AN ITERATED RANDOM FUNCTION WITH LIPSCHITZ NUMBER ONE

AARON ABRAMS, HENRY LANDAU, ZEPH LANDAU, JAMES POMMERSHEIM, AND ERIC ZASLOW

ABSTRACT. Consider the set of functions $f_\theta(x) = |\theta - x|$ on $\mathbb{R}$. Define a Markov process that starts with a point $x_0 \in \mathbb{R}$ and continues with $x_{k+1} = f_{\theta_{k+1}}(x_k)$ with each θ_{k+1} picked from a fixed bounded distribution μ on $\mathbb{R}^+$. We prove the conjecture of G. Letac that if μ is not supported on a lattice, then this process has a unique stationary distribution π_μ and any distribution converges under iteration to π_μ (in the weak-* topology). We also give a bound on the rate of convergence in the special case that μ is supported on a two-point set. We hope that the techniques will be useful for the study of other Markov processes where the transition functions have Lipschitz number one.

1. INTRODUCTION

In their recent paper [D-F], P. Diaconis and D. Freedman describe a general method for studying Markov chains. This consists of viewing a state transition of the chain as the action of a function chosen at random from a family of state transformations; the Markov process can then be expressed as an iteration of these functions. When the state space is a metric space, e.g. $\mathbb{R}$, they show that if the functions exhibit a certain average contractivity (described in terms of Lipschitz numbers), then running the iteration backward rather than forward produces geometric convergence to a unique stationary distribution for the chain. Many applications of this idea are developed.

An example of a Markov process not immediately amenable to this method, originally proposed for study by Letac [L], is defined on the non-negative reals $\mathbb{R}^+ = [0, \infty)$ by starting with a point $x_0 \in \mathbb{R}^+$ and for $k = 0, 1, \ldots$, setting

$$x_{k+1} = f_{\theta_{k+1}}(x_k) \equiv |\theta_{k+1} - x_k|, \tag{1}$$

where the $\theta_k \in \mathbb{R}^+$ are chosen independently from a fixed distribution μ (on $\mathbb{R}^+$). If we choose the starting point x_0 from a distribution π_0, then the above Markov process defines a distribution π_n for x_n.

We will consider the case that μ has finite expectation; it is known (e.g. [F], p. 208) that the distribution

$$\pi_\mu(x) = \frac{1}{E(\theta)} \int_0^x \Pr\{\theta > y\} dy \tag{2}$$

is a stationary probability distribution for this process (where θ is chosen according to the distribution μ and $E(\theta)$ is its expectation). G. Letac conjectured in [L] that if μ is not supported on a lattice (that is, the set of integer multiples of a fixed real number), then π_μ is the unique stationary distribution, and that the distributions

Date: October 25, 2001.

π_n of x_n should converge to π_μ in the weak-* topology. Both these conjectures would follow from Theorem 1 of [D-F] if the f_θ's had Lipschitz numbers less than one (on average).

In this paper we prove Letac's conjecture for compactly supported μ.

Theorem 3.5: *Let μ be a distribution on a bounded interval $Y \subset \mathbb{R}^+$, with μ not supported on a lattice. Then π_μ is the unique stationary distribution for the process (1). Moreover, if x_0 is chosen according to an initial distribution π_0 then the distributions π_n of x_n converge to π_μ in the weak-* topology.*

We prove this theorem in Sections 2 and 3, and in Sections 4 and 5 we analyze more carefully the special case of μ supported on just two points. In this case we provide a second independent proof of Letac's conjecture. Our hope is that the techniques will be useful for the study of other Markov processes where the transition functions have Lipschitz number one.

After completing this work the authors learned of a paper by J. Mattingly [M] addressing similar questions; that paper obtains results similar to ours.

Let μ be a probability measure on a space Y and let $\{f_\theta : \theta \in Y\}$ be a set of maps from a metric space X into itself. The *iterated random function induced by μ* is the Markov process of repeatedly picking $\theta \in Y$ according to μ and then applying f_θ. For $x \in X$, and $\Theta = (\theta_1, \theta_2, \dots)$ with $\theta_i \in Y$, let $B_n(\Theta, x)$ be the nth *backward iterate* of x and $F_n(\Theta, x)$ the nth *forward iterate* of x; that is,

$$\begin{aligned} B_n(\Theta, x) &= f_{\theta_1} \circ \cdots \circ f_{\theta_n}(x), \\ F_n(\Theta, x) &= f_{\theta_n} \circ \cdots \circ f_{\theta_1}(x). \end{aligned}$$

Letac observed that for fixed n and variable Θ, the distributions of $B_n(\Theta, x)$ and $F_n(\Theta, x)$ are identical; yet for fixed Θ, as n tends to infinity the sequences $B_n(\Theta, x)$ and $F_n(\Theta, x)$ behave quite differently. He used this to prove the following general principle [L].

Letac's Principle [L]**:** *If the backward iterates $B_n(\Theta, x)$ converge for almost all Θ to a limit $B(\Theta)$ which is independent of the starting point $x \in X$, then the distribution π_μ on X induced*[1] *by the map $\Theta \mapsto B(\Theta)$ is the unique stationary distribution for the iterated random function induced by μ. Moreover, if x_0 is chosen according to an initial distribution π_0, then the distributions π_n of x_n converge to π_μ in the weak-* topology.*

In Section 2 we prove the following general result which implies the hypotheses of Letac's principle. Note that this theorem requires checking a property of only the *forward* iterates.

Theorem 2.2: *Consider a probability measure μ on a space Y and a set of functions $\{f_\theta : \theta \in Y\}$ on a complete metric space X. Suppose there exists a subset $S \subset X$ such that*

(i) *$f_\theta(S) \subset S$ almost surely for θ picked from μ, and*
(ii) *$\lim_{n\to\infty} diam(F_n(\Theta, S)) = 0$ almost surely, where $\Theta = (\theta_1, \theta_2, \ldots)$ and the θ_i are picked independently from μ.*

[1] For any set $S \subset X$, the induced probability distribution $\pi_\mu(S)$ is the measure of the set $\{\Theta \mid B(\Theta) \in S\}$.

Then the backward iterates applied to an element of S converge almost surely and the limit is independent of the starting point (in S).

In Section 3 we establish the hypotheses of Theorem 2.2, completing our proof of Letac's conjecture.

