Bull Braz Math Soc, New Series 33(2), 213-224

Structure Theorem for (d, g, h)-Maps

A. V. Kontorovich and Ya. G. Sinai

— Dedicated to IMPA on the occasion of its 50^{th} *anniversary*

Abstract. The $(3x + 1)$-Map, T, acts on the set, Π, of positive integers not divisible by 2 or 3. It is defined by $T(x) = \frac{3x+1}{2^k}$, where k is the largest integer for which $T(x)$ is an integer. The $(3x + 1)$-Conjecture asks if for every $x \in \Pi$ there exists an integer, n, such that $T^n(x) = 1$. The Statistical $(3x + 1)$-Conjecture asks the same question, except for a subset of Π of density 1. The Structure Theorem proven in [S] shows that infinity is in a sense a repelling point, giving some reasons to expect that the $(3x + 1)$-Conjecture may be true. In this paper, we present the analogous theorem for some generalizations of the $(3x + 1)$-Map, and expand on the consequences derived in [S]. The generalizations we consider are determined by positive coprime integers, d and g, with $g > d \geq 2$, and a periodic function, $h(x)$. The map T is defined by the formula $T(x) = \frac{gx + h(gx)}{d^k}$, where k is again the largest integer for which $T(x)$ is an integer. We prove an analogous Structure Theorem for (d, g, h)-Maps, and that the probability distribution corresponding to the density converges to the Wiener measure with the drift $\log g - \frac{d}{d-1} \log d$ and positive diffusion constant. This shows that it is natural to expect that typical trajectories return to the origin if $\log g - \frac{d}{d-1} \log d < 0$ and escape to infinity otherwise.

Keywords: $3x + 1$ Problem, $3n + 1$ Problem, Collatz Conjecture, Structure Theorem, (d, g, h)-Maps, Brownian Motion.

1 Introduction

1.1 The $(3x + 1)$-Map and $(3x + 1)$-Conjecture

Recall the definition of the $(3x + 1)$-Map, (see [L]). Take an integer $x > 0$, with x odd. Then $3x + 1$ divides 2, so we can find a unique $k > 0$ such that

Received 18 April 2002.

$y = \dfrac{3x+1}{2^k}$ is again odd. In this way, we get a mapping $T : x \longmapsto y$ defined on the set Π of strictly positive numbers not divisible by 2 or 3. Write $\Pi = 6\mathbb{Z}^+ + E$, where $E = \{1, 5\}$, is the set of possible congruence classes modulo 6.

For every integer, x, with $0 < x < 2^{60}$, a computer has checked that enough iterations of the $(3x + 1)$-Map eventually send x to 1 (see [L]). The natural conjecture asks if the same statement holds for all $x \in \Pi$:

Conjecture 1. **($(3x + 1)$-Conjecture).** *For every $x \in \Pi$, there is an integer n, such that $T^n(x) = 1$.*

The Statistical $(3x + 1)$-Conjecture asks the same question, except for a subset of Π of density 1.

For every x, we can associate a value, which is the k used in the definition of T. When we apply T repeatedly, we get a set of k values, called the path of x. We shall call the ordered set of positive integers, $(k_1, ..., k_m)$, the "m-path of x," denoted by $\gamma_m(x)$, if these are the k values that appear in m repeated iterations of T.

E.g. $T(17) = \dfrac{3 \cdot 17 + 1}{2^2} = 13$, so $k = 2$, and $\gamma_1(17) = (2)$. $T^2(17) = T(13) = \dfrac{3 \cdot 13 + 1}{2^3} = 5$, so here $k = 3$, and thus $\gamma_1(13) = (3)$, and $\gamma_2(17) = (2, 3)$.

Assume that we are given an m-path, $(k_1, ..., k_m)$. We can ask the following question: what is the set of $x \in \Pi$ for which $\gamma_m(x) = (k_1, ..., k_m)$?

