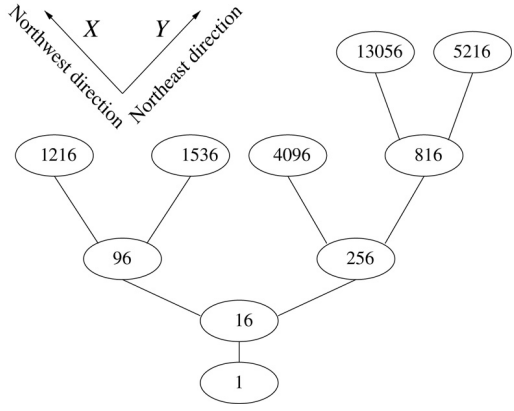

Fig. 2. Tree constructed of the nodes of B_1.

vertices, and these numbers will correspond in some sense to the *logarithms* of those numbers assigned to the nodes of B_1. In this paper we study the properties of this stochastic model, as providing a guide to the properties of the $5x + 1$ iteration. We introduce a quantity $Q(K)$ in the stochastic model which is intended to simulate the behavior of $\tilde{Q}(K)$. We are able to prove rigorous results for the stochastic model. However, we do not attempt to answer rigorously the number-theoretical question about the speed of growth of $\tilde{Q}(K)$, which appears to be a very difficult problem.

The idea of this paper was inspired by a talk of Yakov Sinai talk in Cambridge on the stochastic approach to the famous $3x + 1$ problem, based on [8] and [9]. Besides this, [10] and many other papers by the same author contain similar combinatorial constructions applied to the $3x + 1$ problem. In particular, [11] is a relevant general reference for the tree of inverse iterates, analogous to B_1 above.

Finally, in [6], which is the previous work most closely related to the current paper, the authors construct a tree (see Fig. 1, formula (3.2) on pp. 240–241) which is again essentially a $3x + 1$ analog of the tree B_1 constructed in our paper. The authors refer to this tree as a *branching process model*, and analyze the corresponding branching process to obtain, for example, estimates of the speed of growth of this tree. Note that, unlike [6], we will treat our tree rather as *non-random* but equipped with random labels.

2. Randomly labelled binary tree

Let B be the regular binary rooted tree with the root v_0, that is a tree for which each vertex is adjacent to exactly three other vertices, with the exception of the root v_0 which is adjacent only to two vertices. For any two vertices u and v of this tree, let $\ell(u, v)$ be the unique self-avoiding path connecting u and v, let $|\ell(u, v)|$ be its length, and let $|v| := |\ell(v, v_0)|$. Let $V_n := \{u \in B : |\ell(u, v_0)| = n\} \equiv \{u \in B : |u| = n\}$ denote the set of 2^n vertices at the distance n from the root.

Let X and Y be some random variables. To each e edge of B assign a random variable Z_e ("label") such that $\{Z_e\}$ are independent, and the distribution of Z_e is the same as of X (Y resp.) when e points to the northwest (northeast resp.).

Definition 1. For a vertex $v \in B$, let $T(v)$ be the sum of the random variables on the self-avoiding path $\ell(v_0, v)$ connecting v to the root. For definiteness, consistently set $T(v_0) = 0$.

The quantity $T(v)$ is the analogue to the value $\log n$ where n is a label of the vertex of the tree of inverse iterates of the $5x + 1$ problem.

Definition 2. For a given number $b \in \mathbb{R}$, let the random variable $Q(b)$ denote the number of elements of the (possibly disconnected) set $\{v : T(v) \le b\}$. The quantity can take value $+\infty$.

The quantity $Q(b)$ is the analogue in the stochastic model of the quantity $\tilde{Q}(b)$ for the $5x + 1$ problem. However, the stochastic model is so general that it allows $Q(b) = +\infty$, so it becomes an interesting question to determine conditions under which $Q(b)$ is finite almost surely.