Our second proof of Letac's conjecture applies only in the special case that μ is uniform on a two-point set $\{\alpha,\beta\}$; however the argument is independent of Letac's principle. Our method is to analyze the iterated random function by studying the orbit of a point; we begin this study in Section 3 and continue in more detail in Section 4. This technique leads to a bound (Theorem 5.4) on the rate of convergence of the distributions π_n in this case.

2. A General Result

In this section we prove a general theorem which lets us conclude that the backward iterates converge by looking only at the forward iterates.

Let μ be a probability measure on a set Y and $\{f_\theta : \theta \in Y\}$ a set of functions from a complete metric space X into itself. Denote by $\Theta = (\theta_1, \theta_2, \dots)$ an element of $Y^\infty = Y \times Y \times \cdots$ and by $\mu^\infty = \mu \times \mu \times \cdots$ the infinite product measure on Y^∞. The following lemma is used in the proof of Theorem 2.2 to move from considering backward iterates to considering forward iterates.

Lemma 2.1. *Let $\Gamma \subset Y^\infty$, and let $k_1, \dots, k_n$ be a finite set of distinct indices. Let*

$$\Gamma' = \{\Theta' : \textit{there exists a } \Theta \in \Gamma \textit{ such that } \theta'_{k_j} = \theta_j,\ j = 1 \dots n\},$$

where $\Theta = (\theta_1, \theta_2, \dots)$ and $\Theta' = (\theta'_1, \theta'_2, \dots)$. Then $\mu^\infty(\Gamma) \le \mu^\infty(\Gamma')$.

Proof. Let σ be any finite permutation of $\mathbb{N}$ with $\sigma(j) = k_j$ for $j = 1, \dots, n$. (A finite permutation is one with $\sigma(j) = j$ for all j sufficiently large.) Since μ^∞ is invariant under the action of finite permutations, we have $\mu^\infty(\Gamma) = \mu^\infty(\sigma(\Gamma))$. But $\sigma(\Gamma) \subset \Gamma'$, so the result follows. □

For $S \subset X$, set $B_n(\Theta, S) = \{B_n(\Theta, x) : x \in S\}$ and $F_n(\Theta, S) = \{F_n(\Theta, x) : x \in S\}$. Let $\operatorname{diam}(S) = \sup_{x,y\in S} d(x,y)$ be the diameter of S (where d denotes the distance function on X).

Theorem 2.2. *Suppose there exists a subset S of X such that*

(i) *$f_\theta(S) \subset S$ for almost every θ picked from μ,*
(ii) *$\lim_{n\to\infty} diam(F_n(\Theta, S)) = 0$ for almost every Θ picked from μ^∞.*

Then for all $x \in S$, the backward iterates $B_n(\Theta, x)$ converge for almost all Θ, and the limit is independent of x (in S).

Proof. Fix $\epsilon > 0$. Set $\Gamma^i_N = \{\Theta : \operatorname{diam}(F_n(\Theta, S)) < 2^{-i}$ for all $n \ge N\}$. Since $\Gamma^i_N \subset \Gamma^i_{N+1}$ and $\mu^\infty(\cup_n \Gamma^i_n) = 1$ (by hypothesis (ii)), there exists $N(i)$ such that $\mu^\infty(\Gamma^i_{N(i)}) > 1 - \epsilon 2^{-i}$. Applying Lemma 2.1 to the set $\Gamma = \Gamma^i_{N(i)}$ with $k_j = N(i) + 1 - j$, $j = 1, \dots, N(i)$, we see that the measure of the set $\Lambda^i_{N(i)} = \{\Theta : \operatorname{diam}(B_{N(i)}(\Theta, S)) < 2^{-i}\}$ is greater than or equal to $\mu^\infty(\Gamma^i_{N(i)}) > 1 - \epsilon 2^{-i}$. Hypothesis (i) implies that for $n \ge N$, $B_n(\Theta, S) \subset B_N(\Theta, S)$ for almost all Θ; hence

$$(3)\quad \mu^\infty\{\Theta : \operatorname{diam}(B_n(\Theta, S)) < 2^{-i} \text{ and } B_n(\Theta, S) \subset B_{N(i)}(\Theta, S) \text{ for all } n \ge N(i)\}$$

$$= \mu^\infty(\Lambda^i_{N(i)}) > 1 - \epsilon 2^{-i}.$$

Letting

$$\Omega = \bigcap_i \{\Theta : \mathrm{diam}(B_n(\Theta, S)) < 2^{-i} \text{ and } B_n(\Theta, S) \subset B_{N(i)}(\Theta, S) \text{ for all } n \geq N(i)\},$$

(3) implies $\mu^\infty(\Omega) \geq 1 - \sum_i \epsilon 2^{-i} = 1 - \epsilon$. It follows from the definition of Ω that the $\{B_n(\Theta, x)\}_n$ are Cauchy sequences (i.e. the backward iterates converge) for every $x \in S$, $\Theta \in \Omega$. Furthermore, since for $x, y \in S$, $|B_n(\Theta, x) - B_n(\Theta, y)| \leq \mathrm{diam}(B_n(\Theta, S))$, we see that for $\Theta \in \Omega$, $\lim B_n(\Theta, x) = \lim B_n(\Theta, y)$. Thus we have shown that on the set Ω of measure at least $1 - \epsilon$, the backward iterates applied to an element of S converge to a limit independent of the starting point. Since ϵ is arbitrary, the result follows. □

3. Proof of Letac's Conjecture

We now return to the iterated random function described by equation (1). In this section, we consider μ supported on a bounded subset $Y = [0, b]$ of $\mathbb{R}^+$. Setting $S = [0, b']$ for any $b' \geq b$, it is clear that condition (i) of Theorem 2.2 is satisfied. The rest of this section shows that S and $\{f_\theta : \theta \in [0, b]\}$ satisfy condition (ii) of Theorem 2.2, and then applies the conclusion of Theorem 2.2 to complete the proof of Theorem 3.5.

