*The Annals of Probability*
1999, Vol. 27, No. 4, 1952–1979

# LOCALLY CONTRACTIVE ITERATED FUNCTION SYSTEMS

BY DAVID STEINSALTZ

*University of California, Berkeley*

An iterated function system on $\mathscr{X} \subset \mathbb{R}^d$ is defined by successively applying an i.i.d. sequence of random Lipschitz functions from $\mathscr{X}$ to $\mathscr{X}$. This paper shows how $F_n = f_1 \circ \cdots \circ f_n$ may converge even in the absence of the strong contraction conditions, for instance, Lipschitz constant smaller than 1 on average, which earlier work has required. Instead, it is posited that there be a region of contraction which compensates for the noncontractive or even expansive part of the functions. Applications to queues, to self-modifying random walks and to random logistic maps are given.

**1. Introduction.** Let $(\mathscr{X}, \rho)$ be a metric space, and $\mathscr{H}$ the set of maps from $\mathscr{X}$ to itself. An *iterated function system* is defined by a probability measure $\nu$ on $\mathscr{H}$ with respect to a $\sigma$-algebra $\mathscr{F}$ which makes the map $(x, f) \mapsto f(x)$ measurable. (The space $\mathscr{X}$ is outfitted with the Borel $\sigma$-algebra.) This allows us to speak of an i.i.d. sequence $f_1, f_2, \ldots$ of $\mathscr{H}$-valued random variables with distribution $\nu$, from which arise two distinct compositions,

$$F_n(x) = f_1 \circ f_2 \circ \cdots \circ f_n(x) \quad \text{and} \quad \widetilde{F}_n(x) = f_n \circ f_{n-1} \circ \cdots \circ f_1(x).$$

An important observation is that when $X_0$ is an $\mathscr{X}$-valued random variable independent of the sequence $(f_i)$, the sequence of iterates $X_n = \widetilde{F}_n(X_0)$ is a Markov chain. The system $F_n(X_0)$, though, is a very different sort of object. The distinction is not seen for any fixed $n$, of course: $F_n(x)$ and $\widetilde{F}_n(x)$ have the same distribution for any $x$ and any $n$. But the process $F_n(x)$ is not a Markov chain and in many circumstances tends to converge pointwise. This coupling between two dissimilar processes which nonetheless have the same marginal distributions has proved fruitful in several contexts. In particular, since most Markov chains may be represented as iterated function systems (cf. Chapter 1 of [10]), the reverse sequence gives an alternative window into the Markov chain. Applied to finite state spaces $\mathscr{X}$, this is the crux of the exciting new method known as "coupling from the past," introduced by Propp and Wilson, for simulating the stationary distribution of a Markov chain (see [14] and references therein).

The fundamental idea there is to follow the sequence $F_n(x)$ until it reaches its limit point; when $\mathscr{X}$ is finite, this means that the sequence becomes fixed. The inference then is that the distribution of this limit point is exactly the stationary distribution for $\widetilde{F}_n$. Intuitively, this is clear: if $X_\infty = \lim_{n\to\infty} F_n(x)$,

Received December 1998; revised June 1999.
*AMS* 1991 *subject classifications*. Primary 58F08; secondary 26A18, 60J05.
*Key words and phrases*. Iteration, random attractors, logistic maps, Zeno's walk, single-server queues, Lyapunov functions.

and $f$ is chosen independent of $F_n$ from the distribution $\nu$, then

$$f(X_\infty) =_d f(f_1(f_2(\cdots(x)\cdots))) =_d f_1(f_2(f_3(\cdots(x)\cdots))),$$

simply by renumbering. The following theorem is due to Letac [11].

*If the distribution $\nu$ is concentrated on continuous functions, and if $F_\infty(x) = \lim_{n\to\infty} F_n(x)$ almost surely exists and is independent of $x$, then the distribution of $X_\infty = F_\infty(x)$ is the unique stationary distribution for the Markov chain $\widetilde{F}_n(x)$, and it is attractive, in the sense that any compactly supported initial distribution converges to it under the action of the Markov operator.*

It is also straightforward to see that the pointwise rate of convergence of the sequence $F_n(x)$ is an upper bound for the rate of convergence in distribution of the Markov chain $\widetilde{F}_n(x)$. Here I use the Wasserstein metric, which is defined for $\mu_1, \mu_2$ probability measures on $X$ as (cf. [19])

$$W(\mu_1, \mu_2) = \sup_{\mathrm{Lip}(f)\le 1} |\mu_1(f) - \mu_2(f)| = \inf_{\substack{Y\sim\mu_1\\ Z\sim\mu_2}} \{E(\rho(Y, Z))\},$$

where $\mathrm{Lip}(f)$ is the Lipschitz constant

$$\mathrm{Lip}(f) = \sup_{x\neq y\in\mathscr{X}} \frac{\rho\big(f(x), f(y)\big)}{\rho(x, y)},$$

and the infimum is taken over random variables $Y$ and $Z$ defined on any common probability space such that $Y$ has marginal law $\mu_1$ and $Z$ has marginal law $\mu_2$.

This fact motivates the following definition.

DEFINITION 1. The random variable $F_\infty(x)$ is defined to be $\lim_{n\to\infty} F_n(x)$ if this exists. On the event where the limit is independent of $x$, it will be denoted $X_\infty$, and it will be said that $X_\infty$ *exists*. An iterated function system is *attractive* if $X_\infty$ almost surely exists and is finite.

(The substance of this definition is due to G. Letac, though he called the property *contractive*.) The problem then offers itself, how to determine whether a given system is attractive. A simple but significant lemma which is often applied to this end (cf. [4]) is the following.

LEMMA 1. *A sufficient condition for $F_\infty(x)$ to exist and be finite almost surely is that*

$$\sum_{n=1}^{\infty} \rho\big(F_n(x), F_{n+1}(x)\big)$$

*be almost surely finite.*

Most immediate is the case, first addressed by Hutchinson [8], in which the distribution $\nu$ is concentrated on finitely many contractions, in the sense that there is a positive constant $r < 1$ such that for all $f$ in the support of $\nu$,

$$\rho(f(x), f(y)) \leq r\rho(x, y). \tag{1}$$

This may be generalized without difficulty to the following definition.

DEFINITION 2. An iterated function system is *strongly contractive* if $\mathrm{E} \log \mathrm{Lip} f < 0$.

The following result is then fairly immediate.

PROPOSITION 1. *If the system defined by $\nu$ is strongly contractive and if there exists $\delta > 0$ such that*

$$\mathrm{E}\big[\log^{1+\delta} \rho(f(x), x)\big] < \infty,$$

*then the system is attractive.*

The proof differs only in minor technical points from the one in Hutchinson's paper. Fundamentally, it follows from the observation that strong contractivity makes $\mathrm{Lip}(F_n) < cr^n$ almost surely for $n$ sufficiently large, so $\sum_{n=0}^{\infty} \mathrm{Lip}(F_n)$ is almost surely finite. Because I do not assume here that the random functions are all contractions, I lose Hutchinson's conclusion that the limit point is concentrated on a compact set (which is the unique invariant set of the contractions), but it is still true that the random function induces a strict contraction on compactly supported measures (in the Wasserstein metric).

Barnsley and Elton [4] have extended these results to cases in which the functions $T_i$ are not contractions, but in which they are contractive on the average between any two points.

DEFINITION 3. An iterated function system is *average contractive* if for some positive $q$,

$$\sup_{x \neq y} \mathrm{E}\left[\left(\frac{\rho(f(x), f(y))}{\rho(x, y)}\right)^q\right] < 1. \tag{2}$$

We have then the following result due to [4].

PROPOSITION 2. *If the system defined by $\nu$ is average contractive, and if for some positive $q$,*

$$\mathrm{E}\big[\rho(f(x), x)^q\big] < \infty,$$

*then the system is attractive.*

(The definition of average contractivity given by Barnsley and Elton is apparently a weaker condition, involving as it does only the expectation of the logarithm, rather than any $q$th power. It should be observed, though, that they assume $\nu$ to be supported on finitely many functions, and then show that this, together with the logarithmic bound, implies the condition defined here.) As in Hutchinson's setting, the convergence in distribution results from a strict contraction on compactly supported probability measures. Barnsley and Elton point out, in addition, that the condition of average contractivity may be replaced by this slightly weaker condition.

DEFINITION 4. An iterated function system is *eventually average contractive* if for some positive $q$,

$$\lim_{n\to\infty} \sup_{x\neq y} \mathrm{E}\left[\left(\frac{\rho(F_n(x), F_n(y))}{\rho(x, y)}\right)^q\right] < 1. \tag{3}$$

Arnold and Crauel [2] have further extended and varied these results, identifying the existence of (possibly nonattractive) invariant measures with the sign of the Lyapunov exponents in the case that $\nu$ is supported on finitely many affine maps.

In an earlier paper [16], I analyzed one specific example of an iterated function system which does not seem to fit into any of the above categories. In considering the self-modifying random walk which I called "Zeno's walk," I was led to redefine the walk as $X_n = F_n(0)$, where $F_n = f_1 \circ f_2 \circ \ldots \circ f_n$ and $(f_i)$ are chosen independently, with probability $p$ and $1-p$, respectively, from the two piecewise linear functions $f^+$ and $f^-$, shown in Figure 1:

$$\begin{aligned} f^+(x) &:= \begin{cases} x+1, & \text{if } x \geq 0, \\ \dfrac{x}{2}+1, & \text{if } 0 \geq x \geq -2, \\ x+2, & \text{if } -2 \geq x\ ; \end{cases} \\ f^-(x) &:= \begin{cases} x-1, & \text{if } x \leq 0, \\ \dfrac{x}{2}-1, & \text{if } 0 \leq x \leq 2, \\ x-2, & \text{if } 2 \leq x \end{cases} . \end{aligned} \tag{4}$$

Let $P_n = \mathfrak{L}(X_n)$, the law of $X_n = F_n(0)$. Observe that $\mathrm{Lip}(f_i)$, hence $\mathrm{Lip}(F_n)$ as well, is always 1, so $\nu_p$ cannot be strongly contractive; nor is the condition for average contractivity satisfied, except in the range $[-2, 2]$. For any $x, y \geq 2$, $f^+(x) - f^+(y) = f^-(x) - f^-(y) = x - y$.

