Economic Theory **23**, 13–38 (2004)
DOI: 10.1007/s00199-003-0357-4

Economic Theory

Random dynamical systems: a review*

Rabi Bhattacharya[1] **and Mukul Majumdar**[2]

[1] Department of Mathematics, University of Arizona, Tucson, AZ 85721, USA
(e-mail: rabi@math.arizona.edu)
[2] Department of Economics, Cornell University, Ithaca, NY 14853, USA
(e-mail: mkm5@cornell.edu)

Received: August 16, 2002; revised version: January 9, 2003

Summary. This paper provides a review of some results on the stability of random dynamical systems and indicates a number of applications to stochastic growth models, linear and non-linear time series models, statistical estimation of invariant distributions, and random iterations of quadratic maps.

Keywords and Phrases: Random dynamical systems, Stability, Splitting, Contractions, Stochastic growth, Time series, Estimation.

JEL Classification Numbers: C1, C6, D8, D9.

1 Introduction

In this review we bring together some results on random dynamical systems and indicate their applications to economics. A random dynamical system is described by a triplet (S, Γ, Q) where S is the state space, Γ an appropriate family of maps from S into itself (interpreted as the family of all admissible laws of motion), and Q is a probability distribution on (some σ-field of) Γ. The evolution of the system is described somewhat informally as follows: initially, the system is in some state x in S; an element α_1 of Γ is chosen randomly according to the distribution Q, and the system moves to the state $X_1 = \alpha_1(x)$ in period one. Again, independently of α_1, α_2 is chosen from Γ according to Q, and the state of the system in period two is obtained as $X_2 = \alpha_2(X_1)$, and the story is repeated. The primary focus of the review is the behavior of the random variables X_n starting from an arbitrary initial state. A precise description of the questions that are taken up follows after a more formal account of the structure of (S, Γ, Q).

* Thanks are due to A. Goswami and B.V. Rao for their detailed comments on an earlier version.
Correspondence to: M. Majumdar

Let S (the state space) be a Borel subset of a complete separable metric space and $\mathcal{S}$ be the Borel σ-field of S. Endow Γ (the laws of motion: a family of maps from S into S) with a σ-field Σ such that the map $(\gamma, x) \to \gamma(x)$ on $(\Gamma \times S,\ \Sigma \otimes \mathcal{S})$ into $(S, \mathcal{S})$ is measurable. Let Q be a probability measure on (Γ, Σ). On some probability space $(\Omega, \digamma, P)$, let $(\alpha_n)_{n=1}^{\infty}$ be a sequence of i.i.d. random functions from Γ with a common distribution Q. For a given random variable X_0 (with values in S), independent of the sequence $(\alpha_n)_{n=1}^{\infty}$, define

$$X_1 \equiv \alpha_1(X_0) \equiv \alpha_1 X_0, \tag{1.1}$$

$$X_{n+1} = \alpha_{n+1}(X_n) \equiv \alpha_{n+1}\alpha_n ... \alpha_1 X_0. \tag{1.2}$$

We write $X_n(x)$ for the case $X_0 = x$. Then X_n is a Markov process with a stationary transition probability $p(x,\, dy)$ given as follows: for $x \in S$, $C \in\ \mathcal{S}$,

$$p(x, C) = Q(\{\gamma \in \Gamma : \gamma(x) \in C\}). \tag{1.3}$$

In this review we focus primarily on *stable* random dynamical systems. The Markov process X_n [defined by (1.2)] is *stable* if X_n converges *weakly* (in distribution) to a unique invariant distribution π (see (2.2)) irrespective of the initial X_0. In Section 3.1 we elaborate on two properties: 'splitting' and 'contraction'. In Section 3.2 we derive one of the "classical" results on the stability of Markov processes (Doeblin's theorem) as an application of Theorem 3.1. In Section 4, we sketch some further applications of our results to models of stochastic growth (descriptive as well as optimal), linear and non-linear time series models, statistical estimation of invariant distributions and random iterates of quadratic maps.

Of the main results, Theorem 3.3 on random Lipschitz maps seems to be new in its present generality.