The argument showing that S and $\{f_\theta : \theta \in [0, b]\}$ satisfy condition (ii) of Theorem 2.2 can be summarized as follows: on an interval I, f_θ acts in one of two ways. If $\theta \notin I$, then f_θ translates I; whereas if $\theta \in I$, then f_θ translates and "folds" (and in particular shrinks) I. Furthermore, for θ near the center of I, f_θ will shrink I by about half. Condition (ii) follows from showing that the process defined by equation (1) (for suitable μ) almost halves an interval infinitely many times. The key fact we will use is that the orbit of a point is dense.

3.1. **The orbit of a point.**

In this section we discuss the orbit of a point x under iterates of the two functions $f_\alpha(x) = |\alpha - x|$ and $f_\beta(x) = |\beta - x|$, where $0 < \alpha < \beta$. We use the standard notation $x \bmod y$ to mean the unique number z, $0 \leq z < y$ such that $z - x$ is an integer multiple of y. Let $x \in [0, \beta]$.

The *orbit* $\mathcal{O}_x$ *of* x is the set of points y such that $y = f_{\theta_n} \circ \cdots \circ f_{\theta_1}(x)$ for some n, with each $\theta_i \in \{\alpha, \beta\}$.

Lemma 3.1. *For each $x \in [0, \beta]$, we have $\mathcal{O}_x = \{n\alpha + \epsilon x \bmod \beta : n \in \mathbb{Z}, \epsilon = \pm 1\}$.*

Proof: Let $S_x = \{n\alpha + \epsilon x \bmod \beta : n \in \mathbb{Z}, \epsilon = \pm 1\}$. As shorthand, we represent the real number $n\alpha + \epsilon x \bmod \beta$ by the ordered pair $(n, \epsilon) \in \mathbb{Z} \times \{\pm 1\}$. Note that $f_\beta((n, \epsilon)) = \beta - (n\alpha + \epsilon x \bmod \beta) = (-n, -\epsilon)$. Also,

$$f_\alpha((n, \epsilon)) = \begin{cases} (n - 1, \epsilon) & \text{if } (n, \epsilon) \geq \alpha, \\ \alpha - (n\alpha + \epsilon x \bmod \beta) = (-(n - 1), -\epsilon) & \text{if } (n, \epsilon) < \alpha; \end{cases}$$

thus the set S_x is closed under f_θ for each $\theta \in \{\alpha, \beta\}$ and we have $\mathcal{O}_x \subseteq S_x$.

To show $S_x \subseteq \mathcal{O}_x$, it suffices to show that

(1) $(n, \epsilon) \in \mathcal{O}_x \implies (-n, -\epsilon) \in \mathcal{O}_x$, and
(2) $(n, \epsilon) \in \mathcal{O}_x \implies (n - 1, \epsilon) \in \mathcal{O}_x$.

The first claim follows by applying f_β. The second follows by applying f_α if $n\alpha+\epsilon x \bmod \beta \geq \alpha$, or by first applying f_α and then applying f_β if $n\alpha+\epsilon x \bmod \beta < \alpha$. □

Corollary 3.2. *Let $0 < \alpha < \beta$. If α/β is irrational then the orbit of any point under iteration of the functions f_α, f_β is dense in $[0,\beta]$. If $\alpha/\beta = p/q$ with $p, q \in \mathbb{N}$ relatively prime, then the orbit of any point intersects each closed subinterval of $[0,\beta]$ of length at least β/q.*

Proof. If α/β is irrational then the set $\{n\alpha \bmod \beta\}$ is dense in $[0,\beta]$, so the orbit of a point is also dense in $[0,\beta]$. If $\alpha/\beta = p/q$ then the orbit of x contains the set $\{n\alpha + \epsilon x \bmod \beta\} = \{\frac{k}{q}\beta + x \bmod \beta | k = 0, \ldots, q-1\}$ which intersects each closed subinterval of $[0,\beta]$ of length at least β/q. □

3.2. **Shrinking intervals.** Denote the length of an interval I by $||I||$.

The following lemma gives a criterion for recognizing when the interval $f_{\theta_n} \circ \cdots \circ f_{\theta_1}(I)$ is no bigger than about half the size of I.

Lemma 3.3. *Let $\Theta = (\theta_1, \theta_2 \ldots)$, with $\theta_i \in \mathbb{R}^+$, and let $I \subset \mathbb{R}^+$ be a bounded interval with midpoint m. Fix $\delta_1, \delta_2 > 0$. If some x with $|x - m| < \delta_1$ satisfies $F_n(\Theta, x) < \delta_2$, then*

$$||F_n(\Theta, I)|| \leq \frac{1}{2}||I|| + \delta_1 + \delta_2.$$

Proof. The functions f_θ have Lipschitz number 1, so for all $y \in I$, we have

$$|F_n(\Theta, x) - F_n(\Theta, y)| \leq |x - y| \leq \frac{1}{2}||I|| + \delta_1.$$

Since $F_n(\Theta, x) \leq \delta_2$, it follows that $F_n(\Theta, y) \leq \frac{1}{2}||I|| + \delta_1 + \delta_2$. Also $F_n(\Theta, y) \geq 0$, so the result follows. □

We say that $x \in \mathbb{R}^+$ is a *point of increase* for the probability distribution μ if $\mu(I) > 0$ for every open interval I containing x. The nonzero real numbers α and β are *irrationally related* if α/β is irrational.

Lemma 3.4. *Let μ be a probability distribution supported on a bounded interval $Y \subset \mathbb{R}^+$. Suppose that μ is not supported on a lattice. Then for any interval $I \subset \mathbb{R}^+$ and any $\epsilon > 0$, there exist $N = N(I, \epsilon)$ and intervals $J_1, \ldots, J_N$ with $\mu(J_i) > 0$, such that for all $(\theta_1, \ldots, \theta_N) \in J_1 \times \cdots \times J_N$,*

$$||f_{\theta_N} \circ \cdots \circ f_{\theta_1}(I)|| < \epsilon$$

Proof. We first assert that for any interval K, and for any $r > 0$, there exist intervals $J_1, \ldots J_n$ with $\mu(J_i) > 0$ such that for all $(\theta_1, \ldots, \theta_n) \in J_1 \times \cdots \times J_n$,

$$||f_{\theta_n} \circ \cdots \circ f_{\theta_1}(I)|| < (\frac{1}{2} + r)||K||.$$

The result then follows by repeating this shrinking process.