The answer is given by the so-called Structure Theorem, proven in [S]. The theorem states that if $x \in \Pi$ has $\gamma_m(x) = (k_1, ..., k_m)$, then the next value in Π which will have the same m-path and congruence class modulo 6 is $x + 6 \cdot 2^{k_1+...+k_m}$. In other words, there is some first $x \in \Pi = 6\mathbb{Z}^+ + E$, call it x_0, which has $\gamma_m(x) = (k_1, ..., k_m)$. Writing $x_0 = 6 \cdot q + \varepsilon$, with $\varepsilon \in E$, we get all x with the same ε and m-path from the sequence $x_p = 6\left(2^{k_1+...+k_m} p + q\right) + \varepsilon$. The theorem tells us how to solve uniquely for q given the m-path and ε, and shows that $q < 2^{k_1+...+k_m}$, so the representation of x_p is unique.

E.g. Let $k_1 = 2$, $k_2 = 3$, and $\varepsilon = 5$. Then $x_0 = 17 = 6\left(2^5 \cdot 0 + 2\right) + 5$, and we know that $\gamma_2(17) = (2, 3)$. Look at $x_1 = 6\left(2^5 \cdot 1 + 2\right) + 5 = 209$: $T(209) = 157$, with $k_1 = 2$, and $T^2(209) = 59$ with $k_2 = 3$, so $\gamma_2(209) = (2, 3)$. We can verify that there are no elements of Π between 18 and 208 that are congruent 5 modulo 6 and have the 2-path $(2, 3)$.

Moreover, the Structure Theorem tells us that if the image of x_0 is $y_0 = T^m(x_0) = 6 \cdot r + \delta$, with $\delta \in E$ (since y_0 is also in Π), then we get the next image by adding $6 \cdot 3^m$. In other words, if y_p is the image of x_p, then $y_p = T^m \left(x_p\right) = 6\left(3^m p + r\right) + \delta$. The theorem also solves explicitly for r and δ given the m-path and ε, and finds that $r < 3^m$.

E.g. $T^2\left(17\right) = 5 = 6\left(3^2 \cdot 0 + 0\right) + 5$, and $T^2\left(209\right) = 59 = 6\left(3^2 \cdot 1 + 0\right) + 5$.

The Structure Theorem also shows that infinity is in a sense a repelling point. This gives some reasons to expect that the $(3x + 1)$-Conjecture may be true.

In this paper, we present the analogous theorem for some generalizations of the $(3x + 1)$-Map, and expand on the consequences derived in [S].

1.2 The (d, g, h)-Maps and (d, g, h)-Problem

The generalizations we consider are a particular case of maps proposed in [FR]. They are determined by positive coprime integers, d and g, with $g > d \geq 2$, and a periodic function, $h\left(x\right)$, satisfying:

1. $h\left(x + d\right) = h\left(x\right)$,
2. $x + h\left(x\right) \equiv 0 \pmod{d}$,
3. $0 < \left|h\left(x\right)\right| < g$.

The map T is defined by the formula

$$T\left(x\right) = \frac{gx + h\left(gx\right)}{d^k},$$

where k is uniquely chosen so that the result is not divisible by d. Property 2 of h guarantees $k \geq 1$. The natural domain of this map is the set Π of positive integers not divisible by d and g. Let E be the set of integers between 1 and dg that divide neither d nor g, so we can write $\Pi = dg\mathbb{Z}^+ + E$. The size of E can easily be calculated: $|E| = \left(d - 1\right)\left(g - 1\right)$.

In the same way as before, we have m-paths, which are the values of k that appear in iterations of T, and we again denote them by $\gamma_m\left(x\right)$.

The original problem corresponds to $g = 3$, $d = 2$, and $h\left(1\right) = 1$. The $(3x - 1)$-problem corresponds to $g = 3$, $d = 2$, and $h\left(1\right) = -1$. The $(5x + 1)$-problem corresponds to $g = 5$, $d = 2$, and $h\left(1\right) = 1$, and so on.

The Structure Theorem for (d, g, h)-Maps will be slightly different, in that given an m-path, $\left(k_1, ..., k_m\right)$, and congruence class, ε, modulo dg, we do not have a unique x_0. Instead, we have $\left(d - 1\right)^m$ values of what was x_0 in the original case, which we will denote by $x_0^{(i)}$, with $i = 1, \dots, (d-1)^m$. Each of these can be written as $x_0^{(i)} = dg \cdot q^{(i)} + \varepsilon$, with $q^{(i)} < d^{k_1+\dots+k_m}$. Then we get every x with the given m-path by adding $dg \cdot d^{k_1+\dots+k_m}$. In other words, letting

$$x_p^{(i)} = dg\left(d^{k_1+\dots+k_m} p + q^{(i)}\right) + \varepsilon,$$

we get every $x \in \Pi$ with $\gamma_m(x) = (k_1, \dots, k_m)$ and $x \equiv \varepsilon \,(\mathrm{mod}\, dg)$ in the set $\left\{x_p^{(i)}\right\}_{p \geq 0, 1 \leq i \leq (d-1)^m}$.