The earlier paper included the following results:

$$\text{The system defined by } \nu_p \text{ is attractive iff } p \in \left(\frac{1}{3}, \frac{2}{3}\right); \tag{5}$$

$$\begin{gathered} W(P_n, P_\infty) \leq \mathrm{E}|X_n - X_\infty| \leq c_p r_p^n, \text{ where } c_p \text{ is a constant and} \\ r_p = \left(\frac{27p(1-p)(p\vee(1-p))}{4}\right)^{1/3}; \end{gathered} \tag{6}$$

$$\mathrm{supp}(P_\infty) = \mathbb{R}. \tag{7}$$

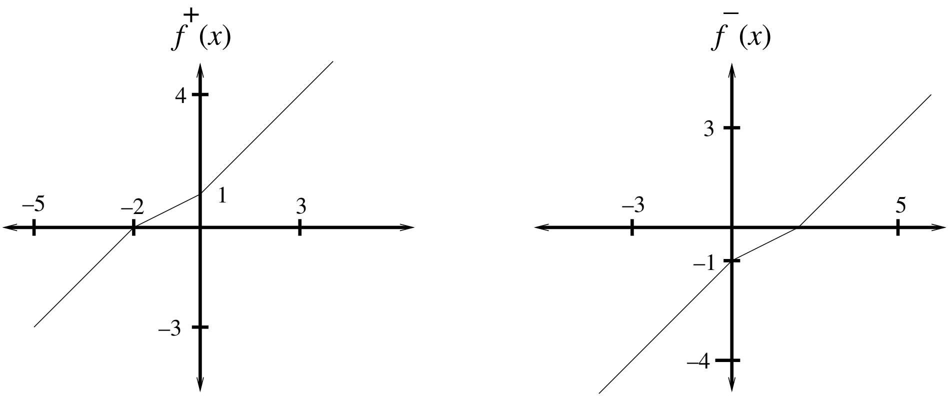


FIG. 1.

(To be sure, the language and the focus of that paper were rather different.) I also computed $P_\infty$ explicitly, and gave its Hausdorff dimension, which is always smaller than 0.96. As in the case of average and strong contractivity, the measures still converge at a geometric rate. On the other hand, the action of the random function is not a contraction on compactly supported probabilities. For instance, $\delta_2$ (the unit point mass at 2) is mapped to $p\delta_3 + (1-p)\delta_0$, while $\delta_3$ is mapped to $p\delta_4 + (1-p)\delta_1$. The Wasserstein distance is 1 after the transformation, as before. And yet, while average contractive everywhere, the system does have a region which contracts. Thus, while the iterated function $F_n$ will always have infinite tails with slope 1, viewed within any finite window the function really will be seen to flatten out.

A problem of a similar flavor arises in an article by Letac and Chamayou [5]. Among their multitude of stationary distributions for particular iterated function systems appear several whose attractivity seems doubtful; they are certainly not strongly, or even average contractive. They may, however, be locally contractive, and in Section 7.4 I address one of these examples, the randomized logistic mappings

$$f_y(x) = 4yx(1-x)$$

for $x \in \mathscr{X} = [0,1]$ and $y$ chosen from a $\beta_{a+1/2,\,a-1/2}$ distribution on $[0,1]$. There I show that the system is indeed attractive, at least for $a \geq 2$.

This motivates the following definition:

DEFINITION 5. An iterated function system is *locally contractive* if there exists a *drift function* $\phi: \mathscr{X} \to [1,\infty)$ and $r \in (0,1)$ such that

$$G_n(x) := \mathrm{E}[D_x F_n] \leq \phi(x) r^n,$$

where $D_x f = \limsup_{y\to x} \rho(f(x), f(y))/\rho(x,y)$ is the local Lipschitz constant at $x$.

This definition is valuable because, on the one hand it can easily be shown to imply attractivity (Theorem 1), while on the other hand there are sufficient conditions for local contractivity which are themselves often readily verified (Theorem 2).

The three conditions for attractivity, Propositions 1 and 2 and Theorem 1, in a sense trade globality against the strength of the moment conditions. To make this clear, I will crudely interpret the conditions of strong and average contractivity as

$$\mathrm{E}\left[\log \sup_{x\in\mathscr{X}} D_x f\right] < 0$$

and

$$\sup_{x\in\mathscr{X}} \mathrm{E}[(D_x f)^q] < 1 \quad \text{for some } q > 0,$$

respectively. Average contractivity is a weaker condition because it moves the supremum outside the expectation, but this seems to demand the compensation of a stronger moment condition: $q$th power instead of the logarithm. The distance between $x$ and $f(x)$ also requires correspondingly stronger moments.

Moving on to local contractivity, the globality is weaker still. The slightly stronger condition of Theorem 2 requires only that for some function $\phi$,

$$\sup_{x\in\mathscr{X}} \frac{\mathrm{E}[\phi(f(x))D_x f]}{\phi(x)} < 1. \tag{8}$$

This allows considerably more local variation in the average behavior of $D_x f$, but this time at the cost of placing the condition on $\mathfrak{L}^1$ norms, instead of $\mathfrak{L}^q$ for arbitrary positive $q$.

The drift function might be compared to a Lyapunov function, which serves to prove the stability of deterministic and stochastic dynamical systems (cf. [9]). In the context of Markov processes, a Lyapunov function is a function $\phi$ such that $\mathscr{L}\phi(x) < -c\phi(x)$, where $\mathfrak{L}$ is the generator of the Markov process and $-c$ is a negative constant. For chains, the corresponding property is $\mathrm{E}[\phi(X_{n+1})|X_n] \le r\phi(X_n)$ where $r < 1$. This would be the same as 8 if the factor $D_x f$ were removed. Similar drift conditions are used to prove the geometric ergodicity of Markov chains (cf. [13]).

**2. Notation.** In what follows, $\mathscr{X}$ will be a convex subset of $\mathbb{R}^n$ and $\nu$ a probability concentrated on the set of Lipschitz functions from $\mathscr{X}$ to itself.

Let $f_1, f_2, \dots$ be i.i.d. samples from the distribution $\nu$, and define

$$F_n(x) := f_1\Big(f_2(\cdots(f_n(x))\cdots)\Big);$$

$$\widetilde{F}_n(x) := f_n\Big(f_{n-1}(\cdots(f_1(x))\cdots)\Big).$$

Also define

$$G_n(x) := \mathrm{E}[D_x F_n]. \tag{9}$$

When $f_i$ is invertible, let $\tau_i(x) := f_i^{-1}(x) - x$.

For $\phi\colon \mathscr{X} \to \mathbb{R}^+$ any measurable function I define, for $x, y \in \mathscr{X}$,

$$\Phi(x; y) := \sup_{0 \le t \le 1} \{\phi(x + t(y - x))\}. \tag{10}$$

I also define the *growth rate of $\phi$ at $x$ with respect to $\nu$* to be

$$r_x := \mathrm{E}\left[\frac{\phi(f(x))}{\phi(x)} D_x f\right], \tag{11}$$

and the *growth rate of $\phi$ with respect to $\nu$*,

$$r := \sup_{x \in \mathscr{X}} r_x. \tag{12}$$

Finally,

$$C_x(\phi, \nu) := \mathrm{E}\big[\|f(x) - x\| \Phi(x; f(x))\big]. \tag{13}$$

Where there is no risk of confusion, $C_x(\phi, \nu)$ will simply be written as $C_x$.

For real numbers $x$ and $y$, $(x; y)$ or $[x; y]$ will represent the open or closed interval with endpoints $x$ and $y$, without regard for their order.

**3. Local contractivity.** I can now state the elementary results on local contractivity. First, local contractivity plus a bound on the tails of $f(x)$ (expressed in terms of $C_x$) suffice to establish attractivity.

THEOREM 1. *If $\nu$ is locally contractive with drift function $\phi$, and if $C_x(\phi, \nu)$ is finite for all $x \in \mathscr{X}$, then $\nu$ is attractive. Furthermore, there is a bound on the rate of convergence*

$$\mathrm{E}\|F_n(x) - F_\infty(y)\| \le \left(\frac{C_x}{1 - r} + \|x - y\| \Phi(x; y)\right) r^n. \tag{14}$$

Second, a simple (sufficient) test for local contractivity is the following.

THEOREM 2. *If $\phi\colon \mathscr{X} \to [1, \infty)$ is a continuous function whose growth rate $r$ with respect to $\nu$ is smaller than 1, then the iterated function system defined by $\nu$ is locally contractive with drift function $\phi$.*

PROOF. Since $F_0(x) = x$, the relation

$$G_n(x) \le \phi(x) r^n \tag{15}$$

is satisfied for $n = 0$ and any $x \in \mathscr{X}$. Now, by the chain rule

$$\begin{aligned} G_{n+1}(x) &= \mathrm{E}\big[D_x(F_n \circ f_{n+1})\big] \\ &\le \mathrm{E}\big[D_{f_{n+1}(x)} F_n D_x f_{n+1}\big] \\ &\le \mathrm{E}\big[G_n(f_{n+1}(x)) D_x f_{n+1}\big], \end{aligned}$$

where the last line makes use of independence. Given (15),

$$\begin{aligned} G_{n+1}(x) &\le r^n \mathrm{E}\big[\phi(f(x))D_x f\big] \\ &= r^n \phi(x)\mathrm{E}\left[\frac{\phi(f(x))}{\phi(x)} D_x f\right] \\ &\le r^{n+1}\phi(x). \end{aligned}$$

By induction, the bound (15) holds for all natural numbers $n$ and all $x \in \mathscr{X}$.

PROOF OF THEOREM 1. Local contractivity implies that for every natural number $i$,

$$\begin{aligned} \mathrm{E}\|F_i(x) - F_{i+1}(x)\| &= \mathrm{E}\|F_i(x) - F_i\big(f_{i+1}(x)\big)\| \\ &\le \mathrm{E}\left[\|x - f_{i+1}(x)\| \int_0^1 D_{x+t(f_{i+1}(x)-x)} F_i \, dt\right] \\ &\le \sup_{0\le t\le 1} \mathrm{E}\Big[\|x - f_{i+1}(x)\| D_{x+t(f_{i+1}(x)-x)} F_i\Big] \\ &= \sup_{0\le t\le 1} \mathrm{E}\Big[\|x - f_{i+1}(x)\| G_i\big(x + t(f_{i+1}(x) - x)\big)\Big] \\ &\le \mathrm{E}\Big[r^i \Phi(x; f_{i+1}(x))\|x - f_{i+1}(x)\|\Big] \\ &= C_x(\phi) r^i. \end{aligned}$$

Then

$$\mathrm{E}\sum_{i=n}^{\infty}\Big\|F_i(x) - F_{i+1}(x)\Big\| \le \frac{C_x(\phi)}{1-r} r^n.$$

By assumption, this is finite for all $x$, which proves the almost-sure convergence, by Lemma 1. Thus $F_\infty(x)$ exists almost surely and

$$\mathrm{E}\Big\|F_n(x) - F_\infty(x)\Big\| \le \frac{C_x(\phi)}{1-r} r^n.$$

Similarly,

$$\mathrm{E}\big\|F_n(x) - F_n(y)\big\| \le r^n \|x - y\| \Phi(x; y).$$

from which attractivity and the conclusion (14) follow.