This question is non-trivial when $\mathbb{E}\,X = \mu_x > 0$, $\mathbb{E}\,Y = \mu_y > 0$ and yet at least one of these two random variables, say Y, is not strictly positive. Indeed, leaving aside the degenerate case where $\mu_x \mu_y = 0$, if at least one of μ_x or μ_y is negative, the strong law of large numbers implies that along one of the branches of the tree all $T(v)$, except at most finitely many, will be negative. On the other hand, if both $X > 0$ a.s. and $Y > 0$ a.s. then for all $v \neq v_0 T(v) > 0$. Hence, from now on we will make the following assumptions:

(H1) $\mu_x = \mathbb{E}\,X > 0$ and $\mu_y = \mathbb{E}\,Y > 0$,
(H2) $\mathbb{P}(Y < 0) > 0$.

Next, let

$$\varphi_X(\theta) = \mathbb{E}\,\mathrm{e}^{\theta X}, \qquad \varphi_Y(\theta) = \mathbb{E}\,\mathrm{e}^{\theta Y}$$

denote the moment generating functions of X and Y respectively. We will additionally assume that

(H3) the moment generating functions $\varphi_X(\theta)$ and $\varphi_Y(\theta)$ exist and are finite for all $\theta \in \mathbb{R}$.

Randomly labelled trees have been extensively studied in the literature, as the are relevant to random walks in random environment, first-passage percolation, as well as many other important problems. The paper [7] studied a related tree model; however, our set-up is somewhat different from theirs; also we ask different questions about the process on the tree. Yet some ideas from [7] are applied in the present paper.

Later we will make use of the following statement.

Lemma 1 (*Chernoff–Cramer*). *Let W be a random variable with mean μ and suppose that its Laplace transform*

$$\varphi_W(\theta) = \mathbb{E}\, \mathrm{e}^{\theta W}$$

is finite for all θ. Let S_n denote the sum of n i.i.d. copies of W. Then for any $a > \mu$

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}[S_n \geq na] \to -\gamma(a)$$

where the rate function

$$\gamma(a) = \sup_{\theta\in\mathbb{R}} [a\theta - \log \varphi_W(\theta)] \tag{2.2}$$

is strictly positive for $a > \mu$.

The proof of this well-known statement about large deviations can be found in various forms, e.g. in [2] Chapter 5.11 or [3] Theorem I.4.

Now write

$$\varphi_W(\theta) = \frac{\varphi_X(-\theta) + \varphi_Y(-\theta)}{2} \tag{2.3}$$

which is the Laplace transform of a random variable W equal to either $-X$ or $-Y$ with equal probability.

The random variable W is relevant to studying the distribution of weights assigned to the vertices of the tree. We set

$$\mu := \mu_W = \mathbb{E}\, W = -\frac{\mathbb{E}\, X + \mathbb{E}\, Y}{2} = -\frac{\mu_x + \mu_y}{2}.$$

As will be shown below, the average expected weight $T(v)$ of a vertex at depth n in there is $-n\mu$. We will be interested in vertices whose weight is significantly smaller than the mean.

Definition 3. For any real a, and fixed depth n, define the random variable

$$R_n(a) := |\{v \in V_n :\ T(v) \leq na\}|$$

and

$$R(a) := |\{v \in B : T(v) \leq |v|a\}| = \sum_{n=0}^{\infty} R_n.$$

Note that for a given infinite tree the quantity $Q(b)$ is finite if $R(a)$ is finite for some $a > 0$. A necessary condition for finiteness of $R(a)$ is that $a < -\mu$, as shown in the next section.

In the following section we establish a phase transition in a for the event $\{R(a) < \infty\}$.

3. Finiteness criterion for $R(a)$

Theorem 1. *Suppose that* (H1),(H2) *and* (H3) *hold, and additionally that* $a < (\mu_x + \mu_y)/2$. *Then*

(i) *If* $\gamma(-a) > \log 2$ *then* $R(a) < \infty$ *a.s.*
(ii) *If* $\gamma(-a) < \log 2$ *then* $R(a) = \infty$ *a.s.*

Remark 1. In the case $a \geq (\mu_x + \mu_y)/2$ one can show $R(a) = \infty$ a.s. as well. This immediately follows from the second part of Theorem 1, the fact that $\gamma(x) \to 0$ as $x \to \mu$ (see Proposition 1 further on in the text), and the monotonicity of $R(a)$ in a.