Here is the precise formulation of the Structure Theorem for (d, g, h)-Maps.

Theorem 2 (Structure Theorem). *Given an m-path, $(k_1, \dots, k_m)$, and $\varepsilon \in E$, let $k = k_1 + \dots + k_m$. Then there exist $(d-1)^m$ triples, $\left(q^{(i)}, r^{(i)}, \delta^{(i)}\right)$, $i = 1, \dots, (d-1)^m$, with $0 \leq q^{(i)} < d^k$, $0 \leq r^{(i)} < g^m$, and $\delta^{(i)} \in E$, such that*

$$\begin{aligned}\{x \in \Pi : x \equiv \varepsilon (\mathrm{mod}\, dg),\ \gamma_m(x) = (k_1, \dots, k_m)\} \\ = \{dg(d^k p + q^{(i)}) + \varepsilon\}_{p \geq 0, 1 \leq i \leq (d-1)^m}.\end{aligned}$$

Moreover, $T^m\left(dg\left(d^k p + q^{(i)}\right) + \varepsilon\right) = dg\left(g^m p + r^{(i)}\right) + \delta^{(i)}$.

The proof of the theorem is given in the next section.

In section 3, we prove that the probability distribution corresponding to the density converges to the Wiener measure with the drift $\log g - \frac{d}{d-1} \log d$ and positive diffusion constant. This shows that it is natural to expect that typical trajectories return to the origin if $\log g - \frac{d}{d-1} \log d < 0$ and escape to infinity otherwise. This question is discussed in more detail in section 4.

2 Proof of the Structure Theorem

The proof goes by induction on m. At each stage, we assume x has the given m-path and modulo class, and write $x = dg\left(d^k p + q\right) + \varepsilon$ and $y = T^m(x) = dg(g^m s + r) + \delta$. This can be done for any number, since we are simply writing out the modulo classes. After some algebra, we come to some equation for the triplets (q, r, δ), and show that it has $(d-1)^m$ solutions.

2.1 Case $m = 1$

Say we are given a 1-path, (k), and let us take an $\varepsilon \in E$. Write $x = dg \cdot t + \varepsilon$, and assume that x has the 1-path, (k). One can further break t into the form: $t = d^k p + q$, with $0 \leq q < d^k$. Let $y = T(x)$, so by our assumption, $d^k y = gx + h(gx)$. By periodicity, $h(gx) = h(g\varepsilon)$, so since ε is fixed, h does not depend on x, and is fixed. Thus we will write just h for $h(gx)$ from now on. Since $y \in \Pi$, we can write $y = dg \cdot t' + \delta$ for some $\delta \in E$, and expand $t' = g \cdot s + r$, for $0 \leq r < g$. The first step of our analysis is to show that $s = p$. We write $gx + h = d^k y$, and substitute for x, y, t, and t':

$$g\left(dg \cdot \left(d^k p + q\right) + \varepsilon\right) + h = d^k \left(dg \cdot (g \cdot s + r) + \delta\right).$$

We expand this to see:

$$g^2 d^{k+1} \cdot p + \left(dg^2 q + g\varepsilon + h\right) = g^2 d^{k+1} \cdot s + \left(d^{k+1} gr + d^k \delta\right). \qquad (1)$$

Next, we apply the following simple Lemma.

Lemma 3. *If $a \cdot b + c = a \cdot b' + c'$ with $0 \leq c, c' < a$, then $b = b'$ and $c = c'$.* □

To apply the lemma (with $a = g^2 d^{k+1}$), we need to show that the parts in parentheses on both sides of (1) are contained in $\left[0, g^2 d^{k+1} - 1\right]$. We will derive upper and lower bounds for the left side, and leave similar calculations for the right side to the reader.