A case of special interest is the "parabolic" setting, where the Lipschitz constant of $f$ is equal to 1. (If it is strictly smaller, then strong contraction applies.)

PROPOSITION 3. *Suppose that* $\operatorname{Lip} f \le 1$ *almost surely, and that* $\|f(x) - x\|$ *has subexponential tails uniformly in* $x$*, in the sense that there are positive constants* $k_1$ *and* $k_2$ *such that for all* $x \in \mathscr{X}$ *and all positive* $t$*,*

$$\mathrm{P}\big\{\|f(x) - x\| > t\big\} \le k_1 \exp(-k_2 t). \tag{16}$$

*Suppose too that there exists a critical radius $R$ such that*

$$\sup_{\|x\|<R} \mathrm{E}\big[D_x f\big] < 1 \tag{17}$$

*and*

$$\sup_{\|x\|=R} \mathrm{E}\|f(x)\| < R. \tag{18}$$

*Then the system is locally contractive, and*

$$\lim_{n\to\infty}\Big(\mathrm{E}\|F_n(x)-X_\infty\|\Big)^{1/n} \le \inf_{\lambda>0}\sup_{\|x\|=R}\mathrm{E}\Big[\exp(\lambda(\|x\|-\|f(x)\|))\Big].$$

PROOF. I will apply Theorem 2 with $\phi(x)=\exp\{\lambda\|x\|\}$, for $\lambda$ slightly larger than 0. Then $\Phi(x;y)=\exp\{\lambda(\|x\|\vee\|y\|)\}$.

Let $\rho = R-\sup_{\|x\|=R}\mathrm{E}\|f(x)\|$. Observe that Lip $f\le 1$ implies that for all $x$ such that $\|x\|\ge R$,

$$\|f(x)\|-\|x\| \le \left\|f\Big(\frac{R}{\|x\|}x\Big)\right\| - R.$$

By assumption (18),

$$\sup_{\|x\|\ge R}\mathrm{E}\big[\|f(x)\|-\|x\|\big] = \sup_{\|x\|=R}\mathrm{E}\big[\|f(x)\|-R\big] \le -\rho < 0.$$

Define

$$r_x(\lambda) = \mathrm{E}\Big[\exp\big(\lambda(\|f(x)\|-\|x\|)\big)D_x f\Big].$$

The object is to show that there exists some positive $\lambda$ such that $\sup_{x\in\mathscr{X}} r_x(\lambda) < 1$. Since $D_x f \le 1$, it is certainly true that $r_x(0)\le 1$ for all $x$ in $\mathscr{X}$. By condition (17), $r_x(\lambda)$ is uniformly less than 1 for $\|x\|<R$ and $\lambda$ sufficiently small. I need then only to show that the stated conditions make the derivative of $r_x(\lambda)$ at $\lambda=0$ negative for $\|x\|\ge R$ and provide as well the necessary uniformity in $x$.

The assumption of subexponential tails implies that $r_x(\lambda)$ is finite and smooth for $|\lambda|<k_2$. Let $\Lambda$ be the interval $[-k_2/2, k_2/2]$. For $\lambda\in\Lambda$, the $k$th derivative of $r_x(\lambda)$ (with respect to $\lambda$) is

$$r_x^{(k)}(\lambda) = \mathrm{E}\Big[\big(\|f(x)\|-\|x\|\big)^k \exp\big(\lambda(\|f(x)\|-\|x\|)\big)D_x f\Big].$$

Because of the subexponential tails,

$$c_1 := \sup_{\substack{x\in\mathscr{X}\\ \lambda\in\Lambda}} |r_x''(\lambda)|$$

and

$$c_2 := \sup_{x\in\mathscr{X}}\mathrm{E}\Big[\big(\|f(x)\|-\|x\|\big)^2\Big]^{1/2}$$

are both finite. For all $x$,

$$\begin{aligned}\left|r'_x(0) - \mathrm{E}\big[\|f(x)\| - \|x\|\big]\right| &\le \mathrm{E}\Big[|1 - D_x f|\ \big|\|f(x)\| - \|x\|\big|\Big]\\ &\le c_2\big(1 - \mathrm{E}[D_x f]\big)^{1/2}\end{aligned}$$

by an application of the Cauchy–Schwarz inequality. For $\lambda \in \Lambda$,

$$r_x(\lambda) \le \mathrm{E}\big[D_x f\big]^{1/2}\,\mathrm{E}\big[\exp\{2\lambda(\|f(x)\| - \|x\|)\}\big]^{1/2}. \tag{19}$$

Alternatively, if $\|x\| \ge R$, by Taylor's formula,

$$\begin{aligned}r_x(\lambda) &\le 1 + r'_x(0)\lambda + c_1\lambda^2\\ &\le 1 + \big(-\rho + c_2\big(1 - \mathrm{E}[D_x f]\big)^{1/2}\big)\lambda + c_1\lambda^2.\end{aligned}$$

Then for $\lambda > 0$ sufficiently small,

$$\sup\big\{r_x(\lambda)\colon \|x\| \ge R \quad \text{and} \quad \mathrm{E}\big[D_x f\big] \ge 1 - \rho^2/2c_2^2\big\} < 1.$$

On the other hand, where $\mathrm{E}\big[D_x f\big]$ is smaller than $1 - \rho^2/2c_2^2$, the right-hand side of (19) is bounded away from 1. Putting these together with the earlier bound for $\|x\| < R$ tells us that $\sup_{x\in\mathscr{X}} r_x(\lambda) < 1$ for positive $\lambda$ sufficiently small.

**4. Inverse functions.** In this and the following section it will be assumed that $\mathscr{X} = \mathbb{R}$ and that $\nu$ is concentrated on nondecreasing functions. This will be denoted the *monotone context*. For real-valued functions $f$, $f'$ will be taken for definiteness to be the right-hand derivative, and the inverse will be defined as $f^{-1}(x) = \sup\{y\colon f(y) \le x\}$, where the supremum of the empty set is assigned the value $-\infty$. [Extended real-valued functions will be composed by the convention that $f(\infty) = \lim_{x\to\infty} f(x)$ and $f(-\infty) = \lim_{x\to-\infty} f(x)$. In the monotone context these are well defined.]

In this setting it is possible to formulate an alternative condition for attractivity which is slightly weaker than local contractivity and which has the additional advantage of being necessary and sufficient.

LEMMA 2. *Let $F_n\colon \mathbb{R} \to \mathbb{R}$ be a sequence of monotone nondecreasing functions. Then the statement that, for a given extended real number $x_\infty$, $\lim_{n\to\infty} F_n(x) = x_\infty$ for all $x \in \mathbb{R}$ is equivalent to*

$$\lim_{n\to\infty} F_n^{-1}(y) = \begin{cases} +\infty, & \text{if } y > x_\infty,\\ -\infty, & \text{if } y < x_\infty.\end{cases} \tag{20}$$

*(The limit simply does not exist in general for $y = x_\infty$.)*

PROOF. I will give the proof for $x_\infty$ real; for $x_\infty = \pm\infty$, only trivial modifications are required.

Suppose $\lim_{n\to\infty} F_n(x) = x_\infty$ for all $x \in \mathbb{R}$. Then for any $y > x_\infty$ and any real number $x_0$, there exists $N$ such that $F_n(x_0) < y$ for all $n > N$. That means that

$$F_n^{-1}(y) = \sup\{x: F_n(x) \le y\} \ge x_0$$

for all $n \ge N$, so $\liminf_{n\to\infty} F_n^{-1}(y) \ge x_0$. Since $x_0$ was arbitrary, it follows that $\lim_{n\to\infty} F_n^{-1}(y) = +\infty$. Similarly, $\lim_{n\to\infty} F_n^{-1}(y) = -\infty$ for $y < x_\infty$.

Conversely, suppose that the limit condition (20) is satisfied for some $x_\infty$. Then for any real number $x_0$ and any $y > x_\infty$, for $n$ sufficiently large,

$$\sup\{x: F_n(x) \le y\} = F_n^{-1}(y) > x_0.$$

By monotonicity, it follows that $F_n(x_0) \le y$ for $n$ sufficiently large. Since $y > x_\infty$ was arbitrary, $\limsup_{n\to\infty} F_n(x_0) \le x_\infty$. Similarly, $\liminf_{n\to\infty} F_n(x_0) \ge x_\infty$. Thus the limit exists and is equal to $x_\infty$.

The advantage of this approach is that $F_n^{-1} = f_n^{-1} \circ \cdots \circ f_1^{-1}$ is a Markov chain. This allows proofs of convergence using conventional transience criteria, without demanding any special moment conditions. A bonus is that the convergence criteria become necessary as well as sufficient in some special cases.

DEFINITION 6. The event that $\lim_{n\to\infty} F_n^{-1}(x)$ exists and is $+\infty$ or $-\infty$ will be described by saying that *x has an infinite inverse limit*. The *infinite inverse set* is the set of $x$ which have an infinite inverse limit almost surely.

PROPOSITION 4. *In the monotone context, $X_\infty$ exists almost surely if and only if $\mathscr{X}'$, the complement of the infinite inverse set, is countable. In this case, the stationary distribution is given by*

$$\mathrm{P}\{X_\infty < t\} = \mathrm{P}\left\{\lim_{n\to\infty} F_n^{-1}(t) = +\infty\right\},$$

*and* $\mathscr{X}' = \{x: \mathrm{P}\{X_\infty = x\} > 0\}$.

REMARK. Attractivity includes the additional condition that the limit be finite almost surely. Here this is equivalent to the condition that

$$\lim_{t\to\infty} \mathrm{P}\left\{\lim_{n\to\infty} F_n^{-1}(t) = +\infty\right\} = 1 = \lim_{t\to-\infty} \mathrm{P}\left\{\lim_{n\to\infty} F_n^{-1}(t) = -\infty\right\}.$$

PROOF. Choose any $x$ in $\mathscr{X}$. By Lemma 2, the event that $X_\infty$ exists and is less than $x$ is contained in the event that $\lim_{n\to\infty} F_n^{-1}(x) = +\infty$, and the event that $X_\infty$ exists and is greater than $x$ is contained in the event that $\lim_{n\to\infty} F_n^{-1}(x) = -\infty$. This means that

$$\{x \text{ does not have an i.i.l.}\} \subset \{X_\infty = x\} \cup \{X_\infty \text{ does not exist}\}.$$

Suppose that $X_\infty$ exists almost surely. Then $\mathscr{X}'$ is contained in the set of $x$ such that $\mathrm{P}\{X_\infty = x\} > 0$, which must be countable.