To prove the assertion, we will first construct the midpoints m_i of the intervals J_i. We consider two cases; as μ is not supported on a lattice, either (a) μ has two irrationally related points of increase, or (b) μ has an infinite set of points of increase which are pairwise rationally related.

Let m be the midpoint of the interval K.

First, suppose (a) holds. Let $\alpha < \beta$ be two irrationally related points of increase for μ. By Corollary 3.2 the orbit of m under iteration of the functions f_α and f_β is dense in $[0, \beta]$; thus there exists a sequence $m_1, \ldots, m_n$ with each m_i equal to α or β such that

$$f_{m_n} \circ \cdots \circ f_{m_1}(m) < \frac{r}{2}||K||. \tag{4}$$

Now suppose (b) holds. In this case we can find points of increase α, β of μ with $\alpha/\beta = p/q$, with $p, q \in \mathbb{N}$ relatively prime and q arbitrarily large. (Note that we cannot simultaneously guarantee that $\alpha < \beta$.) Find such α, β with $q > 2\beta/r||K||$. By Corollary 3.2, the orbit of m under the functions f_α and f_β intersects the closed interval $[0, \beta/q] \subset [0, r||K||/2]$. Thus again there exist $m_1, \ldots, m_n$ with each m_i equal to α or β such that (4) holds.

We have now defined the m_i in each case. Set J_i to be an interval with midpoint m_i and length no bigger than $\frac{r}{2n}||K||$. Note that J_i has positive measure because every m_i is a point of increase of μ.

To verify the claim, note that for any $(\theta_1, \ldots, \theta_n) \in J_1 \times \cdots \times J_n$, we have

$$f_{\theta_n} \circ \cdots \circ f_{\theta_1}(m) < \frac{r}{2}||K|| + \frac{nr}{2n}||K|| = r||K||,$$

since each f_θ has Lipschitz number 1. Thus we can apply the previous lemma with $x = m$, any $\delta_1 > 0$, and $\delta_2 = r||K||$.

□

Theorem 3.5. *Let μ be a distribution on a bounded interval $Y \subset \mathbb{R}^+$, with μ not supported on a lattice. Then π_μ is the unique stationary distribution for the process (1). Moreover, if x_0 is chosen according to an initial distribution π_0 then the distributions π_n of x_n converge to π_μ in the weak-* topology.*

Proof. We will show that $\{f_\theta : \theta \in Y\}$, $S = [0, b']$, $b' \geq b$ satisfy conditions (i) and (ii) of Theorem 2.2. Consequently, the backward iterates converge independent of a starting point chosen in $\bigcup_{b' \geq b}[0, b'] = \mathbb{R}^+$ and the application of Letac's principle completes the proof.

It is clear that S satisfies condition (i) of Theorem 2.2 for all $\Theta \in Y^\infty$. To establish (ii), fix $\epsilon > 0$. By Lemma 3.4 (with $I = [0, b']$), there exist intervals $J_1, \ldots, J_N$ such that

$$||f_{\theta_N} \circ \cdots \circ f_{\theta_1}(I)|| < \epsilon$$

for all $(\theta_1, \ldots \theta_N) \in J_1 \times \cdots \times J_N$. As the f_θ's do not lengthen intervals, we have $||F(\Theta, S)|| < \epsilon$ for any $\Theta = (\theta_1, \theta_2, \ldots)$ such that for some K, $\theta_{K+i} \in J_i$ for each $i = 1, \ldots, N$. Since each $\mu(J_i) > 0$, this occurs with probability 1. Since ϵ was arbitrary, condition (ii) of Theorem 2.2 is satisfied. □

4. Special Case: Two-point Distributions

Our second proof of Theorem 3.5 relies on a detailed understanding of the orbits $\mathcal{O}_x$. In this section suppose θ is chosen from the uniform distribution μ on the set $\{\alpha, \beta\}$ with $0 < \alpha < \beta$; i.e. $\Pr(\theta = \alpha) = \Pr(\theta = \beta) = 1/2$. By scaling, we may assume $\beta = 1$; we do so for the remainder of the paper. We use the notation $\langle x \rangle$ to denote the number x mod 1.

4.1. **The graph of an orbit.** Recall Lemma 3.1, and note that the proof of Lemma 3.1 (with $0 < \alpha < \beta = 1$) reveals the complete structure of the orbit of $x \leq 1$. From now on assume α is irrational. Notice that the numbers $\langle n\alpha + \epsilon x\rangle$ are distinct unless x is in the orbit of 0 (which is also the orbit of 1 and α), the orbit of $1/2$, the orbit of $\alpha/2$, or the orbit of $\frac{1+\alpha}{2}$. These four orbits we call *singular*; other orbits are *generic*.

To visualize the orbit $\mathcal{O}_x$, we associate to each point x a labeled directed graph G_x, as follows. The vertex set of G_x is $\mathcal{O}_x$, with the vertex $\langle n\alpha + \epsilon x\rangle$ labeled with the ordered pair (n, ϵ). This label is unique if $\mathcal{O}_x$ is generic. A directed edge goes from (n, ϵ) to (n', ϵ') if there exists a θ such that $f_\theta(\langle n\alpha + \epsilon x\rangle) = \langle n'\alpha + \epsilon' x\rangle$; the proof of Lemma 3.1 tells us exactly where to put these edges.

Note that if $x' \in \mathcal{O}_x$ then $G_{x'}$ looks exactly like G_x but with different labels. Also note that if $\mathcal{O}_x$ is singular then a vertex of G_x will have multiple labels due to coincidences among the numbers $\langle n\alpha + \epsilon x\rangle$; e.g. a vertex labeled (n, ϵ) in G_0 is also labeled $(n, -\epsilon)$.

We illustrate the case $\alpha = 1/\sqrt{2}$. Figure 1(a) shows the (generic) graph $G_{.2}$, and Figure 1(b) shows the (singular) graph G_0. In the singular orbit we have consolidated the pair of labels (n, ϵ) and $(n, -\epsilon)$ into the single label n. Both graphs have been drawn so that the vertical edges represent an application of f_1 and the horizontal and diagonal edges represent an application of f_α. An undirected edge indicates two oppositely oriented edges. Note that each vertex has out-degree two, since θ has two possible values.