Remark 2. In this paper we do not study the critical case when $\gamma(-a) = \log 2$.

Proof of Theorem 1. Recall that V_n is the set of vertices at the graph-theoretical distance n from the root. First, we obtain an upper bound on the probability

$$p_n = \mathbb{P}(T(v) \leq na \text{ for some } v \in V_n) = \mathbb{P}(R_n(a) > 0).$$

If it turns out that this quantity decays exponentially, then by the Borel–Cantelli lemma $T(v) \leq na$ a.s. only for finitely many v, thus immediately yielding $R(a) < \infty$ a.s. To prove this, it is convenient to consider the following construction, which will be used in further proofs as well.

Let D_k be a "climbing" random walk on B with $D_0 = v_0$ which goes either up left in the northwest direction or up right in the northeast direction with equal probabilities of $1/2$. Then for any $v \in V_n$ we have $\mathbb{P}(D_n = v) = 2^{-n}$. Let $\tilde{S}_n$ be the sum of the labels on the edges traversed by the path of the random walk up to time n. Then $\tilde{S}_n$ happens to have the same distribution as $\tilde{W}_1 + \cdots + \tilde{W}_n$, where $\tilde{W}_i$'s are i.i.d. random variables such that $\tilde{W}_i$ equals to X with probability $1/2$ and Y with probability $1/2$, and the Laplace transform of $W_i = -\tilde{W}_i$ is exactly $\varphi_W(\theta)$ given by (2.3). Hence setting $S_n = -\tilde{S}_n$ we obtain

$$\begin{aligned} p_n &\leq \sum_{v \in V_n} \mathbb{P}(T(v) \leq na) = 2^n \mathbb{P}(\tilde{S}_n \leq na) = 2^n \mathbb{P}(S_n \geq -na) \\ &= \exp\left[n \left\{ \log 2 + \frac{\log \mathbb{P}(S_n \geq -na)}{n} \right\} \right] \end{aligned}$$

which, provided $\gamma(-a) > \log 2$, by Lemma 1 indeed decays exponentially for large n.

Now suppose that $\gamma(-a) < \log 2$. Recall that $-a > \mu$, and note that in this region $\gamma(-a)$ is decreasing in a. By continuity of the rate function under the assumption that it is finite everywhere (see e.g. [3] Lemma I.14), there is an $a' < a$ such that still $\gamma(-a') < \log 2$. Fix some $\nu \in (0, \log 2 - \gamma(-a'))$. Then by Lemma 1 there is a positive integer m such that

$$2^m \mathbb{P}(S_m \geq -ma') = \exp\left\{ m \left(\log 2 + \frac{\log \mathbb{P}(S_m \geq -ma')}{m} \right) \right\} \geq \mathrm{e}^{\nu m}.$$

Construct the following branching process on B. The original particle is v_0. The members of the first generation are vertices $v \in V_m$ for which $T(v) \leq ma'$. A vertex $u \in V_{2m}$ is a member of the second generation if it is a descendant of some member v of the first generation, such that $T(u) - T(v) \leq ma'$. Recursively, define the members of the k-th generation as a subset of V_{km}: a vertex $u \in V_{km}$ belongs to this generation if its m-th ancestor v (that is $v = ancestor_m(u) = V_{(k-1)m} \cap \ell(v_0, u)$) belongs to the $(k-1)$-st generation of the branching

process and also $T(u) - T(v) \le ma'$. Then for each member v of this branching process we necessarily have $T(v) \le |v|a'$.

Observe that for different members of the $(k-1)$-st generation the number of offsprings in the k-th generation are i.i.d. random variables, also independent of k. Thus each member of this branching process behaves independently of the others, and the number of children has the same distribution. On the other hand, the average number of children per member is

$$\mathbb{E} \sum_{v \in V_m} 1_{T(v) \le ma'} = 2^m \mathbb{P}(\tilde{S}_m \le ma') \ge \mathrm{e}^{vm} > 1.$$

Therefore this is a supercritical branching process (see e.g. [2], p. 173) which survives with a positive probability, say $\rho > 0$.