Suppose now that $\mathscr{X}'$ is countable, and let $\mathscr{Y}$ be a countable subset of $\mathbb{R}\backslash\mathscr{X}'$ which is dense in $\mathbb{R}$. Except on a null event, $y$ has an infinite inverse limit for all $y \in \mathscr{Y}$; from here on we may pretend that this null event does not exist. By monotonicity of $F_n^{-1}$, the two random variables

$$\inf\{y \in \mathscr{Y} \text{ s.t. } \lim_{n\to\infty} F_n^{-1}(y) = +\infty\} \quad \text{and} \quad \sup\{y \in \mathscr{Y} \text{ s.t. } \lim_{n\to\infty} F_n^{-1}(y) = -\infty\}$$

are equal. Call this $Y_\infty$. If $\varepsilon$ is any positive number, there exist $y_1$ and $y_2$ in the infinite inverse set such that $y_1 < Y_\infty < y_2$ and $|y_1 - y_2| < \varepsilon$.

Choose any $x \in \mathbb{R}$. For $n$ sufficiently large, $F_n^{-1}(y_2) > x$, which is equivalent to saying that $F_n(x) \le y_2$. Using the same argument for $y_1$, and letting $\varepsilon$ go to 0, it follows that $\lim_{n\to\infty} F_n(x) = Y_\infty$. Thus the limit of $F_n(x)$ exists almost surely, and is equal to $Y_\infty$, independent of $x$.

COROLLARY 1. *Suppose that* Lip $f \le 1$ *almost surely, and $\nu$ has no almost-sure fixed points, that is, points $x$ such that $\nu\{f(x) = x\} = 1$. Then the monotone system defined by $\nu$ is attractive if and only if*

$$\lim_{x\to\infty} \mathrm{E}[f^{-1}(x) - x] > 0 \quad \textit{and} \quad \lim_{x\to-\infty} \mathrm{E}[f^{-1}(x) - x] < 0. \tag{21}$$

PROOF. Suppose condition (21) is satisfied and there are no almost-sure fixed points. Define $\tau(x) = f^{-1}(x) - x$. The random function $\tau(x)$ is almost surely cadlag and nondecreasing. The jump made by the Markov process $\widetilde{F}_n(s)$ at step $n$ is precisely $\tau_n(\widetilde{F}_{n-1}(x))$. The expectations $\mathrm{E}\tau(R)$ and $-\mathrm{E}\tau(-R)$ exist and are positive for $R$ sufficiently large.

We would like to see that for any starting point $x$, $F_n^{-1}(x)$ eventually leaves the interval $[-R, R]$ almost surely. For each $x$ in this interval, let

$$\varepsilon^+(x) := \sup\{\varepsilon : \nu\{\tau(x) \ge \varepsilon\} > \varepsilon\},$$
$$\varepsilon^-(x) := \sup\{\varepsilon : \nu\{\tau(x) \le -\varepsilon\} > \varepsilon\}.$$

Note that $\varepsilon^-(x)$ is decreasing in $x$, while $\varepsilon^+(x)$ is increasing. This means that if for a given $x$ $\varepsilon := \max\{\varepsilon^-(x), \varepsilon^+(x)\}$ is not zero, the probability that $\widetilde{F}_n^{-1}(x)$ leaves $[-R, R]$ by time $2R/\varepsilon + 1$ is at least $\varepsilon^{2R/\varepsilon+1}$. Consequently, the process escapes almost surely if $\inf_{|x|\le R} \max\{\varepsilon^-(x), \varepsilon^+(x)\}$ is positive. Suppose then that this infimum is 0. Since $\varepsilon^+$ and $\varepsilon^-$ cannot be simultaneously zero (on pain of violating the no-fixed-points requirement), there must be a unique point $x_0$ such that $\lim_{x\uparrow x_0} \varepsilon^-(x) = 0$ and $\varepsilon^+(x) = 0$ for $x \le x_0$. If $\widetilde{F}_n^{-1}(x)$ is ever below $x_0$, then it moves only downward ever after, so it leaves $[-R, R]$ eventually. Escaping from $[-R, R]$ is thus identical with escaping from $[x_0, R]$. As long as $\widetilde{F}_n^{-1}(x) \in [x_0, R]$, there is a probability at least $\varepsilon(x_0)^{2R/\varepsilon^+(x_0)+1}$ of escape within the next $2R/\varepsilon^+(x_0) + 1$ steps, so the process does escape almost surely.

Let $n_0 = \min\{n : |F_n^{-1}(x)| > R\}$, and suppose without loss of generality that $F_{n_0}^{-1}(x) > R$. Define random variables $T_n := \sum_{i=1}^n \tau_i(R)$. This is a sum of i.i.d. random variables with positive expectation. By the strong law of large numbers, $\lim_{n\to\infty} T_n = +\infty$ almost surely, and there is a positive probability $p$ that $T_n > 0$ for all $n$. On the latter event, $F_{n_0+n}^{-1}(x) \geq R + T_n$, so $\lim_{n\to\infty} F_n^{-1}(x) = +\infty$. Thus, every time the Markov chain $F_n^{-1}(x)$ leaves $[-R, R]$, there is a probability at least $p$ that it never returns, and that $\lim_{n\to\infty} F_n^{-1}(x)$ is $+\infty$ or $-\infty$. It follows that $x$ has an infinite inverse limit almost surely. Since $x$ was arbitrary, it follows by Proposition 4 that $X_\infty$ exists almost surely.

It remains only to show that $X_\infty$ is finite. However, in the above notation,

$$\lim_{t\to\infty} \mathrm{P}\Big\{\min_{n\geq 1} T_n < -t\Big\} = 0,$$

which implies that

$$\lim_{t\to\infty} \mathrm{P}\Big\{\lim_{n\to\infty} F_n^{-1}(R+t) = -\infty\Big\} = 0.$$

Along with the corresponding statement for $t \to -\infty$, this completes the "if" part of the proof.

Suppose now that $\lim_{x\to\infty} \mathrm{E}\tau(x) \leq 0$, which is the same as to say that $\mathrm{E}\tau(x) \leq 0$ for all $x$. The limit $\tau(\infty) := \lim_{x\to\infty} \tau(x)$ exists, and by the monotone convergence theorem it has finite, nonpositive expectation. Given a realization of $(f_i)$,

$$F_n^{-1}(x) \leq x + \sum_{i=1}^{n} \tau_i(\infty).$$

Since the $\tau_i(\infty)$ are i.i.d. random variables with expectation less than or equal to zero, this tells us that $\liminf_{n\to\infty} F_n^{-1}(x) = -\infty$ for all $x$ almost surely. But this means that the system cannot be attractive, since this would require that $\lim_{n\to\infty} F_n^{-1}(x) = +\infty$ for $x > X_\infty$. An identical argument covers the case $\lim_{x\to-\infty} \mathrm{E}\tau(x) \geq 0$.

**5. Convergence rate.** Theorem 1 gives an upper bound on the convergence of $F_n(x)$ to $X_\infty = F_\infty(x)$. In some cases this may in fact be the correct rate of convergence. As in Section 4, I assume the "monotone context" on $\mathbb{R}$ here. I will also use the following special definition: for real numbers $t$ and $R$,

$$p_{t,R} = \inf_{x\geq R} \mathrm{P}\big\{f(x) - x > t\big\},$$

$$p_{t,R}^{*} = \inf_{x\geq R} \mathrm{P}\big\{|f(x) - x| > t\big\}.$$

THEOREM 3. *Let $\nu$ be a strictly monotone system, with $1 \geq f'(x) > 0$ for all $x$ almost surely. Suppose:*

(i) *There exists $R > 0$ such that $f'(x) = 1$ for $|x| \leq R$.*
(ii) *$f(R) - R$ and $-f(-R) - R$ have negative expectation as well as finite exponential moments of some positive order. Also, $f(R) \geq 0$ and $f(-R) \leq 0$.*

(iii) $\mathrm{E}\, f'(x) < 1$ *for* $|x| \le R$.
(iv) *The system has no almost-sure fixed points.*

*Let*

$$r = \inf_{\lambda > 0} \max\bigl\{\mathrm{E}\,\exp\bigl\{\lambda\bigl(f(R) - R\bigr)\bigr\}, \mathrm{E}\,\exp\bigl\{\lambda\bigl(-f(-R) - R\bigr)\bigr\}\bigr\}.$$

*Then the system is attractive, and for all* $x \in \mathbb{R}$,

$$\lim_{n\to\infty} \bigl(\mathrm{E}|F_n(x) - X_\infty|\bigr)^{1/n} = r,$$

*as long as* $r \ge \sup_{|x|<R} \mathrm{E}\bigl[f'(x)\exp\bigl\{\lambda\bigl(|f(x)| - |x|\bigr)\bigr\}\bigr]$.

This result is a simple combination of the upper bounds from Section 3 with the following general lower bound.

PROPOSITION 5. *Suppose*

$$\limsup_{t\to\infty} t^{-1} \log\, p^*_{t,R} < 0 \tag{22}$$

*for some positive* $R$. *Suppose, too, that for some* $c \in (0, 1]$, *and* $\nu$*-almost every* $f$,

$$\inf_{x \ge R} f'(x) \ge c.$$

*Then if* $x \in \mathbb{R}$ *is such that* $\mathrm{P}\bigl\{F_m(x) > R\bigr\}$ *is nonzero for some* $m$,

$$\liminf_{n\to\infty} \bigl(\mathrm{E}|F_n(x) - X_\infty|\bigr)^{1/n} \ge c \inf_{\lambda>0} \int_{-\infty}^{\infty} \lambda e^{\lambda t} p_{t,R}\, dt.$$

PROOF. Let $\xi_1, \xi_2, \dots$ be i.i.d. random variables with $\mathrm{P}\bigl\{\xi_i > t\bigr\} = p_{t,R}$, and let $S_n = \xi_1 + \cdots + \xi_n$. Also define $r^* = \inf_{\lambda>0} \mathrm{E}\bigl[e^{\lambda\xi}\bigr]$. The condition (22) implies that $r^*$ is finite. The set of $x$ such that $\mathrm{P}\bigl\{F_m(x) > R\bigr\}$ is nonzero for some $m$ is an open interval of the form $(y_0, \infty)$. Define $G_n(y) = \mathrm{E}\bigl[F'_n(y)\bigr]$. (This is the same as the old definition of $G_n$, except that the right-derivative is substituted for the local Lipschitz constant.)