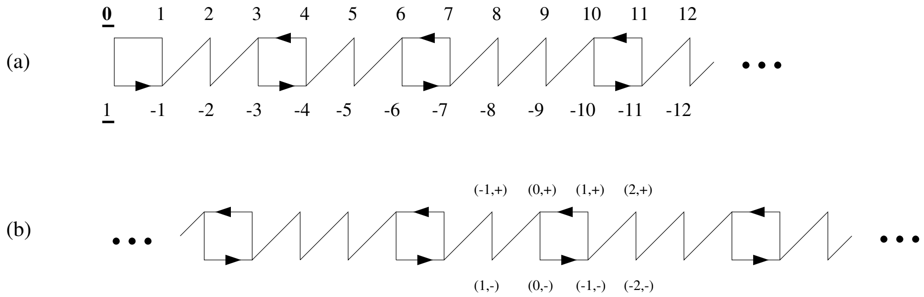

FIGURE 1. Two orbits for $\alpha = 1/\sqrt{2} > 1/2$. The orbit of $x = .2$ is generic (a), and the orbit of $x = 0$ is singular (b). In the orbit of 0, vertices are labeled only with n, since ϵ doesn't matter. The boldface labels **0** and **1** refer to the actual numbers 0 and 1.

The large-scale properties of the graphs G_x (for general irrational α) are important as well. Suppose for a moment that $\alpha > 1/2$. There are three types of vertices in a generic graph (see Figure 1):

- *small*, where $\langle n\alpha + \epsilon x\rangle < 1 - \alpha$;
- *medium*, where $1 - \alpha < \langle n\alpha + \epsilon x\rangle < \alpha$; and
- *large*, where $\langle n\alpha + \epsilon x\rangle > \alpha$.

The large vertices have in-degree one, and appear in G_x as upper-right and lower-left corners of boxes. The small vertices have in-degree three, and form the other corners of boxes. The medium vertices are the remaining ones, which are not part of any box. The fact that $\langle n\alpha\rangle$ is uniformly distributed mod 1 implies that the ratio of the numbers of small : medium : large vertices is $1-\alpha : 2\alpha - 1 : 1-\alpha$. If we write $\alpha = q(1-\alpha)+r$ with q a positive integer and $0 < r < 1-\alpha$, then the number of diagonal edges between consecutive boxes is either q or $q+1$. These numbers occur in a pattern depending on x but always in the ratio $(1-\alpha-r) : r$. For $\alpha = 1/\sqrt{2}$ we see that there are either 2 or 3 diagonal edges between consecutive boxes, this number being 2 exactly $\frac{3-4\alpha}{1-\alpha} \approx 58.6\%$ of the time.

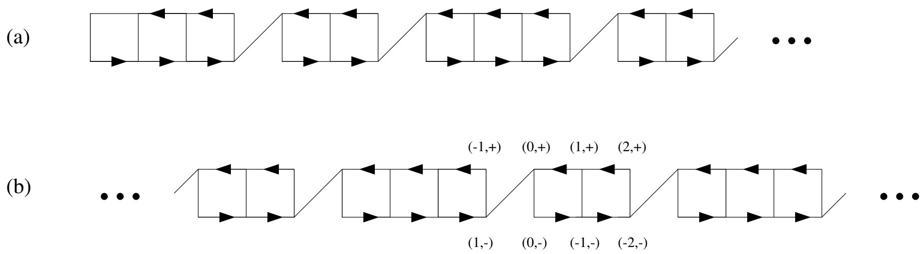

FIGURE 2. Two orbits for $\alpha = 1 - 1/\sqrt{2} < 1/2$. Again, the orbit of $x = .2$ is generic (a) and the orbit of $x = 0$ is singular (b).

In the case $\alpha < 1/2$, the graphs are similar, but we exchange the roles of α and $1-\alpha$ in the definitions of small, medium, and large. The graphs G_x in this case have more boxes and fewer diagonal edges. See Figure 2, where $\alpha = 1 - 1/\sqrt{2}$. The large vertices are the upper-right and lower-left corners of the rectangular strings of boxes; the small vertices are the other corners of the rectangles; and the medium vertices are those vertices which are part of two distinct boxes. The ratio of numbers of small : medium : large vertices is $\alpha : 1-2\alpha : \alpha$. This time writing $1-\alpha = q\alpha + r$, the number of boxes between consecutive diagonal edges is again either q or $q+1$, and these occur in the ratio $(\alpha - r) : r$.

Incidentally if α is rational then the graphs of orbits are finite, but they exhibit many properties analogous to the infinite case.

4.2. **Uniqueness of stationary distribution.** Assume α is irrational and μ is uniform on $\{\alpha, 1\}$. Recall that by (2), the stationary distribution π_μ is

$$\pi_\mu(x) = \begin{cases} 2x/(1+\alpha) & \text{if } 0 \le x \le \alpha \\ (x+\alpha)/(1+\alpha) & \text{if } \alpha \le x \le 1. \end{cases}$$

Let π be any stationary distribution for the iterated random function of equation (1) for μ. We will think of π as its distribution function, namely a right-continuous increasing function $\pi : [0,1] \to [0,1]$; thus $\pi(x)$ is the amount of mass which is concentrated on $\{y : y \le x\}$.

We will show that the values of π are determined by α; the values we compute agree with the corresponding values for π_μ.

Being right-continuous, π is determined by its values on a dense set; with the aid of Figure 2 we shall calculate $\pi(x)$ for $x \in \mathcal{O}_0 = \{\langle n\alpha \rangle\}$.

By stationarity, the mass of π is concentrated on $[0,1]$; thus $\pi(1) = 1$. We begin by establishing that π cannot have an atom (that is, a jump discontinuity) at $\langle n\alpha \rangle$ for any $n \in \mathbb{Z}$.

Lemma 4.1. *Let $0 < \alpha < 1$ be irrational, let μ be uniform on $\{\alpha, 1\}$, and let π be a stationary distribution for the associated iterated random function. Then $\pi(0) = 0$ and π is continuous at $x = \langle n\alpha \rangle$, $n = \pm 1, \pm 2, \ldots$.*

Proof. Suppose $1/2 < \alpha < 1$. If a distribution π_0 has an atom of mass m at 0, then (referring to the graph G_0) we see that π_1 will have an atom of mass $m/2$ at 1, since the only way for x_1 to equal 1 is if $x_0 = 0$ and $\theta_1 = 1$. Thus if $\pi = \pi_0$ is stationary and π has an atom of mass m at 0, then π also has an atom of mass $m/2$ at 1.