2 Evolution

To study the evolution of the process (1.2), it is convenient to define the map T^* [on the space $M(S)$ of all finite signed measures on $(S,\, \mathcal{S})$] by

$$T^*\mu(C) = \int_S p(x, C)\mu(dx) = \int_\Gamma \mu(\gamma^{-1}C)Q(d\gamma), \qquad \mu \in M(S). \tag{2.1}$$

Let $\mathcal{P}(S)$ be the set of all probability measures on $(S,\, \mathcal{S})$. An element π of $\mathcal{P}(S)$ is *invariant* for $p(x,\, dy)$ (or for the Markov process X_n) if it is a fixed point of T^*, i.e.,

$$\pi\ is\, invariant \qquad iff\ \ T^*\pi = \pi. \tag{2.2}$$

Now write $p^{(n)}(x,\, dy)$ for the n-step transition probability with $p^{(1)} \equiv p(x,\, dy)$. Then $p^{(n)}(x,\, dy)$ is the distribution of $\alpha_n .....\alpha_1 x$. Define T^{*n} as the n-th iterate of T^*:

$$T^{*n}\mu = T^{*(n-1)}(T^*\mu)\ (n \geq 2),\, T^{*1} = T^*,\, T^{*0} = Identity. \tag{2.3}$$

Then for any $C \in \mathcal{S}$,

$$(T^{*n}\mu)(C) = \int_S p^{(n)}(x, C)\mu(dx), \tag{2.4}$$

so that $T^{*n}\mu$ is the distribution of X_n when X_0 has distribution μ. To express T^{*n} in terms of the common distribution Q of the i.i.d. maps, let Γ^n denote the usual Cartesian product $\Gamma \times \Gamma \times ... \times \Gamma$ (n terms), and let Q^n be the product probability $Q \times Q \times ... \times Q$ on $(\Gamma^n, \Sigma^{\otimes n})$ where $\Sigma^{\otimes n}$ is the product σ-field on Γ^n. Thus Q^n is the (joint) distribution of $\alpha = (\alpha_1, \alpha_2, ..., \alpha_n)$. For $\gamma = (\gamma_1, \gamma_2, ..., \gamma_n) \,\epsilon\, \Gamma^n$ let $\widetilde{\gamma}$ denote the composition

$$\widetilde{\gamma} := \gamma_n \gamma_{n-1} ... \gamma_1. \tag{2.5}$$

Then, since $T^{*n}\mu$ is the distribution of $X_n = \alpha_n ... \alpha_1 X_0$, one has, by the independence of $\tilde{\alpha}$ and X_0, $(T^{*n}\mu)(A) = \text{Prob}(X_0 \,\epsilon\, \widetilde{\alpha}^{-1} A)$, where $\widetilde{\alpha} = \alpha_n \alpha_{n-1} .... \alpha_1$.

$$(T^{*n}\mu)(A) = \int_{\Gamma^n} \mu(\widetilde{\gamma}^{-1} A) Q^n(d\gamma) \;\; (A \,\epsilon\, \mathcal{S},\, \mu \,\epsilon\, \mathcal{P}(S)). \tag{2.6}$$

The stationary transition probability $p(x, dy)$ is said to be *weakly continuous* or to have the *Feller property* if for any sequence x_n converging to x, the sequence of probability measures $p(x_n, \cdot)$ converges weakly to $p(x, \cdot)$. It is easy to show that if Γ consists of a family of continuous maps, $p(x, dy)$ has the Feller property.

Finally, we come to the definition of *stability*. A Markov process X_n is *stable in distribution* if there is a unique invariant probability measure π and $p^{(n)}(x, dy)$ converges weakly to π for all x. In the case one has $(1/n) \sum_{m=1}^{n} p^{(m)}(x, dy)$ converging weakly to the same invariant π for all x, we may define the Markov Process to be *stable in distribution on the average*.

3 Splitting and contraction

In this section we review two classes of random dynamical systems that are stable in distribution. In Section 3.1 we deal with the concept of splitting. We apply the basic Theorem 3.1 to derive a well-known result on the stability of Markov processes (Corollary 3.3), and then move on to Lipschitz maps in Section 3.3.