I show first that

$$\liminf_{n\to\infty} \inf_{y > y_0} \bigl(G_n(y)\bigr)^{1/n} \ge c r^*.$$

Begin with the relation

$$\begin{aligned} G_{n+1}(y) &= \mathrm{E}\bigl[G_n(f_{n+1}(y)) f'_{n+1}(y)\bigr] \\ &\ge c\mathrm{E}\bigl[G_n(f_{n+1}(y))\bigr]\mathbf{1}_{\{y \ge R\}}. \end{aligned} \tag{23}$$

Suppose that $y \ge R$. Iterating (23) yields

$$G_n(y) \ge c^n \mathrm{P}\bigl\{\widetilde{F}_i(y) \ge R \text{ for } i = 0, \dots, n-1\bigr\}.$$

(Essentially, the bound here is the probability that $\widetilde{F}_n(y)$ stays outside of the central region of contraction until time $n$.) If $y \geq R$, then we can couple the sequences $(\widetilde{F}_n)$ and $(S_n)$ so that $\widetilde{F}_n(y) \geq S_n + R$, as long as $S_n$ stays positive. The problem becomes then one of showing that

$$\liminf_{n\to\infty} \mathrm{P}\left\{\min_{i\leq n} S_i \geq 0\right\}^{1/n} \geq r^*. \tag{24}$$

For the endpoint $S_n$ in place of the minimum this would be simply the Cramér large deviation theorem (cf. [18]). To extend this to the minimum requires the theorem of Varadhan and Mogulskii for the large deviations of random-walk paths (Theorem 5.1.2 of [6]).

It remains only to consider the case $y_0 < y \leq R$. Use the relation

$$\begin{aligned} G_{n+m}(y) &= \mathrm{E}\bigl[G_n(F_m(y))F'_m(y)\bigr] \\ &\geq \mathrm{P}\bigl\{F_m(y) > R,\ F'_m(y) \geq \gamma\bigr\}\gamma \inf_{z>R} G_n(z). \end{aligned}$$

Since $y > y_0$, the probability in the last line above is nonzero for some given $m$ and positive $\gamma$,

$$\begin{aligned} &\liminf_{n\to\infty} G_n(y)^{1/n} \\ &\quad = \lim_{n\to\infty} G_{n+m}(y)^{1/(n+m)} \\ &\quad \geq \bigl(\mathrm{P}\bigl\{F_m(y) > R,\ F'_m(y) \geq \gamma\bigr\}\gamma\bigr)^{1/(n+m)} \liminf_{n\to\infty} \inf_{z>R} G_n(z)^{1/(n+m)} \\ &\quad = \liminf_{n\to\infty} \inf_{z>R} G_n(z)^{1/n}, \end{aligned}$$

which proves the claim.

Now, for every positive $\varepsilon$,

$$\begin{aligned} \mathrm{E}|F_n(x) - F_\infty(x)| &= \mathrm{E}\bigl|\widetilde{F}_n(x) - \widetilde{F}_n(X_\infty)\bigr| \\ &= \mathrm{E}\int_x^{X_\infty} \widetilde{F}'_n(y)\,dy \\ &\geq \varepsilon \mathrm{P}\bigl\{|X_\infty - x| > \varepsilon\bigr\} \inf_{|y-x|\leq\varepsilon} G_n(y), \end{aligned}$$

where $X_\infty$ is taken to be an alternative realization independent of $F_n$. (It is here that the monotonicity of the functions $f$ is essential.) If $X_\infty$ were concentrated at a single point, this would be a fixed point for $f$ almost surely, violating the assumptions. Since $x > y_0$, for $\varepsilon$ sufficiently small,

$$\begin{aligned} \liminf_{n\to\infty} \bigl(\mathrm{E}|F_n(x) - F_\infty(x)|\bigr)^{1/n} &\geq \liminf_{n\to\infty} \Bigl(\varepsilon \mathrm{P}\bigl\{|X_\infty - x| > \varepsilon\bigr\}\Bigr)^{1/n} \inf_{y>y_0} G_n(y)^{1/n} \\ &\geq cr^*. \end{aligned}$$

**6. Stationary systems.** In every one of my specific examples of locally contractive systems, presented in Section 7, the random functions $f_i$ are independent. At the same time, it has long been recognized that contractive iterated function systems reside comfortably in the context of stationary sequences. No new problems arise for locally contractive systems. The purpose of this section is to explain briefly how stationary iterated function systems behave, following the account in, and to restate the main results on local contractivity in this new setting.

We begin now with a two-sided stationary sequence of random Lipschitz functions from $\mathscr{X}$ to itself $(f_n, -\infty < n < \infty)$. Define for nonnegative integers $k$ and $n$,

$$F_{k,n}(x) = f_k \circ f_{k-1} \circ \cdots \circ f_{-n}(x).$$

We say the system is attractive now if $Y_k = \lim_{n\to\infty} F_{k,n}(x)$ exists and is independent of $x$. As is standard for random dynamical systems, when seeking random point attractors (the existence of which is equivalent to the system being attractive) it is necessary to push the starting time back to $-\infty$, rather than push the end time forward to $+\infty$. Elton [7] proved that attractivity follows from the condition that I have here called strong contractivity and showed that the following results hold for attractive stationary systems.

1. The sequence $(Y_k)$ is stationary. Thus $Y_0$ is a random variable of starting values that makes the iteration chain stationary.

2. For any $x \in \mathscr{X}$, the random sequence $\big(F_{k,0}(x), F_{k+1,0}(x), \dots\big)$ converges in distribution as $k$ goes to $\infty$, to $(Y_0, Y_1, \dots)$.

3. If $(f_n)$ is ergodic, then so is $(Y_n)$.

Other results of that paper seem to depend directly upon the strong contractivity.

In the stationary setting, we may define local contractivity as before, substituting $F_{0,n}$ for $F_n$. The growth rate must be slightly redefined, by taking now the expectations conditional on the whole past of the sequence,

$$\begin{aligned} r_x &:= \text{ess sup}\, \mathrm{E}\bigg[\frac{\phi(f_0(x))}{\phi(x)} D_x f_0 \,\Big|\, \mathscr{F}_{-1}\bigg],\\ C_x &:= \text{ess sup}\ \mathrm{E}\big[\|f_0(x) - x\| \Phi(x; f_0(x)) \,\big|\, \mathscr{F}_{-1}\big],\\ G_n(x) &:= \text{ess sup}\ \mathrm{E}\big[D_x F_n \,\big|\, \mathscr{F}_{-n-1}\big], \end{aligned}$$

where $\mathscr{F}_n$ is the $\sigma$-algebra generated by $\{f_n, f_{n-1}, \dots\}$, the infinite past of $f_n$, and the essential supremum is taken with respect to the given probabilities on the set of possible pasts. With these definitions in place, the preceding results remain valid for stationary sequences. For example, if under this new definition $r := \sup r_x$ is smaller than 1 and $G_n(x) \le r^n \phi(x)$, then

$$\begin{aligned}
G_{n+1}(x) &= \operatorname{ess\,sup} \mathrm{E}\big[D_x F_{n+1} \,\big|\, \mathscr{F}_{-n-2}\big] \\
&\le \operatorname{ess\,sup} \mathrm{E}\big[D_{f_{-n-1}(x)} F_n D_x f_{-n-1} \,\big|\, \mathscr{F}_{-n-2}\big] \\
&\le \operatorname{ess\,sup} \mathrm{E}\big[G_n\big(f_{-n-1}(x)\big) D_x f_{-n-1} \,\big|\, \mathscr{F}_{-n-2}\big] \\
&\le \operatorname{ess\,sup} \mathrm{E}\big[r^n \phi\big(f_0(x)\big) F_n D_x f_0 \,\big|\, \mathscr{F}_{-1}\big] \quad \text{(by stationarity)} \\
&\le \phi(x) r^{n+1},
\end{aligned}$$

which proves the stationary analogue to Theorem 2. The other results follow from similar computations. Whether this extends the range of applications in any interesting way is uncertain.

## 7. Applications.

7.1. *Queueing.* The first example is more an illustration than an application, since the problem is simple enough to dispatch by other means. This is the "stability theorem" of Loynes [12], which demonstrates the existence of a "remaining work" process for general single service queues. A single-server queueing system is defined by an i.i.d. sequence of pairs of positive real-valued random variables $(\sigma_n, \tau_n)$, which are to be thought of as the service time and the interarrival time, respectively, of customer number $n$; that is, $\tau_n$ is the time between the arrival of customer $n-1$ and the arrival of customer $n$, while $\sigma_n$ is the time between customer $n$'s arrival at the head of the queue and her departure. Independence of the different pairs is assumed purely for convenience; stationary sequences may be handled as well, as discussed in Section 6. In such a case we would assume that the sequence is ergodic.

An essential scaffolding for the queueing system is the process of "remaining work" when customer $n$ arrives, $\widetilde{W}_n$. In the case of a first-in-first-out service protocol, this is the time between customer $n$'s first arrival and her departure. Loynes proved that $\mathrm{E}\sigma_0 < \mathrm{E}\tau_0$ is a sufficient condition for the existence of a unique, almost-surely finite stationary distribution for this workload process.

Baccelli and Brémaud [3] rederive this result essentially by representing it as an iterated function system, and proving that the reversed system $W_n$ is attractive. It is clear that $\widetilde{W}_n = (\widetilde{W}_{n-1} + \sigma_n - \tau_n)^+$. Defining the random function $f_n(x) = (x + \sigma_n - \tau_n)^+$, it follows that $\widetilde{W}_n = \widetilde{F}_n(0)$.

In fact, it is elementary to show that this system is attractive, with

$$W_\infty = \max_{1 \le n < \infty} \sum_{i=1}^{n} (\sigma_i - \tau_i).$$

The purpose here is simply to point out how this result fits naturally into the context of locally contractive systems, which also offers a free estimate on the rate of convergence.