Now fix a positive integer n and for each $v \in V_n$ consider a subtree B_v rooted at v. Since B_v is a replica of B and these trees do not overlap for different v's, the probability that the supercritical branching process described above but started at v rather than at v_0 will survive for at least one $v \in V_n$ is $1 - (1-\rho)^n$. In the event of survival, the construction of the branching process would imply that for infinitely many vertices $u \in B_v$ we have $T(u) - T(v) < (|u| - n)a'$. This, in turn, yields

$$T(u) < |u|a$$

as long as

$$|u| \ge \frac{T(v) - na'}{a - a'},$$

and therefore $R(a) = \infty$ with probability at least $1 - (1-\rho)^n$. Since the latter expression can be made arbitrarily close to 1 by choosing n large, the second part of the Theorem follows. □

4. Cardinality of $Q(b)$

Recall that $Q(b) = \mathsf{card}\,\{v : T(v) \le b\}$. In this section we obtain a criterion for finiteness of $Q(b)$. When the criterion holds, we will also compute the asymptotic speed of growth of $\mathbb{E}\,Q(b)$ depending on b.

If $\mu > 0$, then for any b, $Q(b) = \infty$ a.s. Indeed, by strong law of large numbers applied to the sequence $T(D_n) = \tilde{S}_n$ where D_n is the random walk described in the proof of Theorem 1, $T(D_n)/n \to -\mu$ a.s., and also D_n must be at different vertices at different times n, since $D_n \in V_n$. Hence $T(D_n) \to -\infty$ and thus for any $b \in \mathbb{R}$, $Q(b)$ is infinite a.s.

Ignoring the special case $\mu = 0$, from now on we suppose that

(H4) $\mu = \mu_W = -(\mathbb{E}\,X + \mathbb{E}\,Y)/2 < 0$.

The next result gives a condition for $Q(b)$ to be finite. This relates to the case $a = 0$ of Theorem 1.

Theorem 2. *Suppose that* (H1),(H2),(H3)*, and* (H4) *hold, and let b be any non-negative number.*

(i) *If* $\gamma(0) = -\sup_{\theta \in \mathbb{R}} \log \varphi_W(\theta) > \log 2$ *then* $Q(b) < \infty$ *a.s.*
(ii) *If* $\gamma(0) < \log 2$*, then* $Q(b) = \infty$ *a.s.*

Proof. First, assume $\gamma(0) > \log 2$. Following the lines of the first part of the proof of Theorem 1, we obtain that

$$\begin{aligned}\tilde{p}_n &:= \mathbb{P}(T(v) \le b \text{ for some } v \in V_n)\\ &\le \exp\left[n\left\{\log 2 + \frac{\log \mathbb{P}(S_n \ge -n(b/n))}{n}\right\}\right].\end{aligned}$$

By continuity of the rate function, there exists $\varepsilon > 0$ such that $\gamma(-\varepsilon) > \log 2$. Fix $\nu \in (0, \gamma(-\varepsilon) - \log 2)$. Then Lemma 1 implies that there is n_0 such that

$$\frac{\log \mathbb{P}(S_n \ge -n\varepsilon)}{n} < -\log 2 - \nu$$

as soon as $n > n_0$. Therefore for all $n > \max\{n_0, b/\varepsilon\}$

$$\frac{\log \mathbb{P}(S_n \ge -n(b/n))}{n} < \frac{\log \mathbb{P}(S_n \ge -n\varepsilon)}{n} < -\log 2 - \nu$$

yielding $\tilde{p}_n < \exp(-\nu n)$, which again decays exponentially and therefore by the Borel–Cantelli lemma $Q(b)$ is finite a.s.