Consider the lower bound of the left side. Since $q \geq 0$, $\varepsilon \geq 1$ and $h \geq -g+1$ (by Condition 3), we have that

$$dg^2 \cdot q + g\varepsilon + h \geq g \cdot 1 + (-g + 1) = 1,$$

and thus is positive.

For the upper bound of the left side, we notice that $q \leq d^k - 1$, $\varepsilon \leq dg - 1$ (since $\varepsilon \in E$) and $h \leq g - 1$. So

$$\begin{aligned} dg^2 \cdot q + g\varepsilon + h &\leq g^2 d \cdot \left(d^k - 1\right) + g(dg - 1) + (g - 1) \\ &= g^2 d^{k+1} - 1. \end{aligned}$$

- The Lemma gives us that $p = s$, and from now on we write just p. We want to characterize q, r and δ, showing that they are independent of p.

To continue, we recall that the Lemma implies that the parts in parentheses of (1) also concur. So:

$$g^2 d \cdot q + g\varepsilon + h = d^k gd \cdot r + d^k \delta. \qquad (2)$$

The next step is to break δ into $\delta = \delta' g + \delta''$, with $0 \leq \delta'' < g$. Since $\delta \in E$, we have $\delta < dg$, implying $0 \leq d' < d$. We now look at (2) modulo g to solve for δ'':

$$d^k \delta'' = h \,(\bmod\, g)\,. \tag{3}$$

Since g and d are relatively prime, d^k has a multiplicative inverse in $(\mathbb{Z} \setminus g\mathbb{Z})^*$, meaning δ'' is uniquely determined. Exactly one of the d possible values of δ' will make $\delta = \delta' g + \delta''$ divisible by d, and we throw this value away since $\delta \in E$.

This leaves us with $d - 1$ possible values for δ, which we denote by $\delta^{(1)}, \delta^{(2)}, \ldots, \delta^{(d-1)}$. It suffices to solve (2) uniquely for $q^{(i)}$ and $r^{(i)}$ given $\delta^{(i)}$.

Now we assume we have fixed $\delta^{(i)}$, and rearrange (2), adding a superscript to q and r to correspond to δ:

$$g \cdot q^{(i)} - d^k r^{(i)} = \frac{d^k \delta^{(i)} - g\varepsilon - h}{dg} = v.$$

Everything on the right hand side is known, so v is now just an integer (and independent of p). We solve for $q^{(i)}$ and $r^{(i)}$ by applying the Chinese Remainder Theorem to the equation $g \cdot a - d^k b = 1$, then setting $q^{(i)} = v \cdot a (\bmod\, d^k)$ and $r^{(i)} = v \cdot b (\bmod\, g)$.

- Having found the triplets $(q^{(i)}, r^{(i)}, \delta^{(i)})$, we are done with the case $m = 1$.

Summarizing the first step of the induction, we pick some $\varepsilon \in E$, assume $x \in \Pi$ is of the form $dg \cdot t + \varepsilon$, and write $t = d^k p + q$. Under the same assumptions for the image, $y = T(x)$, we write $y = dg \cdot t' + \delta$ and $t' = gp + r$. We find that δ is unique modulo g, and there are $d - 1$ values, $\delta^{(1)}, \ldots, \delta^{(d-1)}$, which $\delta \in E$ may take. For each one, we solve for $q^{(i)}$ and $r^{(i)}$. All of the calculations depend only on k and ε.

2.2 Induction on $m > 1$

For $m > 1$, the induction goes as follows. To know which x have a given m-path, $(k_1, \ldots, k_m)$, we first assume we know the answer for the $(m-1)$-path, $(k_1, \ldots, k_{m-1})$.

Let $k = k_1 + k_2 + \ldots + k_{m-1}$, and assume by the induction hypothesis that there are $(d-1)^{m-1}$ values for the triplet $(q_{m-1}, r_{m-1}, \delta_{m-1})$ which satisfy our equations. Fix one such triplet, pick any integer, p_{m-1}, and set $x =$

$dg(d^k p_{m-1} + q_{m-1}) + \varepsilon$, and $y = dg(g^{m-1} p_{m-1} + r_{m-1}) + \delta_{m-1}$. Then we have $\gamma_m(x) = (k_1, ..., k_{m-1})$, and $y = T^{m-1}(x)$. Here we write p_{m-1} instead of just p to distinguish from the p we will have in the next paragraph. The triplet $(q_{m-1}, r_{m-1}, \delta_{m-1})$ is still gotten independently of p_{m-1}.