PROPOSITION 6. *The waiting-time process $W_n$ is attractive if and only if* $\mathrm{E}[\sigma] < \mathrm{E}[\tau]$. *If $\sigma - \tau$ also has subexponential tails, then*

$$\limsup_{n\to\infty} \left(\mathrm{E}|F_n(x) - X_\infty|\right)^{1/n} \le \inf_{\lambda>0} \mathrm{E} \exp\bigl(\lambda(\sigma - \tau)\bigr).$$

PROOF. The first part is a direct application of Corollary 1. The definition gives

$$f_n^{-1}(x) = \begin{cases} x - \sigma_n + \tau_n, & \text{if } x \ge 0, \\ -\infty, & \text{if } x < 0. \end{cases}$$

If $\mathrm{E}\tau > \mathrm{E}\sigma$, it follows that $F_n^{-1}(x)$ will eventually run off to $+\infty$, unless it falls off below 0 and then to $-\infty$ first; the probability of this latter event is smaller than 1 for $x \ge 0$. The second part applies Proposition 3 to the same fact.

7.2. *Zeno's walk in dimension 1.* In [16] I considered the iterated function system where $\nu$ is concentrated on $\{f^+, f^-\}$, given in (4), with $\nu(f^+) = 1 - \nu(f^-) = p$. I showed there that $F_n(0)$ converges almost surely for $\frac{1}{3} < p < \frac{2}{3}$ and computed the rate

$$r_p^* := \lim_{n\to\infty} \left(\mathrm{E}|F_n(0) - X_\infty|\right)^{1/n} = \left(\frac{27p(1-p)\max\{p, 1-p\}}{4}\right)^{1/3}.$$

In addition I gave an explicit description of the distribution of $X_\infty$ and computed its Hausdorff dimension, which is smaller than 1.

By the present methods this rate may be derived more simply. First, the system is locally contractive, hence also attractive, for $\frac{1}{3} < p < \frac{2}{3}$ by Theorem 2. $|F_n(x) - x|$ is bounded, and

$$\mathrm{E}\left[f'(x)\right] = \begin{cases} 1 - \dfrac{p}{2}, & \text{if } -2 \le p < 0, \\ \dfrac{1+p}{2}, & \text{if } 0 \le p < 2, \end{cases}$$

which is always smaller than 1, while

$$\mathrm{E}|f(-2)| = 3(1-p) < 2 \quad \text{for } p > \tfrac{1}{3},$$
$$\mathrm{E}|f(2)| = 3p < 2 \quad \text{for } p < \tfrac{2}{3}.$$

The necessity supplement of Corollary 1 shows furthermore that $F_n(x)$ does not converge for $p \ge \frac{2}{3}$ and $p \le \frac{1}{3}$.

The rate of convergence, by Theorem 3, is

$$\begin{aligned}\lim_{n\to\infty} \left(\mathrm{E}|F_n(x) - X_\infty|\right)^{1/n} &= \max\left\{\inf_{\lambda>0} pe^{\lambda} + (1-p)e^{-2\lambda},\ \inf_{\lambda>0} (1-p)e^{\lambda} + pe^{-2\lambda}\right\} \\ &= \max\left\{\left(\frac{p^2(1-p)}{4}\right)^{1/3}, \left(\frac{p(1-p)^2}{4}\right)^{1/3}\right\}.\end{aligned}$$

More generally, define for positive parameters $r, b$,

$$f^+_{b,r}(x) = \begin{cases} x + b\cdot r, & \text{if } x \le -b\cdot r,\\ \dfrac{x}{r} + b, & \text{if } -b\cdot r < x \le 0,\\ x + b, & \text{if } 0 < x,\end{cases}$$

and $f^-_{b,r}(x) = -f^+_{b,r}(-x)$. Then define an iterated function system by a distribution on $(\pm, b, r)$. The system is attractive whenever $\mathrm{E}[1/r] < 1$, and

$$\mathrm{E}[b|+]\mathrm{P}\{+\} < \mathrm{E}[b\cdot r|-]\mathrm{P}\{-\}$$

and

$$\mathrm{E}[b|-]\mathrm{P}\{-\} < \mathrm{E}[b\cdot r|+]\mathrm{P}\{+\}.$$

In particular, if $\pm$, $b$ and $r$ are independent, with $\mathrm{P}\{+\} = p = 1 - \mathrm{P}\{-\}$, the system is attractive if $\mathrm{E}[1/r] < 1$, and

$$\mathrm{E}[r] > \max\left\{\frac{p}{1-p}, \frac{1-p}{p}\right\}.$$

Conversely, if

$$\mathrm{E}[r] < \max\left\{\frac{p}{1-p}, \frac{1-p}{p}\right\},$$

then there is no convergence. In the former case, if $b$ is deterministic we can compute that

$$\lim_{n\to\infty}\left(\mathrm{E}|F_n(x) - X_\infty|\right)^{1/n} = (1-p_*)\mathrm{E}\big[(1+r)\exp(-br\lambda*)\big],$$

where $p_* = \max\{p, 1-p\}$ and $\lambda_*$ is defined to satisfy

$$\exp(b\lambda*) = \frac{(1-p*)}{p*}\mathrm{E}\big[r\,\exp(-br\lambda_*)\big],$$

as long as $\mathrm{E}\big[\exp(-br\lambda_*)\big] \ge \mathrm{E}\big[\frac{1}{r}\big]\exp(b\lambda_*)$. (If $r$ is deterministic, this last condition is automatically satisfied.)

Now take $r$ to be deterministic, and $b \equiv 1$. Then the system is attractive precisely when $r > 1$. The methods of [16] will also describe the distribution of $X_\infty$ explicitly when $r$ is an integer.

7.3. *Zeno's walk in higher dimensions.* Given a positive constant $\rho$, and $\zeta \in \mathbb{R}^d$, we define a function $f^{\rho,\zeta}\colon \mathbb{R}^d \to \mathbb{R}^d$ by

$$f^{\rho,\zeta}(x) = \frac{1}{2}\zeta + \left(x + \frac{\rho}{2}\zeta\right)\phi_\rho\left(\left\|x + \frac{\rho}{2}\zeta\right\|\right), \tag{25}$$

where

$$\phi_\rho(t) = \begin{cases} \dfrac{1}{\rho}, & \text{if } t \le \dfrac{\rho}{2},\\ 1 - \dfrac{\rho - 1}{2t}, & \text{if } t > \dfrac{\rho}{2}.\end{cases}$$

Note that $f^{1,\zeta}(x) = x + \zeta$. For $\rho > 1$, $f^{\rho,\zeta}$ may be understood as a two-stage process: first, there is a translation by $\zeta$, then the disk whose diameter is $(0, \zeta)$ is shrunk by a factor of $\rho$ from its center. That is, in the inverse map, from the point of view of the walker, all distances within the disk are increased by a factor of $\rho$: $f(0) = \zeta$ and $f(-\rho\zeta) = 0$.

I define now an iterated function system by fixing $\rho > 1$, and choosing $\zeta$ according to some given distribution on the unit sphere. For $d = 1$, this is just Zeno's walk. Since $\|D_x f^{\rho,\zeta}\| \le 1$, Theorem 2 tells us that the system is attractive if

$$\mathrm{E}\|f^{\rho,\zeta}(x)\| < \rho \tag{26}$$

for $\|x\| = \rho$. Call this quantity $E(\rho, \theta)$, a function of $\rho$ and of the angle $\theta$ between $\zeta$ and $x$. By an application of the law of cosines,

$$E(\rho, \theta) = \left\{ a(\rho, \theta)^2 + a(\rho, \theta)\frac{\cos\theta - \frac{1}{2}}{\sqrt{\frac{5}{4} - \cos\theta}} + \frac{1}{4} \right\}^{1/2}$$

where

$$a(\rho, \theta) := \sqrt{\frac{5}{4} - \cos\theta} - \frac{\rho - 1}{2}.$$

Let us choose $\zeta$ uniformly on the unit sphere. The criterion for local contractivity then becomes

$$\int_0^{2\pi} \big(E(\rho, \theta) - \rho\big)|\sin\theta|^{d-2} d\theta < 0. \tag{27}$$

(The derivative actually goes to 1 at the boundary of the $\|x\| \le \rho$ disc, but we can take the critical radius $R$ to be $\rho - \varepsilon$ for $\varepsilon$ arbitrarily small.)

This integrand is decreasing in $\rho > 1$, so there is attractivity for all $\rho$ larger than a critical value $\rho_d$. One may compute that $\rho_2 \approx 1.60$, $\rho_3 \approx 1.76$, and it is easy to see that $\lim_{d\to\infty} \rho_d \approx 2.025$. For $d = 1$ the system is attractive for all $\rho > 1$. In higher dimensions it is unclear what happens for $\rho$ between 1 and $\rho_d$.

7.4. *Random logistic maps.* In [5] it was shown that if $y_i$ are chosen i.i.d. from a $\beta_{a+1/2,\,a-1/2}$ distribution, for $a > 1/2$, and $f_i\colon [0, 1] \to [0, 1]$ is defined by $f_i = f_{y_i}$, where

$$f_y(x) = 4yx(1 - x),$$

then $\beta_{a,a}$ is a stationary distribution for the resulting iterated function system $\widetilde{F}_n(x)$. It was not determined there whether there was any convergence.

PROPOSITION 7. *With the above definitions, the Markov chain $\widetilde{F}_n(x)$ converges in distribution to a unique stationary distribution for all $a$. If $a \ge 2$, the system is locally contractive, hence also attractive.*

PROOF. For $x \neq 0$, almost surely $\widetilde{F}_n(x)$ is never 0 or 1, so we may view the process as occurring on the open interval (0,1). (This makes the chain irreducible.)

By Theorem 9.2.2 in [13], the chain is Harris recurrent if there is a compact subset of (0,1) to which the chain returns infinitely often with probability 1; that is, for $x$ outside the subset, there is almost surely some $n$ such that $\widetilde{F}_n(x)$ is inside. I will show that this is the case for the interval $\mathscr{I} = [0.1, 0.9]$. Given $x \in (0, 0.1)$, $0.36 > 4yx(1-x) > 3.6yx$. Writing $f_n(x) = 4y_n x(1-x)$, it must be that either $\widetilde{F}_n(x) \geq 0.1$ for some $n \leq N$, or

$$\log \widetilde{F}_N(x) > \log(3.6y_1) + \log(3.6y_2) + \cdots + \log(3.6y_N).$$

Thus $\widetilde{F}_n(x) < 0.1$ for all $n$ only if $\sum_{i=1}^n \log(3.6y_i) < 0.1$ for all $n$. But this happens with probability 0, since $\log(3.6y_i)$ are i.i.d. random variables with positive expectation. [Letting $\psi$ be the digamma function, the derivative of $\log\Gamma$, $\mathrm{E}\log y_i = \psi(a+1/2) - \psi(2a)$, which is always larger than $-\log 3.6$.] Thus there almost surely exists $n$ such that $\widetilde{F}_n(x) \geq 1$, and this value must lie in the interval $[0.1, 0.36]$. If $x \in (0.9, 1)$, then $\widetilde{F}_1(x) \in (0, 0.36)$; if it is smaller than 0.1, continue as before.