Likewise, iterating the random function shows that at each vertex of G_0 there must be an atom having mass equal to half the sum of the masses of the atoms at all incoming vertices. It follows that there must be an atom of mass $3m/2$ at α, an atom of mass $2m$ at $\langle -\alpha \rangle$, and thereafter an atom of mass $2m$ at every small and medium vertex and an atom of mass m at every large vertex.

As the total mass of a distribution must be 1, it is impossible to have $m > 0$.

Thus $\pi(0) = 0$. We now note that if π were not continuous at some point $\langle n\alpha \rangle$, then π would have an atom there; then by the same reasoning π would have an atom at 0, a contradiction.

The case $0 < \alpha < 1/2$ is similar (see e.g. Figure 2). Here an atom of mass m at 0 implies an atom of mass $m/2$ at 1, an atom of mass $3m/2$ at α, and an atom of mass m at $1 - \alpha$, and thereafter atoms of mass $2m$ at each small vertex and of mass m at each medium and large vertex. As before, m must be 0, and the result follows. □

We are now able to compute the values of $\pi(x)$ for $x \in \mathcal{O}_0$, using a technique similar to that employed in the lemma. Observe that the stationarity of π is equivalent to the condition that for each $x \in [0, 1]$ we have

$$\pi(x) = \sum_{\theta \in \{1, \alpha\}} \mu(\theta) \left(\sum_{\{y : f_\theta(y) = x\}} \pi(y) \right) \tag{5}$$

$$= \frac{1}{2} \left(\pi(1 + x) - \pi_-(1 - x) \right) + \frac{1}{2} \left(\pi(\alpha + x) - \pi_-(\alpha - x) \right), \tag{6}$$

where $\pi_-(y) \equiv \lim_{z \to y^-} \pi(z)$ (the limit exists since π is increasing). We will only need to consider the points $\langle n\alpha \rangle$; Lemma 4.1 establishes that π and π_- are equal at these points. Recall, of course, that $\pi(x) = 1$ for all $x \geq 1$ and $\pi(x) = 0$ for all $x < 0$.

Again we first treat the case $1/2 < \alpha < 1$. We already know $\pi(0) = 0$ by Lemma 4.1, and also $\pi(1) = 1$. Suppose $\pi(\alpha) = z \in [0, 1]$. Evaluating (5) at $x = \alpha$, we get $\pi(\langle -\alpha \rangle) = 2 - 2z$. Similarly, at $x = \langle -\alpha \rangle$ we get $\pi(\langle 2\alpha \rangle) = 3z - 2$, and so on. Using (5) repeatedly, we can compute $\pi(\langle n\alpha \rangle)$ for all $n \in \mathbb{Z}$. In particular, if $n > 0$ and $\langle n\alpha \rangle$ is small, then we get $\pi(\langle n\alpha \rangle) = (2n - L_n)z - (2n - 2L_n)$ where L_n is the number of large vertices among $\langle \alpha \rangle, \ldots, \langle n\alpha \rangle$. Recall that $L_n/n \to 1 - \alpha$.

We now claim that there is only one possible value of z. To see this, choose an increasing sequence of positive integers $\{n_i\}$ such that $\langle n_i\alpha \rangle \to 0$. We may assume the $\langle n_i\alpha \rangle$ are small. Since $\pi(0) = 0$ and π is right continuous, we have $0 = \lim_{i \to \infty} \pi(\langle n_i\alpha \rangle) = \lim_{i \to \infty} [(2n_i - L_{n_i})z - (2n_i - 2L_{n_i})]$ whence $z = \lim_{i \to \infty} \frac{2n_i - 2L_{n_i}}{2n_i - L_{n_i}} = 2\alpha/(1 + \alpha)$.

If $0 < \alpha < 1/2$ we proceed similarly and find that for $n > 0$ and small $\langle n\alpha\rangle$, $\pi(\langle n\alpha\rangle) = (2n - LM_n)z - (2n - 2LM_n)$ where LM_n is the number of large and medium vertices among $\langle\alpha\rangle, \dots, \langle n\alpha\rangle$. We again have $LM_n/n \to 1-\alpha$, and the corresponding limiting argument again shows that $z = 2\alpha/(1+\alpha)$.

Note that by equation (2), $\pi_\mu(\alpha) = 2\alpha/(1+\alpha)$.

Theorem 4.2. *Let μ be uniform on the set $\{\alpha, 1\}$, with $0 < \alpha < 1$ and α irrational. Let π be a stationary distribution for the iterated random function. Then $\pi = \pi_\mu$.*

Proof. Being right continuous, π is determined by its values on the dense set $\mathcal{O}_0$. Therefore, by the above discussion, π is determined by its value at α. Finally, we have shown that there is only one possible value of $\pi(\alpha)$, namely the value $\pi_\mu(\alpha)$. □

5. Rate of convergence

We will use the description of G_x to give us a rate of convergence for the backwards iterates in Theorem 5.4. We shall need some preliminary results before proving the bound in Section 5.3.

5.1. **Continued fractions.** Fix α irrational, and let

$$\alpha = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots}}} \tag{7}$$

be the continued fraction expansion of α (each a_i is an integer, with $a_i > 0$ if $i > 0$). For $n \geq 0$ let p_n and q_n be relatively prime integers such that

$$\frac{p_n}{q_n} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots + \frac{1}{a_n}}}}.$$

It follows that q_n satisfies the recursion $q_{n+1} = a_n q_n + q_{n-1}$ (see e.g. [H-W]). The p_n and q_n are such that

$$|\alpha - \frac{p_n}{q_n}| < \frac{1}{2q_n^2}. \tag{8}$$

We shall make use of the following fact.