3.1 A general theorem under splitting

Recall that $\mathcal{S}$ is the Borel σ-field of the state space S. For a given set $\mathcal{A} \subset \mathcal{S}$, define

$$d(\mu, \nu) := \sup_{A \epsilon \mathcal{A}} |\mu(A) - \nu(A)| \qquad (\mu, \nu \epsilon \mathcal{P}(S)). \tag{3.1}$$

Consider the following *hypothesis* $(\mathbf{H}_1)$:

(1) $$(\mathcal{P}(S), d) \text{ is a complete metric space;} \tag{3.2}$$

(2) there exists a positive integer N such that for all $\gamma \,\epsilon\, \Gamma^N$, outside a set of Q^N-probability zero, one has

$$d(\mu_0 \tilde{\gamma}^{-\mathbf{1}},\ \nu_0 \tilde{\gamma}^{-1}) \leq d(\mu, \nu) \ \ (\mu, \nu \epsilon \mathcal{P}(S)); \tag{3.3}$$

(3) there exists $\delta > 0$ such that $\forall A \epsilon \mathcal{A}$, and with N as in (2), one has

$$P((\alpha_N ... \alpha_1)^{-1} A = S \text{ or } \varphi) \geq \delta. \tag{3.4}$$

Remark 3.1 For (3.4), we have assumed the set within parentheses to be measurable. If it is not, then assume that there exists $F_A \epsilon \mathcal{F}$, with $P(F_A) \geq \delta$, on which $(\alpha_N ... \alpha_1)^{-1} A = S$ or φ.

Theorem 3.1 *Assume the hypothesis* $(\mathbf{H}_1)$. *Then there exists a unique invariant probability* π *for the Markov process* $X_n := \alpha_n ... \alpha_1 X_0$, *where* X_0 *is independent of* $\{\alpha_n : n \geq 1\}$. *Also, one has*

$$d(T^{*n}\mu, \pi) \leq (1 - \delta)^{[n/N]} \qquad (\mu \epsilon \mathcal{P}(S)) \tag{3.5}$$

where $T^{*n}\mu$ *is the distribution of* X_n *when* X_0 *has distribution* μ, *and* $[n/N]$ *is the integer part of* n/N.

Proof. Let $A \epsilon \mathcal{A}$. Then (3.4) holds, which one may express as

$$Q^N(\{\gamma \epsilon \Gamma^N : \tilde{\gamma}^{-\mathbf{1}} A = S \text{ or } \varphi\}) \geq \delta. \tag{3.6}$$

Then, $\forall \mu,\ \nu \epsilon \mathcal{P}(S)$,

$$\left|(T^{*N}\mu)(A) - (T^{*N}\nu)(A)\right| = \left| \int_{\Gamma^N} (\mu(\tilde{\gamma}^{-\mathbf{1}} A) - \nu(\tilde{\gamma}^{-\mathbf{1}} A)) Q^N(d\gamma) \right|. \tag{3.7}$$

Denoting the set in curly brackets in (3.6) by Γ_1, one then has

$$\begin{aligned}
&\left|(T^{*N}\mu)(A) - (T^{*N}\nu)(A)\right| \\
&= \left| \int_{\Gamma_1} (\mu(\tilde{\gamma}^{-\mathbf{1}} A) - \nu(\tilde{\gamma}^{-\mathbf{1}} A)) Q^N(d\gamma) \right. \\
&\quad \left. + \int_{\Gamma^N \setminus \Gamma_1} (\mu(\tilde{\gamma}^{-\mathbf{1}} A) - \nu(\tilde{\gamma}^{-\mathbf{1}} A)) Q^N(d\gamma) \right| \\
&= \left| \int_{\Gamma^N \setminus \Gamma_1} (\mu(\tilde{\gamma}^{-\mathbf{1}} A) - \nu(\tilde{\gamma}^{-\mathbf{1}} A)) Q^N(d\gamma) \right|
\end{aligned} \tag{3.8}$$

since on Γ_1 the set $\tilde{\gamma}^{-\mathbf{1}} A$ is S or φ, so that $\mu(\tilde{\gamma}^{-\mathbf{1}} A) = 1 = \nu(\tilde{\gamma}^{-\mathbf{1}} A)$, or $\mu(\tilde{\gamma}^{-\mathbf{1}} A) = 0 = \nu(\tilde{\gamma}^{-\mathbf{1}} A)$. Hence, using (3.3) and (3.4)

$$\left|(T^{*N}\mu)(A) - (T^{*N}\nu)(A)\right| \leq (1 - \delta) d(\mu, \nu). \tag{3.9}$$