We can alternatively break x into $x = dg(d^{k+k_m} p_m + q_m) + \varepsilon$ for some $q_m < d^{k+k_m}$ and also write $z = T^m(x) = T(y) = dg \cdot t + \delta_m$, with $t = g^m s + r_m$. The key idea is to find the $d-1$ possible values for $\delta_m \in E$, and with each we solve for the corresponding q_m and r_m, knowing q_{m-1}, r_{m-1}, and δ_{m-1}. We will again see that $p_m = s$ and that (q, r, δ) do not depend on this value.

Since $z = T(y)$, by assumption, we have $d^{k_m} z = gy + h(gy)$, (again let $h = h(gy) = h(g\delta_{m-1})$) which expands to:

$$d^{k_m+1} g^{m+1} s + d^{k_m+1} g r_m + d^{k_m} \delta_m = dg^{m+1} p_{m-1} + g^2 d r_{m-1} + g\delta_{m-1} + h. \tag{4}$$

Remembering the two expressions for x, and setting $p_m = d^{k_m} p_1 + p_2$ (with $0 \leq p_2 < d^{k_m}$), we write:

$$\begin{aligned} d^{k+k_m} p_m + q_m &= \frac{x - \varepsilon}{dg} = d^k p_{m-1} + q_{m-1} \\ &= d^{k+k_m} p_1 + d^k p_2 + q_{m-1}. \end{aligned}$$

We easily see that $0 \leq d^k p_2 + q_{m-1} < d^{k+k_m}$, so we again use the Lemma to find:

$$p_m = p_1, \tag{5}$$
$$q_m = d^k p_2 + q_{m-1}. \tag{6}$$

Returning to (4), we expand:

$$\begin{aligned} d^{k_m+1} g^{m+1} s + \left(d^{k_m+1} g r_m + d^{k_m} \delta_m\right) = d^{k_m+1} g^{m+1} p_1 + \\ + \left(dg^{m+1} p_2 + g^2 d r_{m-1} + g\delta_{m-1} + h\right). \end{aligned}$$

Following the same techniques as before, we bound the parts in parentheses on both sides between zero and $d^{k_m+1} g^{m+1}$, and apply the Lemma. This gives us that $p_m = p_1 = s$, and that

$$d^{k_m+1} g r_m + d^{k_m} \delta_m = dg^{m+1} p_2 + g^2 d r_{m-1} + g\delta_{m-1} + h. \tag{7}$$

Again looking modulo g and setting $\delta_m = \delta' g + \delta''$, we solve:

$$\delta'' \equiv g\delta_{m-1} + h \,(\operatorname{mod} g)\,,$$

which again gives us d choices for δ', one of which we throw out because $\delta_m \in E$. Rearranging (7), we get:

$$g^m p_2 - d^{k_m} r_m = \frac{d^{k_m}\delta_m - dg^2 r_{m-1} - g\delta_{m-1} - h}{dg} = v$$

From here, we solve $g^m a - d^{k_m} b = 1$ and set $p_2 = a \cdot v \left(\operatorname{mod} d^{k_m}\right)$ and $r_m = b \cdot v \,(\operatorname{mod} g^m)$, so $q_m = d^k p_2 + q_{m-1}$. We have $(d-1)$ values of (q_m, r_m, δ_m) derived from $(d-1)^{m-1}$ values of $(q_{m-1}, r_{m-1}, \delta_{m-1})$, so there are a total of $(d-1)^m$ triplets, consistent with the induction hypothesis. Now everything in the triplet (q_m, r_m, δ_m) is defined, and we are done. □

3 Brownian Motion of (d, g, h)-Paths

In [FMMT],[LW], it is assumed that the $(3x+1)$-Map behaves as a geometric Brownian motion, and a stochastic model is built from which other conjectures relating to the problem are derived. Here, we prove that the generalized (d, g, h)-Maps do indeed have this behavior.