When $\mu < \log 2$ the proof is much simpler. Indeed, by Theorem 1, for $a = 0$ we have $R(a) = \infty$ a.s. On the other hand, $R(0) = Q(0)$, and since $Q(b)$ is monotonically increasing in b, we conclude that $Q(b) = \infty$ a.s. for all $b \ge 0$. □

Theorem 2 shows that the following assumption is useful for making $Q(b)$ finite almost surely.

(H5) $\gamma(0) = -\sup_{\theta\in\mathbb{R}} \log \varphi_W(\theta) > \log 2$.

In what follows, we will assume this (as well as (H1)–(H4)) and obtain a result concerning the expected growth rate of $Q(b)$ as $b \to \infty$. This will be a large deviation result, and for it we will need basic properties of the rate function, which we state next.

Proposition 1 (*Lemma I.14 in [3]*). *Suppose that the moment generating function* $\varphi(\theta) = \mathbb{E}\, e^{\theta W}$ *of a random variable* W *with mean* μ *is finite for all* $\theta \in \mathbb{R}$. *Then its rate function* $\gamma(a) = \sup_{\theta\in\mathbb{R}}(a\theta - \log \varphi(\theta))$ *satisfies the following:*

1. γ *is lower semi-continuous and convex in* $\mathbb{R}$.
2. γ *has compact level sets.*
3. γ *is continuous and strictly convex on int*$(\mathcal{D})$, *where* $\mathcal{D} = \{a \in \mathbb{R} : \gamma(a) < \infty\}$ *and int*$(\mathcal{D})$ *is the interior of* $\mathcal{D}$.
4. γ *is smooth on int*$(\mathcal{D})$.
5. $\gamma(a) \ge 0$ *with equality if and only if* $a = \mu$.
6. $\gamma''(\mu) = 1/\mathbb{V}\mathsf{ar}\,(W)$.

The following result estimates the expected growth rate of the set $Q(b)$ under our model assumptions.

Theorem 3. *Suppose* (H1)–(H5) *all hold. Then the following limit exists and is finite: there exists*

$$\beta := \lim_{b\to\infty} \frac{\log \mathbb{E}\, Q(b)}{b}.$$

Furthermore, β is given in terms of the rate function $\gamma(a)$ by

$$\beta = \max_{a\in(0,-\mu]} \frac{\log 2 - \gamma(-a)}{a} \tag{4.4}$$

and the minimum is attained strictly inside the interval.

Proof. We let

$$f(a) = \frac{\log 2 - \gamma(-a)}{a} \qquad \text{and} \qquad \beta = \max_{a\in(0,-\mu]} f(a). \tag{4.5}$$

Observe that the maximum of $f(\cdot)$ is achieved strictly inside $(a^*, -\mu)$, where $a^* \in (0, -\mu)$ is such that $f(a^*) = 0$. This, as well as the existence and uniqueness of a^*, follows from the fact that $f(a)$ is a continuous function with

$$\begin{aligned}
&\lim_{a\downarrow 0} f(a) = -\infty,\\
&f(-\mu) = \frac{\log 2}{-\mu} > 0,\\
&f'(-\mu) = \frac{\gamma(-a)a - \log 2 + \gamma(-a)}{a^2}|_{a=-\mu} = -\frac{\log 2}{\mu^2} < 0.
\end{aligned}$$

From the proof of Theorem 1 we remember that for any uniformly randomly chosen $v \in V_n$, $T(v)$ has the same distribution as $\tilde{S}_n = \tilde{W}_1 + \cdots + \tilde{W}_n = -S_n$. Therefore,

$$\begin{aligned}
\mathbb{E}\, Q(b) &= \mathbb{E} \sum_{v\in B} 1_{T(v)\le b} = \sum_{v\in B} \mathbb{P}(T(v) \le b)\\
&= \sum_{n=0}^{\infty} \sum_{v\in V_n} \mathbb{P}(T(v) \le b) = \sum_{n=0}^{\infty} 2^n \mathbb{P}(\tilde{S}_n \le b)\\
&= \sum_{n=0}^{\infty} \mathrm{e}^{bU_n}
\end{aligned} \tag{4.6}$$