Thus

$$d(T^{*N}\mu, T^{*N}\nu) \leq (1 - \delta) d(\mu, \nu). \tag{3.10}$$

Since $(\mathcal{P}(S), d)$ is a complete metric space by assumption $(\mathbf{H}_1)(1)$, and T^{*N} is a uniformly strict contraction on $\mathcal{P}(S)$ by (3.10), there exists a unique *fixed point* π of T^{*N}, i.e., $T^{*N}\pi = \pi$, and

$$\begin{aligned} d(T^{*kN}\mu, \pi) &= d(T^{*N}(T^{*(k-1)N}\mu),\, T^{*N}\pi) \\ &\leq (1-\delta)d(T^{*(k-1)N}\mu, \pi) \leq ... \leq (1-\delta)^k d(\mu, \pi). \end{aligned} \tag{3.11}$$

Also,

$$\begin{aligned} d(T^{*}\pi, \pi) &= d(T^{*}T^{*kN}\pi, \pi) = d(T^{*kN}T^{*}\pi, \pi) \\ &\leq (1-\delta)^k d(T^{*}\pi, \pi) \to 0 \qquad \text{as } k \to \infty. \end{aligned}$$

Hence $d(T^*\pi, \pi) = 0$, which implies $T^*\pi = \pi$. Thus π is a fixed point of T^*. If π_1 is another fixed point of T^* then π_1 is a fixed point of T^{*N}. By the uniqueness of the fixed point of T^{*N}, $\pi_1 = \pi$. Finally, since $d(\mu, \nu) \leq 1$ one has, with $n = [n/N]N + r(r = 1, 2, ..., N-1)$,

$$d(T^{*n}\mu, \pi) = d(T^{*[n/N]N}T^{*r}\mu, T^{*[n/N]N}\pi) \leq (1-\delta)^{[n/N]}. \tag{3.12}$$

We now derive two corollaries of Theorem 3.1 applied to i.i.d. monotone maps. Corollary 3.1 extends a result of Dubins and Freedman (1966, (Thm. 5.10)), to more general state spaces in $\mathbb{R}$ and relaxes the requirement of continuity of α_n. The set of monotone maps may include both nondecreasing and nonincreasing ones.

Let S be an interval in $\mathbb{R}$. Denote by $d_K(\mu, \nu)$ the *Kolmogorov distance* on $\mathcal{P}(S)$: if F_μ and F_ν are the distribution functions of μ and ν, then

$$d_K(\mu, \nu) \equiv \sup_{x\epsilon S} |F_\mu(x) - F_\nu(x)| \equiv \sup_{x\epsilon \mathbb{R}} |F_\mu(x) - F_\nu(x)| \tag{3.13}$$

It should be noted that convergence in the distance d_K on $\mathcal{P}(S)$ implies weak convergence in $\mathcal{P}(S)$.

Corollary 3.1 *Let S be an interval or a closed subset of $\mathbb{R}$. Suppose $\alpha_n (n \geq 1)$ is a sequence of i.i.d. monotone maps on S satisfying the splitting condition $(\mathbf{H})$:*

$(\mathbf{H})$ *There exist $x_0 \epsilon S$, a positive integer N and a constant $\delta > 0$ such that*

$$\begin{aligned} P\left(\alpha_N \alpha_{N-1} ... \alpha_1 x \leq x_0\ \forall x \epsilon S\right) \geq \delta, \\ P\left(\alpha_N \alpha_{N-1} ... \alpha_1 x \geq x_0\ \forall x \epsilon S\right) \geq \delta. \end{aligned}$$

Then there exists a unique invariant probability π for the Markov process X_n generated by $\alpha_n (n \geq 1)$ and, irrespective of the initial distribution μ of X_0,

$$d_K(T^{*n}\mu, \pi) \leq (1-\delta)^{[n/N]}. \tag{3.14}$$

Proof. First let S be a closed set. To apply Theorem 3.1, let $\mathcal{A}$ be the class of all sets $A = (-\infty, y] \cap S$, $y \epsilon \mathbb{R}$. then $d = d_K$. Completeness of $(\mathcal{P}(S), d)$ may be established directly, but also follows from a more general result in Bhattacharya and Lee (1988, Correction (1997)). Hence the condition (1) of $(\mathbf{H}_1)$ holds.