Lemma 5.1. *For any $x \in [0,1]$, there is an integer k with $0 \leq k < q_n$ such that $0 \leq \langle x - k\alpha\rangle < \frac{3}{2q_n}$.*

Proof. Let $S = \{\langle k\alpha\rangle \ : \ k = 0, 1, \dots, q_n - 1\}$, and for each $i = 0, 1, \dots, q_n - 1$ let $\bar{\imath}$ denote the unique integer among $0, 1, \dots, q_n - 1$ satisfying $\bar{\imath}p_n \equiv i (\mathrm{mod} q_n)$. Note that $0 \in S$. It essentially follows from (8) that each of the $q_n - 1$ (disjoint) intervals $(\frac{i}{q_n} - \frac{1}{2q_n}, \frac{i}{q_n} + \frac{1}{2q_n})$, where $i = 1, \dots, q_n - 1$, contains exactly one point of S, namely $\langle\bar{\imath}a\rangle$. To see this, note that by (8),

$$\left|\bar{\imath}\alpha - \bar{\imath}\frac{p_n}{q_n}\right| = \bar{\imath}\left|\alpha - \frac{p_n}{q_n}\right| < \bar{\imath} \cdot \frac{1}{2q_n^2} < \frac{1}{2q_n}.$$

This implies that there is no integer strictly between $\bar{\imath}\alpha$ and $\bar{\imath}p_n/q_n$, and therefore $|\bar{\imath}\alpha - \bar{\imath}p_n/q_n| = |\langle\bar{\imath}\alpha\rangle - \langle\bar{\imath}p_n/q_n\rangle|$. Thus

$$|\langle\bar{\imath}\alpha\rangle - i/q_n| = |\langle\bar{\imath}\alpha\rangle - \langle\bar{\imath}p_n/q_n\rangle| = |\bar{\imath}\alpha - \bar{\imath}p_n/q_n| < \frac{1}{2q_n}.$$

Note also that if $\alpha > p_n/q_n$ then $\langle\bar{\imath}\alpha\rangle > i/q_n$ for all $i = 0, 1, \ldots, q_n - 1$; whereas if $\alpha < p_n/q_n$ then $\langle\bar{\imath}\alpha\rangle < i/q_n$ for all such i. Thus no two consecutive points of S are more than $\frac{3}{2q_n}$ apart.

To summarize, the smallest point of S is 0, any two consecutive points of S are within $\frac{3}{2q_n}$ of each other, and the largest point of S is greater than $\frac{q_n-1}{q_n} - \frac{1}{2q_n} = 1 - \frac{3}{2q_n}$. This proves the lemma. □

5.2. **Random walks on the line.** Fix $0 < \alpha < 1$ irrational, and fix $0 < x_0 < \alpha$ in a generic orbit. The vertex v_0 corresponding to x_0 in the graph $G = G_{x_0}$ is small; thus (see Section 4.1) v_0 is a cut-vertex of G (that is, $G - v_0$ is not connected).

For an arbitrary vertex v of G, let $d(v_0, v)$ be the distance from v_0 to v in G, ignoring orientations. Define a map ρ from the vertex set of G to $\mathbb{Z}$ as follows: let $\rho(v) = n > 0$ if $d(v_0, v) = n$ and v is in the right hand component of $G - v_0$; let $\rho(v) = n < 0$ if $d(v_0, v) = -n$ and v is in the left hand component of $G - v_0$; and let $\rho(v_0) = 0$.

Observation: *If θ is chosen uniformly from $\{\alpha, 1\}$ then for any $x \in \mathcal{O}_{x_0}$, we have*

$$\rho(f_\theta(x)) = \begin{cases} \rho(x) + 1 & w.p.\ 1/2, \\ \rho(x) - 1 & w.p.\ 1/2. \end{cases}$$

The next lemma says that long paths on G must go through vertices which correspond to very small numbers. As x_0 is fixed, we no longer distinguish between a point $x \in \mathcal{O}_{x_0}$ and the corresponding vertex of G.

Lemma 5.2. *Let $x, y \in \mathcal{O}_{x_0}$, and suppose that $\rho(x) - \rho(y) \geq 2q_m$. Then there exists $z \in \mathcal{O}_{x_0}$ with $\rho(x) \geq \rho(z) \geq \rho(y)$ such that $0 \leq z < \frac{3}{2q_m}$.*

Proof. From the graph G (more precisely, from the proof of Lemma 3.1), it is clear that for any $k \in \mathbb{Z}$,

$$k \leq |\rho(\langle x - k\alpha\rangle) - \rho(x)| \leq 2k.$$

Suppose first that $x = \langle r\alpha + x_0\rangle$ for some $r \in \mathbb{Z}$; thus the vertex x is on the top row of the graph G and $\langle x+\alpha\rangle$ is to the right of x. Lemma 5.1 gives a number k with $0 \leq k < q_m$ and $0 \leq \langle x - k\alpha\rangle := z < \frac{3}{2q_m}$. Then $\rho(y) \leq \rho(x) - 2q_m \leq \rho(z) \leq \rho(x)$ as desired.

On the other hand if x is on the bottom row of G, i.e. $x = \langle r\alpha - x_0\rangle$ for some $r \in \mathbb{Z}$, then we apply Lemma 5.1 to the number $\langle x + (q_m - 1)\alpha\rangle$ to get a k with $0 \leq k < q_m$ and $0 \leq \langle x + k\alpha\rangle := z < \frac{3}{2q_m}$. Again we have $\rho(y) \leq \rho(x) - 2q_m \leq \rho(z) \leq \rho(x)$. □

We shall need the following; the simple proof provided was pointed out to us by D. Aldous.

Lemma 5.3. *There exists an integer N_0 and a constant $c_0 > 0$ such that for all $n > N_0$, a simple random walk of length n^3 on a line remains within n units of its starting point with probability less than or equal to $e^{-c_0 n}$.*

Proof. For a walk to stay within n units of its starting point it is necessary (though not sufficient) that no subwalk of the walk moves more than $2n$ units in either direction. By the central limit theorem, there is an integer N_0' and a constant $c_0' > 0$ so that for $k > N_0'$, a walk of length $c_0' k^2$ remains within k units of its starting point with probability less than e^{-1}. Given a walk of length n^3, we break it into pieces of length $c_0'(4n^2)$ and apply the previous fact with $k = 2n$ to each of the $n/4c_0'$ pieces. It follows that the probability that the walk remains within n units of its starting point is at most $e^{-n/4c_0'}$. □

5.3. **The bound.** The content of the following theorem is essentially that with high probability, the entire interval $[0,1]$ gets within ϵ of its limit under backward iteration after $O((1/\epsilon)^3 \log(1/\epsilon))$ steps.