Since the Markov chain $\widetilde{F}_n(x)$ is Harris recurrent it has a unique invariant measure, by Theorem 10.4.4 of Meyn and Tweedie [13]. Since the stationary distribution is finite, the chain converges in distribution. (The idea of using a recurrence argument here was suggested by R. Durrett.)

Now consider the issue of local contractivity. Define for $r > 0$,

$$\phi_r(x) = (x - x^2)^{-r}.$$

This $\phi_r$ is differentiable on (0,1), and

$$\sup_{t\in[x, f_y(x)]} \phi_r(t) \leq \phi_r(x) + \phi_r(f_y(x)),$$

so

$$\begin{aligned}
\mathrm{E}\left[\sup_{t\in[x, f_y(x)]} \phi_r(t)\right] &\leq \phi_r(x) + B(a - \tfrac{1}{2}, a + \tfrac{1}{2}) \int_0^1 y^{a-1/2}(1-y)^{a-3/2}(y-y^2)^{-r}\, dy \\
&\leq \phi_r(x) + B(a - \tfrac{1}{2}, a + \tfrac{1}{2}) \int_0^1 y^{a-1/2-r}(1-y)^{a-3/2-r}\, dy,
\end{aligned}$$

which is finite for $r < a - \frac{1}{2}$. $\big[$Here $B(\alpha, \beta) = (\Gamma(\alpha+\beta)/\Gamma(\alpha)\Gamma(\beta)).\big]$ Thus, $C_x < \infty$. For the other condition in Theorem 1, we have

$$\begin{aligned}
\frac{\mathrm{E}[\phi(f_y(x))|f'_y(x)|]}{\phi(x)} &= \frac{\mathrm{E}[(4Yx(1-x))^{-r}(1-4Yx(1-x))^{-r}4Y|1-2x|]}{(x(1-x))^{-r}} \\
&= 4^{1-r}|1-2x|\mathrm{E}[Y^{1-r}(1-4Yx(1-x))^{-r}] \\
&= 4^{1-r}\frac{\Gamma(2a)\Gamma(a+\frac{3}{2}-r)}{\Gamma(a+\frac{1}{2})\Gamma(2a+1-r)} \\
&\quad \times |1-2x|\,{}_2F_1\Big(\tfrac{3}{2}+a-r, r; 2a+1-r; 4x(1-x)\Big).
\end{aligned}$$

It follows that for any given $a$, the system is attractive if there exists $r < a - \frac{1}{2}$ such that this expression is smaller than 1 for all $x$. It remains then only to show that this is the case if $a \geq 2$. (Actually, this can be pushed down to about $a = 1.98$.)

The details of this computation, while offering a grand tour of hypergeometric arcana, are hardly interesting or germane to the central topic of this paper. I will only sketch them here.

By a change of variables, the problem becomes one of showing for all $a \geq 2$ and some $r \leq a - 0.5$, that $\sup_{z\in[0,1]} \rho_{a,r}(z) < 1$, where

$$\rho_{a,r}(z) := 4^{1-r}\sqrt{1-z}\frac{\Gamma(2a)\Gamma(a+\frac{3}{2}-r)}{\Gamma(a+\frac{1}{2})\Gamma(2a+1-r)}{}_2F_1\Big(\tfrac{3}{2}+a-r, r; 2a+1-r; z\Big).$$

It will suffice to take the value $r = 1.35$ for all $a$. We are considering then the function

$$\rho_a(z) := 4^{-0.35}\sqrt{1-z}\frac{\Gamma(2a)\Gamma(a+0.15)}{\Gamma(a+0.5)\Gamma(2a-0.35)}{}_2F_1(0.15+a, 1.35; 2a-0.35; z).$$

For fixed $a$ and $z$, of course, this is merely a numerical calculation. The problem is to prove that the bound holds uniformly. If $a$ is fixed, it is still no problem to demonstrate that $\rho_a(z) < 1$ for all $z$. Given two monotone increasing real-valued functions $f$ and $g$, with $f(1) < g(1)$, to show that $f(z) < g(z)$ for all $z \in [0,1]$ is equivalent to showing that successive iterates of the function $\tilde{\rho}(z) = g^{-1}(f(z))$, eventually become negative. Here the functions are $g(z) = (1-z)^{-1/2}$ and

$$\tilde{\rho}(z) = 1 - \left(4^{-0.35}\frac{\Gamma(2a)\Gamma(a+0.15)}{\Gamma(a+0.5)\Gamma(2a-0.35)}{}_2F_1(a+0.15, 1.35; 2a-0.35; z)\right)^{-2}.$$

It is then straightforward to iterate this on a computer and find that $\tilde{\rho}^{(83)}(1) < 0$ for $a = 2$.

We could similarly test any particular value of $a$, and in fact, as $a$ gets larger, the maximum of $\rho$ only seems to get smaller. On the other hand, $\rho$ itself is not decreasing in $a$, and showing that the bound holds for all $a$ requires more effort. The second step is then to show that the bound holds for all $a \geq 2$ when $z$ is sufficiently small.

In the power-series representation of the hypergeometric function ${}_2F_1$ (which may be found, for example in Chapter 15 of [1]), the coefficient of $z^n$ is decreasing in $n$, so the whole sum is smaller than the first coefficient, which is 1, multiplied by $(1-z)^{-1}$. Thus

$$\begin{aligned}\rho_a(z) &\leq 4^{-0.35}\frac{\Gamma(2a)\Gamma(a+0.15)}{\Gamma(a+0.5)\Gamma(2a-0.35)}(1-z)^{-1/2}\\ &\leq \lim_{a\to\infty} 4^{-0.35}\frac{\Gamma(2a)\Gamma(a+0.15)}{\Gamma(a+0.5)\Gamma(2a-0.35)}(1-z)^{-1/2}\\ &= 2^{-0.35}(1-z)^{-1/2},\end{aligned}$$

and $\rho_a(z) < 1$ for all $z < 1 - 2^{-0.7} \approx 0.384$.

The final step will then be to show that $\rho_a(z)$ is decreasing in $a$ for each fixed $z > 0.38$. To do this, write

$$\rho_a(z) = c(z) \sum_{n=0}^{\infty} g_n(a) z^n,$$

where

$$g_n(a) = \frac{\Gamma(2a)\Gamma(a + 0.15 + n)\Gamma(1.35 + n)}{\Gamma(a + 0.5)\Gamma(2a - 0.35 + n)\Gamma(n + 1)}.$$

It will suffice to show now that

$$\begin{aligned} &g_0'(a) > 0, \\ &g_1'(a) < -2g_0'(a), \\ &g_2'(a) < -2g_0'(a), \\ &g_n'(a) < 0 \text{ for } n \geq 3, \end{aligned} \tag{28}$$

since that will then imply that

$$\begin{aligned} \frac{d}{da}\rho_a(z) &\leq c(z) \sum_{n=0}^{\infty} g_n'(a) z^n \\ &\leq c(z) g_0'(a)(1 - 2z - 2z^2) \\ &< 0 \quad \text{for } z > 0.37. \end{aligned}$$

The relations (28) may be established by using known properties of the digamma function, again, found in [2]. $\square$

## 8. Open questions.

8.1. *Necessary conditions for convergence.* There is a wide gap between the sufficient conditions for attractivity given in Theorem 2 and any known necessary conditions. The problem is twofold: first, failure of the conditions of Theorem 2 does not necessarily imply that the system is not locally contractive; second, even if the system is not locally contractive, it does not follow that the system is not attractive. For instance, in the $d$-dimensional Zeno's walk with expansion factor $\rho$ below the critical $\rho_d$, there is a ball in which the random function $f$ is average contractive ($Df < 1$), an unbounded outer region in which $\mathrm{E}[\|f(x)\| - \|x\|] < 0$ and a band in between where neither condition holds. Does this prevent convergence?

For the random logistic functions, the situation is perhaps even more interesting. There it does seem that attractivity should fail when the parameter $a$ is very small, smaller than 1, but there still gapes a chasm between that realm and the domain in which attractivity is proved. What is more, while it is indeed possible to show that the system is not locally contractive for $a \leq 1$, this does not immediately rule out the possibility that the system could still be attractive.

PROPOSITION 8. *If the iterated function system on $\mathscr{X} \subset \mathbb{R}$ defined by $\nu$ is attractive, then*

$$\frac{1}{n} \log |F'_n(x)| \to_P \int_{\mathscr{F}} \int_{\mathscr{X}} \log \left|f'(x)\right| d\nu(f)\, d\mu(x),$$

*for $\mu$-almost every $x$, where $\mu$ is the distribution of $X_\infty$. (Here $\to_P$ means convergence in probability).*

PROOF. The almost-sure convergence of $F_n(x)$ implies that the $\mu$-preserving random transformation $f$ is ergodic, from which follows that the $\mu \otimes \nu$-preserving skew-product tranformation,

$$\tau\big(x, (f_1, f_2, \ldots)\big) = \big(f_1(x), (f_2, f_3, \ldots)\big)$$

is also ergodic (Theorem 2.1 of [10]). Let

$$\delta\big(x, (f_1, f_2, \ldots)\big) = \log \left|f'_1(x)\right|.$$

Then

$$\begin{aligned}
\log \left|\widetilde{F}'_n(x)\right| &= \log \left|f'_n\big(\widetilde{F}'_{n-1}(x)\big)\right| + \log \left|\widetilde{F}'_{n-1}(x)\right| \\
&= \delta\Big(\tau^{n-1}\big(x, (f_1, \ldots, f_n, \ldots)\big)\Big) + \log \left|\widetilde{F}'_{n-1}(x)\right| \\
&= \sum_{i=0}^{n-1} \delta\Big(\tau^i\big(x, (f_1, \ldots)\big)\Big).
\end{aligned}$$

The result then follows by Birkhoff's additive ergodic theorem, using the identity in distribution of $F_n$ and $\widetilde{F}_n$.