To check condition (2) of $(\mathbf{H}_1)$ note that if γ is monotone nondecreasing and $A = (-\infty, y] \cap S$, then $\gamma^{-1}((-\infty, y] \cap S) = (-\infty, x] \cap S$ or

$(-\infty, x) \cap S$, where $x = \sup\{z : \gamma(z) \le y\}$. Thus $\left|\mu(\gamma^{-1}A) - \nu(\gamma^{-1}A)\right| = |\mu((-\infty, x] \cap S) - \nu((-\infty, x] \cap S)|$ or, $|\mu((-\infty, x) \cap S) - \nu(-\infty, x) \cap S)|$. In either case, $\left|\mu(\gamma^{-1}A) - \nu(\gamma^{-1}A)\right| \le d(\mu, \nu)$, since $\mu((-\infty, x - 1/n] \cap S \uparrow \mu((-\infty, x) \cap S)$ (and the same holds for ν). If γ is monotone nonincreasing, then $\gamma^{-1}A$ is of the form $[x, \infty) \cap S$ or $(x, \infty) \cap S$, where $x := \inf\{z : \gamma(z) \le y\}$. Again it is easily shown, $\left|\mu(\gamma^{-1}A) - \nu(\gamma^{-1}A)\right| \le d(\mu, \nu)$. Finally, (3.4) holds for all $A = (-\infty, y] \cap S$, by $(\mathbf{H})$.

If S is an interval which is not closed then there is a strictly increasing (continuous) homeomorphism h on S onto one of the following closed subsets of $\mathbb{R} : (-\infty, \infty)$, $(-\infty, a]$, $[a, \infty)$, depending on whether S is an open interval or semi-closed $(c, d]$, $[c, d)$. Since h preserves both the order and the topology of S, one may lift the state space to the closed set $h(S)$.

To state the next corollary, let S be a closed subset of $\mathbb{R}^k$. Define $\mathcal{A}$ to be the class of all sets of the form

$$A = \{y : \phi(y) \le x\},\ \phi \text{ continuous and monotone on } S \text{ into } \mathbb{R}^k,\ x\epsilon\mathbb{R}^k. \quad (3.15)$$

Also, we will mean by " $\le$ " the *partial order*: $x = (x_1, ..., x_k) \le y = (y_1, ..., y_k)$ iff $x_i \le y_i\ \forall i = 1, 2, ..., k$. In the following corollary we will interpret $(\mathbf{H})$ to hold with this partial order $\le$, and with "$x \ge y$" meaning $y \le x$.

Corollary 3.2 *Let S and $\mathcal{A}$ be as above. If $\alpha_n (n \ge 1)$ is a sequence of i.i.d. monotone maps which are continuous, and the splitting condition $(\mathbf{H})$ holds, then there exists a unique invariant probability π and*

$$d(T^{*n}\mu, \pi) \le (1 - \delta)^{[n/N]} \qquad \forall \mu\epsilon\mathcal{P}(S), \qquad n \ge 1 \qquad (3.16)$$

where $d(\mu, \nu) := \sup\{|\mu(A) - \nu(A)| : A\epsilon\mathcal{A}\}$.

Proof. As in the proof of Corollary 3.1, the completeness of $(\mathcal{P}(S), d)$ follows from Bhattacharya and Lee (1988, Correction (1997)), where the proof of completeness does not depend on whether the ϕ in (3.15) are only monotone nondecreasing, or simply monotone. Condition (2) in $(\mathbf{H}_1)$ is immediate. For if $A = \{y : \phi(y) \le x\}$, and γ is continuous and monotone, then $\gamma^{-1}A = \{y : (\phi \circ \gamma)y \le x\}\epsilon\mathcal{A}$ since $\phi \circ \gamma$ is monotone and continuous.

It remains to verify $(\mathbf{H}_1)$ (3). Let A in (3.15) be such that ϕ is monotone nondecreasing. If $\phi(x_0) \le x$, then, by the splitting condition $(\mathbf{H})$,

$$\begin{aligned}\delta &\le P(\alpha_N ... \alpha_1 z \le x_0 \forall z\epsilon S) \le P(\phi\alpha_N ... \alpha_1 z \le \phi(x_0) \forall z\epsilon S)\\ &\le P(\phi\alpha_N ... \alpha_1 z \le x \forall z\epsilon S) = P(\alpha_N ... \alpha_1 z\epsilon A \forall z\epsilon S)\\ &= P((\alpha_N ... \alpha_1)^{-1}A = S). \qquad (3.17)\end{aligned}$$