In order to consider sample (d, g, h)-paths, we must first establish a version of a probability measure on $\mathbb{Z}^+$. The only natural way to do this is through density:

Definition 4. *For $A \subset \mathbb{Z}^+$, define*

$$P\,(A) = \lim_{n\to\infty} \frac{|A \cap [1, n] \cap \Pi|}{|[1, n] \cap \Pi|} = \lim_{n\to\infty} \frac{|A \cap [1, n] \cap \Pi|}{n} \cdot \frac{dg}{|E|}, \qquad (8)$$

provided the limit exists.

A nice consequence of the Structure Theorem is that if we want to consider the set of x that follow a certain m-path, they all fall in one of several arithmetic progressions, and so these sets have a density.

Partition the interval $[0, 1]$ by: $0 = t_0 < t_1 < \ldots < t_r = 1$. Fix m and let $m_i = \lfloor t_i m \rfloor$. For any x, let $x_i = T^{m_i}(x)$.

Theorem 5. *The properly normalized path* $\ln x_i$ *converges as* $m \to \infty$ *to a Brownian Path with drift* $\ln g - \frac{d}{d-1}\ln d$. *More precisely,*

$$\lim_{m\to\infty} P\left\{x : a_i < \frac{\ln x_{i+1} - \ln x_i - (m_{i+1} - m_i)\left(\ln g - \frac{d}{d-1}\ln d\right)}{\sqrt{\frac{d}{(d-1)^2} m \ln d}} < b_i,\right.$$

$$\left. \text{with } i = 0, \ldots, r-1\right\}$$

$$||$$

$$\int_{a_0}^{b_0}\int_{a1}^{b_1}\cdots\int_{a_{r-1}}^{b_{r-1}} \frac{e^{\left(-\frac{1}{2}\sum_{i=0}^{r-1} u_i^2\right)}}{(2\pi)^{\frac{r}{2}}} du_0 du_1 \cdots du_{r-1}.$$

Proof. By an extension of the Structure Theorem, we know that $x_i = T^{m_i}(x)$ can be expressed as $x_i = dg(g^{m_i} d^{k_{m_i+1}+\ldots+k_m} p + q_i) + \delta_i$. Then

$$\ln x_i = m_i \ln g + \left(k_{m_i+1} + \ldots + k_m\right)\ln d + \ln p + O(1), \tag{9}$$

and since we are interested in questions about density, x_i is large, so p is large, and thus $O(1)$ is non-essential. Then we can rearrange (9) to:

$$\begin{aligned}\ln x_i - m_i \ln g - \left(k_{m_i+1} + \ldots + k_m\right)\ln d &= \ln p \\ &= \ln x_{i+1} - m_{i+1}\ln g - \\ &\quad \left(k_{m_{i+1}+1} + \ldots + k_m\right)\ln d,\end{aligned}$$

from which we get:

$$\begin{aligned}&(m_{i+1} - m_i)\frac{d}{d-1}\ln d - \left(k_{m_i+1} + \ldots + k_{m_{i+1}}\right)\ln d = \\ &\quad = \ln x_{i+1} - \ln x_i - (m_{i+1} - m_i)\left(\ln g - \frac{d}{d-1}\ln d\right).\end{aligned} \tag{10}$$

Since the set of x_i consists of precisely $(d-1)^i$ arithmetic progressions, each with step $dg \cdot d^k$ (where $k = k_1 + \ldots + k_m$), we use (8) to find that

$$P\left\{\gamma_m(x) = (k_1, \ldots, k_m), x \equiv \varepsilon \,(\mathrm{mod}\, dg)\right\} = \frac{1}{dg \cdot d^k}\frac{dg}{|E|}(d-1)^m.$$

This holds for each $\varepsilon \in E$, so we see that

$$P\left\{\gamma_m(x) = (k_1, ..., k_m)\right\} = |E| \cdot P\left\{\gamma_m(x) = (k_1, ..., k_m), x \equiv \varepsilon \,(\mathrm{mod}\, dg)\right\}$$
$$= \frac{(d-1)^m}{d^k} = \prod_{j=1}^{m} \frac{(d-1)}{d^{k_j}}. \quad (11)$$