where

$$U_n := \frac{\log 2 + \frac{1}{n}\log \mathbb{P}(S_n/n \ge -b/n)}{b/n}.$$

From (H5) and the continuity of γ we conclude that there are $\varepsilon > 0$, $\delta > 0$ and $n_0 = n_0(\varepsilon, \delta)$ such that

$$\frac{1}{n}\log \mathbb{P}(S_n/n \ge -\varepsilon) < -(\log 2 + \delta)$$

for all $n \ge n_0$. Therefore, for $n \ge \max\{n_0, b/\varepsilon\}$ we have

$$U_n \le -\delta n/b \;\Rightarrow\; \mathrm{e}^{bU_n} \le \mathrm{e}^{-\delta n}.$$

On the other hand, by Markov's inequality, for any $\theta > 0$

$$\frac{1}{n}\log \mathbb{P}(S_n/n \ge -b/n) \le \theta(b/n) + \log \varphi_W(\theta). \tag{4.7}$$

Taking the infimum over $\theta > 0$, we obtain that the LHS of (4.7) does not exceed $-\gamma(-b/n)$ if $-b/n \geq \mu$. Thus for $n \geq b/(-\mu)$ we have

$$U_n \leq \frac{\log 2 - \gamma(-b/n)}{b/n} \leq \beta.$$

Since for $n < b/(-\mu)$ obviously

$$U_n \leq \frac{\log 2}{b/n} < \frac{\log 2}{-\mu} = f(-\mu) \leq \beta,$$

combining the above, we conclude that for any n,

$$\mathrm{e}^{bU_n} \leq \mathrm{e}^{\beta b}.$$

As a result, splitting the sum (4.6) into two parts for b sufficiently large,

$$\sum_{n=0}^{\infty} \mathrm{e}^{bU_n} < (b/\varepsilon + 1)\mathrm{e}^{\beta b} + (1 - \mathrm{e}^{-\delta})^{-1}$$

yielding

$$\limsup_{b\to\infty} \frac{\log \mathbb{E}\, Q(b)}{b} \leq \beta.$$

Secondly, since $\gamma(\mu) = 0$ we can fix an a such that $a^* \leq a < -\mu$ and $\gamma(-a) < \log 2$. By Lemma 1 for any positive $\delta < \log 2 - \gamma(-a)$ there is an $n_0 = n_0(a, \delta)$ such that

$$\log 2 + \frac{1}{n} \log \mathbb{P}(S_n \geq -na) \geq \log 2 - \gamma(-a) - \delta$$

as soon as $n \geq n_0$. Suppose that b is sufficiently large, namely $b > n_0 a$, then

$$\mathbb{P}(S_n \geq -b) \geq \mathbb{P}(S_n \geq -na)$$

for n such that $n_0 \leq n \leq b/a$, which yields

$$\log 2 + \frac{1}{n} \log \mathbb{P}(S_n \geq -b) \geq \log 2 - \gamma(-a) - \delta.$$

Consequently, taking just one element of the sum (4.6), we have

$$\mathbb{E}\, Q(b) \geq \mathrm{e}^{bU_{\lfloor b/a \rfloor}} \geq \exp\{\lfloor b/a \rfloor [\log 2 - \gamma(-a) - \delta]\}$$

where $\lfloor \cdot \rfloor$ denotes the integer part. Now letting $\delta \downarrow 0$ and then maximizing with respect to $a \in (a^*, -\mu)$ to obtain

$$\liminf_{b\to\infty} \frac{\log \mathbb{E}\, Q(b)}{b} \geq \sup_{a\in(0,-\mu]} \frac{\log 2 - \gamma(-a)}{a} = \beta$$

which concludes the proof. □

5. Application to $5x + 1$ problem

Now we are going to link our probabilistic model to the original number-theoretic problem to provide an insight into the asymptotic behaviour of the quantity $\tilde{Q}(K)$ defined at the very beginning of the paper.