If x in the definition of A in (3.15) is such that $\phi(x_0) \not\le x$ (i.e., at least one coordinate of $\phi(x_0)$ is larger than the corresponding coordinate of x), then

$$\begin{aligned}\delta &\le P(\alpha_N ... \alpha_1 z \ge x_0 \forall z\epsilon S) \le P(\phi\alpha_N ... \alpha_1 z \ge \phi(x_0) \forall z\epsilon S)\\ &\le P(\phi\alpha_N ... \alpha_1 z \not\le x \forall z\epsilon S) \le P(\alpha_N ... \alpha_1 z\epsilon A^c \forall z\epsilon S)\\ &= P((\alpha_N ... \alpha_1)^{-1}A = \varphi). \qquad (3.18)\end{aligned}$$

Now let ϕ in the definition of A in (3.15) be monotone decreasing. If $\phi(x_0) \leq x$, then

$$\begin{aligned}\delta &\leq P(\alpha_N ... \alpha_1 z \geq x_0 \,\forall z \epsilon S) \\ &\leq P(\phi\alpha_N ... \alpha_1 z \leq \phi(x_0) \,\forall z \epsilon S) \leq P(\phi\alpha_N ... \alpha_1 z \leq x \,\forall z \epsilon S) \\ &= P(\alpha_N ... \alpha_1 z \epsilon A \,\forall z \epsilon S).\end{aligned} \tag{3.19}$$

If $\phi(x_0) \nleq x$, then

$$\begin{aligned}\delta &\leq P(\alpha_N ... \alpha_1 z \leq x_0 \,\forall z \epsilon S) \leq P(\phi\alpha_N ... \alpha_1 z \geq \phi(x_0) \,\forall z \epsilon S) \\ &\leq P(\phi\alpha_N ... \alpha_1 z \nleq x \,\forall z \epsilon S) = P(\alpha_N ... \alpha_1 z \epsilon A^c \,\forall z \epsilon S).\end{aligned} \tag{3.20}$$

Thus $(\mathbf{H}_1)$ (3) is verified for all $A \epsilon \mathcal{A}$.

Remark 3.2 Corollary 3.2 extends the main theorem in Bhattacharya and Lee (Correction, ibid, 1997) to i.i.d. monotone maps which may include both types – increasing as well as decreasing.

3.2 A link with the classical approach to stability

Recall that in the standard ("classical") approach, a Markov process on a state space $(S, \mathcal{S})$ is described by (a) an initial distribution π_0 and (b) the one-step transition probability $p(x, A)$ which satisfies (i) for each $x \epsilon S, p(x, \cdot)$ is a probability distribution on $\mathcal{S}$ and (ii) for any fixed $A \epsilon \mathcal{S}\, p, (\cdot, A)$ is $\mathcal{S}$-measurable. A fundamental result on the stability of a Markov process is stated below as another example of the application of Theorem 3.1.

Corollary 3.3 *Let $p(x, A)$ be a transition probability on a Borel subset S of a complete, separable metric space with $\mathcal{S}$ as the Borel sigmafield on S. Suppose there exists a nonzero measure λ on S and a positive integer m such that*

$$p^{(m)}(x, A) \geq \lambda(A) \qquad \forall x \epsilon S, \qquad A \epsilon \mathcal{S} \tag{3.21}$$

Then there exists a unique invariant probability π for $p(\cdot, \cdot)$ such that

$$\sup_{x,A} \left| p^{(n)}(x, A) - \pi(A) \right| \leq (1 - \delta)^{[n/m]}, \qquad \delta := \lambda(S). \tag{3.22}$$

For a proof of this result of Doeblin and for some applications see Doob (1953, p. 197), or Bhattacharya and Waymire (1990, pp. 180, 181, 198, 199). We now show that (3.22) is an almost immediate consequence of Theorem 3.1. For this express $p^{(m)}(x, A)$ as

$$p^{(m)}(x, A) = \lambda(A) + (p^{(m)}(x, A) - \lambda(A)) = \delta\lambda_\delta(A) + (1 - \delta)q_\delta(x, A) \tag{3.23}$$

where

$$\lambda_\delta(A) := \frac{\lambda(A)}{\delta}, \qquad q_\delta(x, A) = \frac{p^{(m)}(x, A) - \lambda(A)}{1 - \delta} \tag{3.24}$$