This shows that we can consider the k_j as independent identically distributed random variables, with exponential distribution having the parameter $\frac{1}{d}$. Thus the expected value,

$$\begin{aligned} E[k_1 + ... + k_m] &= \sum_{n \geq m} n \cdot P\{k_1 + ... + k_m = n\} \\ &= \sum_{n \geq m} n \cdot \sum_{s_1 + ... + s_m = n - m,\, s_i \geq 0} P\{(s_1 + 1, ..., s_m + 1)\} \\ &= (d-1)^m \sum_{n \geq m} n \sum_{s_1 + ... + s_m = n - m,\, s_i \geq 0} \frac{1}{d^n} \\ &= (d-1)^m \sum_{n \geq m} \frac{n}{d^n} \binom{n-1}{m-1} \\ &= \frac{d}{d-1} m. \end{aligned}$$

Similarly, we can calculate that $Var[k_1 + ... + k_m] = \frac{d}{(d-1)^2} m$. So by the Central Limit Theorem,

$$\lim_{m \to \infty} P\left\{ \frac{k_1 + ... + k_m - \frac{d}{d-1} m}{\sqrt{\frac{d}{(d-1)^2} m}} \in (a, b) \right\} = \int_a^b \frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}} du.$$

And by (10), we have that

$$P\left\{ \frac{\ln x_{i+1} - \ln x_i - (m_{i+1} - m_i)\left(\ln g - \frac{d}{d-1} \ln d\right)}{\sqrt{m \cdot \frac{d}{(d-1)^2}} \ln d} \in (a_i, b_i) \right\} =$$

$$||$$

$$= P\left\{ \frac{\frac{d}{d-1}(m_{i+1} - m_i) - \left(k_{m_i+1} + ... + k_{m_{i+1}}\right)}{\sqrt{\frac{d}{(d-1)^2} m}} \in (a_i, b_i) \right\},$$

which converges exactly as claimed. Since the k_i are independent, the increments, $\ln x_{i+1} - \ln x_i$ are as well, and we have the statement about the convergence of our distributions to the Wiener measure. □

4 Asymptotic Behavior of Typical Trajectories

The previous section proves that the probability distribution corresponding to the density converges to the Wiener measure with drift $\log g - \frac{d}{d-1} \log d$. Since d and g are relatively prime, there are no values of d and g for which $\log g - \frac{d}{d-1} \log d = 0$, and thus every (d, g, h)-Map has a non-trivial drift. Therefore, the asymptotic behavior of typical trajectories depends entirely on the sign of the drift. When the drift is negative, infinity is a repelling point. In the opposite case, typical trajectories escape to infinity. For the original $(3x + 1)$-Map, the drift is $\log 3 - 2 \log 2 < 0$, and so as a special case, we get the result found in [S].

In the literature, the stopping time of an integer x is defined as the first positive integer, n, such that $T^n(x) < x$. If n does not exist, we say that x has an infinite stopping time. In [E] and [T76], [T79], it is independently proven that for the $(3x + 1)$-Map, the density of integers with a finite stopping time is 1. This paper provides another proof of this statement.

Acknowledgments. The first author thanks L. Kontorovich and S. Payne for discussions and criticism. The second author thanks the NSF for financial support, grant DMR-9813268.

References

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

[FMMT] M.R. Feix, A. Muriel, D. Merlini and R. Tartini, *The* $(3x + 1)/2$ *problem: A Statistical Approach*, in: Stochastic Processes, Physics and Geometry II, Locarno 1991. World Scientific (1995), 289–300.

[FR] M.R. Feix and J.L. Rouet, *The* $(3x + 1)/2$ *problem and its generalization: a stochastic approach*, Proceedings of the International Conference on the Collatz Problem and Related Topics (2001).

[L] J.C. Lagarias, *The* $3x + 1$ *Problem and Its Generalizations*, American Mathematical Monthly, Vol. **92**, Issue 1 (Jan., 1985), 2–23.

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

[S] Ya. G. Sinai, *Statistical* $(3x+1)$-*Problem*, (2002), preprint.

[T76] R. Terras, *A stopping time problem on the positive integers*, Acta Arith. **30** (1976), 241–252.

[T79] R. Terras, *On the existence of a density*, Acta Arith. **35** (1979), 101–102.

A. V. Kontorovich
Mathematics Department of Princeton University
Princeton, NJ, 08544
USA

E-mail: alexk@alumni.princeton.edu.

Ya. G. Sinai
Mathematics Department of Princeton University
Princeton, NJ, 08544
USA

E-mail: sinai@math.princeton.edu.