Let

$$X = \begin{cases} \log 16, & \text{with probability } 4/5, \\ \log 256, & \text{with probability } 1/5; \end{cases}$$

$$Y = \begin{cases} \log 2^1/5, & \text{with probability } 1/5 = 1/4 \times 4/5, \\ \log 2^2/5, & \text{with probability } 1/5, \\ \log 2^3/5, & \text{with probability } 1/5, \\ \log 2^4/5, & \text{with probability } 1/5, \\ \log 2^5/5, & \text{with probability } 1/20 = 1/4 \times 1/5, \\ \log 2^6/5, & \text{with probability } 1/20, \\ \log 2^7/5, & \text{with probability } 1/20, \\ \log 2^8/5, & \text{with probability } 1/20. \end{cases}$$

Construct the tree B labelling its edges with X and Y in the manner described in Section 2.

For a positive integer n, following our notation,

$$Q(\log(n)) = \mathsf{card}\,\{v : T(v) \le \log n\}$$

equals the number of vertices for which $\prod_{e\in\ell(v_0,v)} \mathrm{e}^{Z_e} \le n$ where $\ell(v_0, v)$ is again the set of edges on the unique self-avoiding path connecting the root to v. Then with this choice of X and Y, tree B with randomly assigned labels on edges "resembles" the binary tree of the nodes of B_1, obtained from tree B_1 in the manner described in Section 1. Of course, it equally "resembles" B_{13} and B_{17}, but on the log scale it is sufficient to consider just one of the three trees, as well as to ignore the vertices of B_1 located between the nodes, and also infertile branches.

We conjecture that $\tilde{Q}(K)$, the number of $k \in \{1, 2, \ldots, K\}$ for which the mapping M described in Section 1 is eventually periodic, behaves asymptotically as $Q(\log K)$.

We now determine the behavior of the quantity $Q(b)$ in our stochastic model, as given by our earlier theorems. The first step is to show that $Q(b)$ is finite. Indeed $\mu = -2.0025\cdots < 0$ and

$$\gamma(0) = 0.709\cdots > 0.693\cdots = \log 2,$$

and hence by Theorem 2 $Q(b)$ is a.s. finite for any $b > 0$ and thus $Q(\log K)$ is finite a.s. for $K > 1$. Secondly, using (4.4) we compute β_M, i.e. β for our particular model, which turns out to be

$$\beta_M = \beta \approx 0.678$$

where the maximum in (4.4) is achieved for $a \approx 0.22$, and Theorem 3 now yields

$$\mathbb{E}\, Q(\log K) \sim \mathrm{e}^{0.678 \log K} = K^{0.678}$$

This supports the following conjecture about the $5x + 1$ problem.

Conjecture 1. *$\tilde{Q}(K)$, the number of those positive integers $k \le K$ for which the mapping M does not iterate to ∞, behaves $\sim K^{0.68}$, that is*

$$\lim_{K\to\infty} \frac{\log \tilde{Q}(K)}{\log K} = \beta_M.$$

Note that the above computed β_M is in surprisingly good agreement with simulated data, see Section 5.1 (Fig. 3).

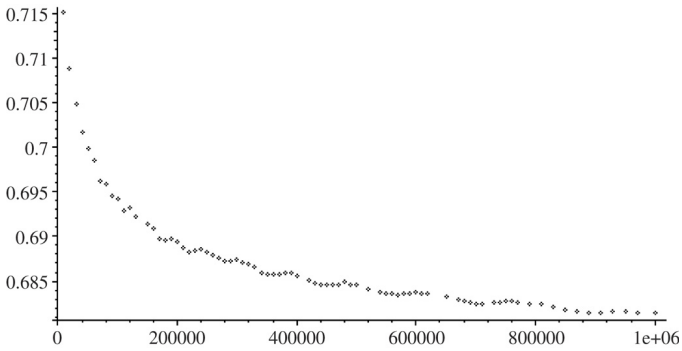

Fig. 3. $\hat{\beta}(m)$, numerical estimates of $\tilde{\beta}$.

Remark 3. As was mentioned by the referee, in fact one can *rigorously* prove a lower bound on the number of integers below K that iterate under mapping M to some fixed number N as a power of K (see [1]).