Let $\beta_n (n \geq 1)$ be an i.i.d. sequence of maps on S constructed as follows. For each n, with probability δ let $\beta_n \equiv Z_n$ where Z_n is a random variable with values in S and distribution λ_δ; and with probability $1-\delta$, let $\beta_n = \alpha_n$ where α_n is a sequence of i.i.d. random maps on S such that $P(\alpha_n x \epsilon A) = q_\delta(x, A)$ (see Kiefer, 1986, p. 8; or Bhattacharya and Waymire (1990, p. 228), for constructing α_n). Then Theorem 3.1 applies for the transition probability $p^{(m)}(x, A)$ (for $p(x, A)$), and with $\mathcal{A} = \mathcal{S}$, $N = 1$. Note that $P(\beta_1^{-1} A = S$ or $\phi) \geq P(\beta_1(\cdot) \equiv Z_1) = \delta$. Hence (3.4) holds. Since $\mathcal{A} = \mathcal{S}$ in this example, completeness of $(\mathcal{P}(S), d)$ and the condition (3.3) obviously hold.

As the corollaries indicate, the significance of Theorem 3.1 stems from the fact that it provides geometric rates of convergence in appropriate metrics for different classes of irreducible as well as nonirreducible Markov processes. The metric d depends on the structure of the process.

We now turn to the necessity of the splitting condition $(\mathbf{H})$ for the existence of a unique invariant probability. For $\alpha_n (n \geq 1)$ i.i.d. nondecreasing and continuous on $S = [a, b]$ it was proved by Dubins and Freedman (1966, Theorem 5.17), that, barring the case of a.s. all maps having a common fixed point, the splitting condition $(\mathbf{H})$ is necessary for the existence of a unique invariant probability. It turns out that if the maps are a.s. nonincreasing then splitting is not in general necessary for the existence of a unique invariant probability, even in the case $S = [a, b]$ and α_n continuous (see Bhattacharya and Majumdar, 1999, Sect. 4).

3.3 Lipschitz maps

Now, let S be a complete separable metric space, and write ρ for the metric. A map f from S into S is *Lipschitz* if for some finite $L > 0$,

$$\rho(f(x),\ f(y)) \leq L\rho(x, y)$$

for all $x,\ y \in S$. Let $\varGamma$ be a family of Lipschitz maps from S into S.

Let $\{\alpha_n : n \geq 1\}$ be an i.i.d. sequence of random Lipschitz maps from $\varGamma$. Denote by $L_j^k (k \geq j)$ the *random Lipschitz coefficient* of $\alpha_k .....\alpha_j$, i.e.,

$$L_j^k(\cdot) := \sup\{\rho(\alpha_k....\alpha_j x,\ \alpha_k...\alpha_j y)/\rho(x, y) : x \neq y\}. \tag{3.25}$$

As before, $X_n(x) = \alpha_n .....\alpha_1 x$, and let $Y_n(x) = \alpha_1 ....\alpha_n x$ denote the *backward iteration* $(n \geq 1;\ x \in S)$. Note that, for every $n \geq 1$, $X_n(x)$ and $Y_n(x)$ have the same distribution for each x, and we denote this by:

$$X_n(x) \stackrel{d}{=} Y_n(x).$$

Before deriving the general result of this subsection, namely, Theorem 3.3, we first state and prove the following simpler special case, which suffices for a number of applications.

Theorem 3.2 *Let (S, ρ) be compact metric, and assume*

$$-\infty \leq E\ \log\ L_1^r < 0 \textit{ for some } r \geq 1. \tag{3.26}$$

Then the Markov process X_n has a unique invariant probability and is stable in distribution.

Proof. First assume (3.26) with $r = 1$. A convenient distance metrizing the weak topology of $\mathcal{P}(S)$ is the *bounded Lipschitzian distance* d_{BL} (Dudley, 1989, p. 317):

$$d_{BL}(\mu, \nu) = \sup\left\{\left|\int f d\mu - \int f d\nu\right| : f\epsilon\mathcal{L}\right\}$$
$$\mathcal{L} := \{f : S \to \mathbb{R}, |f(x)| \leq 1 \forall x, \ |f(x) - f(y)| \leq \rho(x,y) \forall x, y\}. \quad (3.27)$$