Theorem 5.4. *Let $0 < \alpha < 1$ be irrational, and fix ϵ small and positive. Find k such that $8/q_k < \epsilon$ (where q_k is as in Section 5.1), and let $N = 8q_k^3 \log_2 q_k$. There is a constant $c > 0$, which does not depend on ϵ, such that with probability larger than $1 - e^{-c/\epsilon}$, we have*

$$|B_n(\Theta, x) - B(\Theta)| < \epsilon \quad \textit{for all } n \geq N \textit{ and for all } x \in [0,1], \tag{9}$$

where $\Theta = (\theta_1, \theta_2, \ldots)$ is chosen uniformly from $\{\alpha, 1\}^\infty$.

Proof. Let $j = 8q_k^3$. Given an interval I we compute a lower bound p for the probability that

$$||f_{\theta_1} \circ \cdots \circ f_{\theta_j}(I)|| \leq \frac{1}{2}||I|| + \frac{2}{q_k}. \tag{10}$$

It follows then that with probability bigger than p^t,

$$||f_{\theta_1} \circ \cdots \circ f_{\theta_{tj}}([0,1])|| \leq (\frac{1}{2})^t + \frac{4}{q_k}.$$

In this way, with $t = \lfloor \log_{\frac{1}{2}} \frac{4}{q_k} + 1 \rfloor = \lfloor \log_2 q_k - 1 \rfloor$ and $n \geq tj$ fixed, we have that for all $x \in [0,1]$, (9) holds, i.e.,

$$\begin{aligned} |B_n(\Theta, x) - B(\Theta)| &= \lim_{m\to\infty} |B_n(\Theta, x) - f_{\theta_1} \circ \cdots \circ f_{\theta_m}(x)| \\ &= \lim_{m\to\infty} |B_n(\Theta, x) - B_n(\Theta, f_{\theta_{n+1}} \circ \cdots \circ f_{\theta_m}(x)| \\ &< \frac{8}{q_k} < \epsilon, \end{aligned}$$

with probability at least p^t.

We get the estimate p as follows. Pick a point $m^* \in \mathcal{O}_{x_0}$ (x_0 as in section 5.2) such that m^* is within $\delta = \frac{1}{2q_k}$ of the midpoint of I (recall that $\mathcal{O}_{x_0}$ is dense). By the observation in Section 5.2, the sequence

$$\rho(m^*), \rho(f_{\theta_j}(m^*)), \rho(f_{\theta_{j-1}} \circ f_{\theta_j}(m^*)), \ldots, \rho(f_{\theta_1} \circ \cdots \circ f_{\theta_j}(m^*))$$

is a simple random walk on $\mathbb{Z}$.

If this walk does not stay within $2q_k$ units of $\rho(m^*)$ on the line, then we can apply Lemma 5.2 to get a j_0, $1 \leq j_0 \leq j$ such that $f_{\theta_{j_0}} \circ \cdots \circ f_{\theta_j}(m^*) < \frac{3}{2q_k}$. Then Lemma 3.3 implies that $||f_{\theta_{j_0}} \circ \cdots \circ f_{\theta_j}(I)|| \leq \frac{1}{2}||I|| + \delta + \frac{3}{2q_k} \leq \frac{1}{2}||I|| + \frac{2}{q_k}$. Since f_θ never increases the length of an interval, (10) holds.

With c_0 and N_0 as in Lemma 5.3, we therefore have $\mathsf{p} \geq 1-e^{-2c_0 q_k}$ (provided that $2q_k > N_0$). Thus for $n \geq 8q_k^3 \log_2 q_k \geq tj$ we have that (9) holds with probability at least $(1 - e^{-2c_0 q_k})^{\log_2 q_k}$. By the binomial theorem, this last expression is the same as

$$1 - \frac{\log_2 q_k}{e^{2c_0 q_k}} + O(\frac{\log_2^2 q_k}{e^{4c_0 q_k}})$$

which, for large q_k, is at least $1 - c' e^{-2c_0 q_k} \log_2 q_k > 1 - e^{-cq_k}$. This is the desired result. □

As an application of this theorem, let $\alpha = 1/\sqrt{2}$. Here (see Section 5.1) we have $a_0 = 0, a_1 = 1$, and $a_i = 2$ for $i \geq 2$. The recurrence relation for q_n implies that

$$q_n = \frac{1}{2}((1+\sqrt{2})^n + (1-\sqrt{2})^n).$$

The theorem implies in this case that there are constants c and C such that for $n \geq C\epsilon^{-3}\log(1/\epsilon)$, and for all $x \in [0,1]$, the nth backward iterate is within ϵ of its limit with probability at least $1 - e^{-c/\epsilon}$. The same statement holds for any α for which the set of a_i's (see (7)) is bounded; in particular this holds for any quadratic irrational α.

6. Acknowledgements

We thank Persi Diaconis for suggesting this problem and for numerous helpful conversations. We also thank Julie Landau, without whose support and net game this paper would not have been written.

References

[D-F] P. Diaconis and D. Freedman, *Iterated Random Functions,* SIAM Rev. **41** (1999) pp. 45–76.

[F] W. Feller, *An Introduction to Probability Theory and Its Applications, II* 2nd ed., Wiley, New York, 1971.

[H-W] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers* 5th ed., Clarendon Press, Oxford, 1979.

[L] G. Letac, *A Contraction Principle for Certain Markov Chains and Its Applications,* Contemp. Math. **50** (1986), pp. 263–273.

[M] J. Mattingly, *Contractivity and Ergodicity of the Random Map* $x \mapsto |x-\theta|$, preprint (2001).

Department of Mathematics, University of Georgia, Athens GA 30602
E-mail address: abrams@math.uga.edu
URL: http://www.math.uga.edu/∼abrams/

E-mail address: hjl@research.att.com

Mathematical Sciences Research Institute, 1000 Centennial Drive, Berkeley CA 94720
E-mail address: landau@math.berkeley.edu

Department of Mathematics, Pomona College, 610 N. College Ave, Claremont CA 91711
E-mail address: jpommersheim@pomona.edu

Department of Mathematics, Northwestern University, 2033 Sheridan Road, Evanston IL 60208
E-mail address: zaslow@math.nwu.edu
URL: http://www.math.northwestern.edu/∼zaslow/