5.1. Numerical estimates

Since one cannot establish rigorously whether $M^{(n)}(k) \to \infty$ from numerical calculations, we artificially chose a threshold of $T = 4.2 \times 10^8$ to represent "infinity". An integer $k < T$ was called *bad*, if $M^{(n)}(k) > T$ for some $n \geq 0$, and *good* if $M^{(n)}(k) \in \{1, 13, 17\}$ for some n, while the programme did not encounter any other alternatives. For an $m < T$ the total number of "good" k's in $\{1, 2, \ldots, m\}$ was denoted as $q(m) = q(m, T)$. Then we computed the value

$$\hat{\beta}(m) = \frac{\log q(m)}{\log m}$$

for various values of m. Some results are in the table below:

$$\hat{\beta}(1000) = 0.741 \quad \hat{\beta}(10000) = 0.715 \quad \hat{\beta}(100000) = 0.694 \quad \hat{\beta}(1000000) = 0.680$$

(There is no point in doing this for larger values of m as it becomes to close to the threshold T which was dictated by computational arguments).

6. Possible extensions and open problems

So far we have only conjectured but not rigorously proved almost anything about mapping M given by (1.1). It would be interesting to obtain a rigorous proof of some properties of M, at least to show that there are $x \in \mathbb{Z}_+$ for which $M^{(n)}(x) \to \infty$ as $n \to \infty$. However, even this may be quite hard.

The most obvious extensions of the methods described in this paper are applications to various other number theoretical algorithms of similar type, to predict at least what kind of asymptotics can be expected there. From a probabilistic point of view, instead of considering binary trees, we could have considered any $d-$ry regular trees and obtained similar statements, with $\log 2$ replaced by $\log d$. The stochastic model should also be able to provide a heuristic on how many integers below n eventually iterate to some number N, where N does not have to be 1 or even in some cycle. The results should be identical to the ones obtained here, of course possibly with a different power.

Another interesting result would be if one can actually compute the limit of $\log Q(b)/b$, without the expectation sign.

Acknowledgements

The author would like to thank Geoffrey Grimmett, Mikhail Menshikov, and Yakov Sinai for useful discussions and suggestions. The anonymous referee's detailed comments and suggestions were also very valuable for references and the quality of exposition of this paper.

References

[1] D. Applegate, J.C. Lagarias, Density bounds for the $3x + 1$ problem. I. Tree-search method, Math. Comp. 64 (209) 411–426.
[2] G. Grimmett, D. Stirzaker, Probability and Random Processes, 3rd ed., Oxford University Press, New York, 2002.
[3] F. den Hollander, Large Deviations, in: Fields Institute Monographs, vol. 14, American Mathematical Society, Providence, 1992.
[4] J.C. Lagarias, The $3x + 1$ problem and its generalizations, Amer. Math. Monthly 92 (1) (1985) 3–23.
[5] J.C. Lagarias, The $3x + 1$ problem: An annotated bibliography (preprint), http://arxiv.org/abs/math.NT/0309224, 2004.
[6] J.C. Lagarias, A. Weiss, The $3x + 1$ problem: Two stochastic models, Ann. Appl. Probab. 2 (1992) 229–261.
[7] R. Lyons, P. Pemantle, Random walk in random environment and first-passage percolation on trees, Ann. Probab. 20 (1992) 125–136.
[8] Ya.G. Sinai, Uniform distribution in the $(3x + 1)$-problem, Mosc. Math. J. 3 (4) (2003) 1429–1440.
[9] Ya.G. Sinai, Statistical $(3x + 1)$ problem, Comm. Pure Appl. Math. 56 (7) (2003) 1016–1028.
[10] G. Wirsching, On the combinatorial structure of $3n + 1$ predecessor sets, Discrete Math. 148 (1–3) (1996) 265–286.
[11] G. Wirsching, The dynamical system generated by the $3n + 1$ function, in: Lecture Notes in Mathematics, vol. 1681, Springer-Verlag, Berlin, 1998.