COROLLARY 2. *If $\mu$ is a stationary distribution for $\nu$ (where $\mathscr{X} \subset \mathbb{R}$), and the iterated function system defined by $\nu$ is locally contractive, then the Lyapunov exponent*

$$\int_{\mathscr{X}} \int_{\mathscr{X}} \log \left|f'(x)\right| d\nu(f)\, d\mu(x)$$

*is negative.*

Taking $\mu = \beta_{a,a}$, $\nu$ the image of $\beta_{a+1/2,\, a-1/2}(dy)$ on $f_y(x) = 4yx(1-x)$, it is straightforward to compute

$$\begin{aligned}
&\int_0^1 \int_0^1 \log(4y|1-2x|)\, d\beta_{a+1/2,\, a-1/2}(y)\, d\beta_{a,a}(x) \\
&\quad = \log 4 + \int_0^1 \log y\, d\beta_{a+1/2,\, a-1/2}(y) + \int_0^1 \log|1-2x|\, d\beta_{a,a}(x) \\
&\quad = \log 4 + \left(\psi\left(a + \frac{1}{2}\right) - \psi(2a)\right) \\
&\qquad + 4^{-a} \frac{\Gamma(\frac{1}{2})\Gamma(2a)}{\Gamma(a)\Gamma(a+\frac{1}{2})} \left(\psi\left(\frac{1}{2}\right) - \psi\left(a + \frac{1}{2}\right)\right),
\end{aligned}$$

where $\psi(a)$ is the digamma function. This expression is decreasing in $a$ and takes on the value 0 at $a = 1$. It follows then that the system is not locally contractive for $a \leq 1$.

It is still unclear whether the system is locally contractive (or attractive) for $a$ between 1 and 2, though simulations suggest strongly that it is. I have also been so far unable to prove that the system is not attractive for $a$ smaller than 1. (This problem is resolved in a new paper [17].) In this range, Proposition 8 says that the derivative of $F_n$ increases exponentially with $n$, at almost every point. Intuitively, it seems impossible that $F_n$ could still be converging under such conditions to a constant function, uniformly on compact subintervals of [0,1]; again, simulations support this belief, but it remains unproved.

8.2. *Rate of convergence.* What is the correct rate of convergence for the distribution of $F_n(x)$? For locally contractive systems, Theorem 1 gives an upper bound for $\mathrm{E}\|F_n(x) - X_\infty\|$, which is in turn an upper bound for $W(F_n(x), X_\infty)$. In the one-dimensional monotone setting, Proposition 5 provides a lower bound on this expectation, and in some conditions, given in Theorem 3, the upper and lower bounds are of the same exponential order. In this case, the Wasserstein distance to the limit is bounded by

$$\lim_{n\to\infty} W(F_n(x), X_\infty)^{1/n} \leq \lim_{n\to\infty} \bigl(\mathrm{E}\bigl|F_n(x) - X_\infty\bigr|\bigr)^{1/n} = r^* := \inf_\lambda r^*(\lambda),$$

where $r^*(\lambda) = \sup_{x\in\mathbb{R}} \mathrm{E}\bigl[\exp(|f(x)| - |x|)f'(x)\bigr]$. But this is still only an upper bound. While the expected distance from the limit point is itself an interesting measure of the rate of convergence, there are reasons to be less than satisfied with this answer and to suppose that the real rate of convergence of the Wasserstein distance may be faster. Remember that the lower bound on the expectation was essentially just the probability that no contraction had occurred up to time $n$; that is, that the expectation of the future jumps is still on the order of 1. The typical case will see quicker convergence, though this is swamped in the expectation by the contribution of relatively enormous values on a very small set.

Considering $G_n^*(x) := \mathrm{E}\log F_n'(x)$ offers a different picture. There the very small sets for which $F_n'(x)$ is close to 1 lose their influence, since $\log F_n'(x)$ is small as well. There is now the relation

$$G_{n+1}^*(x) = \mathrm{E}G_n^*(f(x)) + \mathrm{E}\log f'(x). \tag{29}$$

If we define

$$r_*(\lambda) := \inf_{x\in\mathscr{X}} \exp\{\lambda\mathrm{E}|f(x)| - \lambda|x| + \mathrm{E}\log f'(x)\},$$

then for all $n$ and $x \in \mathbb{R}$,

$$G_n^*(x) \leq \lambda|x| + r_* n.$$

By Jensen's inequality it is easy to see that $r_*(\lambda) \le r^*(\lambda)$. For Zeno's walk with $p = \frac{1}{2}$, $r_* := \inf_\lambda r_*(\lambda) \approx 0.89$, while $r^* \approx 0.96$.

If we were to look at $G_n^* = \sup_x G_n^*(x)$, we could apply the subadditive ergodic theorem to establish that $(1/n)G_n^*$ converges almost surely. Unlike the strongly contractive setting, though, for locally contractive systems this supremum remains 0, so this does not provide any information. On the other hand, if $(1/n)G_n^*(x)$ did converge almost surely to $\log r_*$ for fixed $x$, this would suggest that $r_*$ is a lower bound for $\lim_{n\to\infty} W(F_n(x), X_\infty)$. The question remains, whether $r^*$, $r_*$, or something in between, is the correct rate of convergence of the distributions.

It should also be noted that $r^*$ was computed essentially as the spectral radius of the operator $\Gamma$ defined by

$$\Gamma\phi(x) = \mathrm{E}\big[\phi(f(x))f'(x)\big].$$

It is unclear (and perhaps worth understanding) why the spectral radius of this operator gives us information about the spectral gap of the Markov operator (of the reverse process $\widetilde{F}_n$),

$$\widetilde{\Gamma}\phi(x) = \mathrm{E}\big[\phi(f(x))\big].$$

It is obtained here by means of a coupling which seems arbitrary, as well as peripheral to the Markov chain. For instance, the proof of Theorem 3 shows that in some cases the same result is obtained by replacing $f'$ in the definition of $\Gamma$ by the indicator of $\{f'(x) = 1\}$.

Another problem is to describe better the rate of convergence in higher-dimensional settings.

8.3. *The limiting distribution.* In the case of Zeno's walk, with expansion factor 2 and probabilities of jumping right or left $p$ and $1-p$, respectively, I described the limiting distribution in [16], and computed its Hausdorff dimension to be

$$\left(1 + \frac{p}{1-p}\alpha_{1-p}^2\right)|\log_4 \alpha_{1-p}| + \left(1 + \frac{1-p}{p}\alpha_p^2\right)|\log_4 \alpha_p|,$$

where

$$\alpha_p = -\frac{1}{2} + \frac{1}{2}\sqrt{\frac{4p}{1-p} + 1}.$$

The same method would work for any integral expansion factor. It may be that, for processes on the real line, the distribution of $X_\infty$ is singular with respect to Lebesgue measure and has Hausdorff dimension smaller than 1 if and only if for every $x$ and $n$ the support of the random variable $F_n(x)$ generates a discrete subgroup of $\mathbb{R}$. For Zeno's walk, this would mean that the limiting measure is singular if the expansion factor is rational. On the other hand, this problem is at least superficially similar to the famous Erdős problem, of determining when the distribution of $\sum_{i=1}^{\infty} \pm\lambda^i$ is absolutely continuous with respect to Lebesgue measure (where $\lambda$ is a fixed constant between $\frac{1}{2}$ and 1 and the signs are chosen independently with probability $\frac{1}{2}$). Although there has been significant progress on this recently (e.g., [15]), there is still no way to answer the question for a given choice of $\lambda$. It would be interesting to try to determine for which choices of parameters in Zeno's walk the limit distribution is singular, and to compute, or at least estimate, the Hausdorff dimensions in such cases.

It is easy to see that

$$\operatorname{supp} X_\infty = \bigcup_{f \in \operatorname{supp} \nu} f(\operatorname{supp} X_\infty),$$

where the support of $\nu$ is understood to be taken with respect to the uniform topology on functions. For Zeno's walk, this implies that the support of $X_\infty$ is all of $\mathbb{R}^d$. What are more general conditions for this to be the case? And which sets can arise as supports?

## REFERENCES


[1] ABRAMOWITZ, M. and STEGUN, I. (1965). *Handbook of Mathematical Functions, with Formulas, Graphs, and Mathematical Tables*. Dover, New York.

[2] ARNOLD, L., and CRAUEL, H. (1992). Iterated function systems and multiplicative ergodic theory. In *Diffusion Processes and Related Problems in Analysis* (M. Pinksy and V. Wihstutz, eds.) **2** 283–305. Birkhäuser, Boston.

[3] BACCELLI, F. and BREMAUD, P. (1994). *Elements of Queueing Theory: Palm–Martingale Calculus and Stochastic Recurrences*. Springer, New York.

[4] BARNSLEY, M. F. and ELTON, J. H. (1988). A new class of Markov processes for image encoding. *Adv. in App. Probab.* **20** 14–32.

[5] CHAMAYOU, J. F. and LETAC G. (1991). Explicit stationary distributions for compositions of random functions and products of random matrices. *J. Theoret. Probab.* **4** 3–36.

[6] DEMBO, A. and ZEITOUNI, O. (1993). *Large Deviation Techniques and Applications*. Jones and Bartlett, Boston.

[7] ELTON, J. H. (1990). A multiplicative ergodic theorem for Lipschitz maps. *Stochastic Process. Appl.* **34** 39–47.

[8] HUTCHINSON, J. E. Fractals and self-similarity. *Indiana Univ. Math. J.* **30** 713–747.

[9] KHASMINSKII, R. Z. (1990). *Stochastic Stability of Differential Equations*. Sijthoff & Noordhoff, Alphen aan den Rijn, Netherlands.

[10] KIFER, Y. (1986). *Ergodic Theory of Random Transformations*. Birkhäuser, Boston.

[11] LETAC, G. (1986). A contraction principle for certain Markov chains and its applications. *Contemp. Math.* **50** 263–273.

[12] LOYNES, R. (1962). The stability of queues with nonindependent interarrival and service times. *Proc. Cambridge Philos. Soc.* **58** 497–520.

[13] MEYN, S. and TWEEDIE, R. (1993). *Markov Chains and Stochastic Stability*. Springer, New York.

[14] PROPP, J. and WILSON, D. (1998). Coupling from the past: a user's guide. In *Microsurveys in Discrete Probability* (D. Aldons and J. Propp, eds.) 181–192. Aer. Math. Soc., Providence, RI.

[15] SOLOMYAK, B. (1995). On the random series $\sum \pm\lambda^n$ (an Erdős problem). *Ann. Math.* **142** 611–625.

[16] STEINSALTZ, D. (1997). Zeno's walk: a random walk with refinements. *Probab. Theory Related Fields* **107** 99–121.

[17] STEINSALTZ, D. (1999). Random interval maps and Lyapunov exponents. Unpublished Manuscript.

[18] STROOCK, D. W. (1993). *Probability Theory, an Analytic View*. Cambridge Univ. Press.

[19] SHORACK, G. and WELLNER, J. *Empirical Processes with Applications to Statistics*. Wiley, New York.



DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720
E-MAIL: dstein@stat.berkeley.edu