By Prohorov's theorem (see, e.g., Billingsley, 1968, p. 37), the space $(\mathcal{P}(S),\ d_{BL})$ is a compact metric space. Since T^* is a continuous map on this latter space, due to Feller continuity of $p(x, dy)$, it has a fixed point π - an invariant probability. To prove uniqueness of π and stability, write $L_j = L_j^j$ (see (3.25) and note that for all $x, y \,\epsilon\, S$ one has

$$\begin{aligned} \rho(X_n(x),\ X_n(y)) &\equiv \rho(\alpha_n....\alpha_1 x,\ \alpha_n...\alpha_1 y) \\ &\leq L_n \rho(\alpha_{n-1}...\alpha_1 x,\ \alpha_{n-1}....\alpha_1 y) \\ &\quad .... \\ &\leq L_n L_{n-1}...L_1 \rho(x,y) \leq L_n L_{n-1}...L_1 M, \end{aligned} \quad (3.28)$$

where $M = \sup\{\rho(x,y) : x, y \,\epsilon\, S\}$. By the strong law of large numbers, the logarithm of the last term in (3.28) goes to $-\infty$ a.s. since, in view of (3.28) (with $r = 1$), $\frac{1}{n}\sum_{j=1}^{n} \log\, L_j \to E\, \log\, L_1 < 0$. Hence

$$\sup_{x,y\epsilon S} \rho(X_n(x), X_n(y)) \to 0 \ a.s.\ \ as\ n \to \infty. \quad (3.29)$$

Let X_0 be independent of $(\alpha_n)_{n\geq 1}$ and have distribution π. Then, arguing conditionally, (3.29) yields

$$\sup_{x\epsilon S} \rho(X_n(x),\ X_n(X_0)) \to 0 \ a.s.\ \ as\ n \to \infty. \quad (3.30)$$

Since $X_n(x)$ has distribution $p^{(n)}(x, dy)$ and $X_n(X_0)$ has distribution π one then has (see (3.27))

$$\begin{aligned} d_{BL}(p^{(n)}(x,\cdot), \pi) &= \sup\{|Ef(X_n(x)) - Ef(X_n(X_0))| : f\epsilon\mathcal{L}\} \\ &\leq E\rho(X_n(x),\ X_n(X_0))\} \to 0 \quad as\ n \to \infty. \end{aligned} \quad (3.31)$$

This completes the proof for the case $r = 1$.

Let now $r > 1$ (in (3.26)). The above argument applied to $p^{(r)}(x, dy)$, regarded as a one-step transition probability, shows that for every $x\epsilon S$, as $k \to \infty,\ p^{(kr)}(x,\cdot)$ converges in d_{BL}-distance (and, therefore, weakly) to a probability π which is invariant for $p^{(r)}(\cdot,\cdot)$. This implies that $T^{*kr}\mu \to \pi$ (weakly) for every $\mu\epsilon\mathcal{P}(S)$. In particular, for each $j = 1, 2, ..., r-1$, $p^{(kr+j)}(x,\cdot) \equiv T^{*kr} p^{(j)}(x,\cdot) \to \pi$ as $k \to \infty$. That is, for every $x\epsilon S$,

$$p^{(n)}(x,\cdot)\ converges\ weakly\ to\ \pi\ as\ n \to \infty \quad (3.32)$$

In turn, (3.32) implies $T^{*n}\mu \to \pi$ weakly for every $\mu\epsilon\mathcal{P}(S)$. Therefore, π is the unique invariant probability for $p(\cdot,\cdot)$ and stability holds.

Remark 3.3 The last argument above shows, in general, that if for some $r > 1$, the skeleton Markov process $\{X_{kr} : k = 0, 1, ...\}$ has a unique invariant probability π and is stable, then the entire Markov process $\{X_n : n = 0, 1, ...\}$ has the unique invariant probability π and is stable.

Remark 3.4 A direct application of Theorem 3.2 leads to a strengthening of the stability results in Green and Majumdar (1975, Theorems 5.1–5.2) who introduced (i.i.d.) random shocks in the Walras-Samuelson tatonnement.

We now turn to the main result of this subsection which extends Theorem 3.2 and which will find an important application in Section 4.

Theorem 3.3 *Let S be a complete separable metric space with metric ρ. Assume that (3.26) holds for some $r \geq 1$, and there exists a point $x_0 \in S$ such that*

$$E \log^+ \rho(\alpha_r....\alpha_1 x_0, x_0) < \infty \tag{3.33}$$

Then the Markov process $X_n(x) := \alpha_n....\alpha_1 x$ $(n \geq 1)$ is stable in distribution.

Proof. First consider the case $r = 1$ for simplicity. For this case we will show that the desired conclusion follows from the two assertions (a), (b) below:

(a) $\sup\{\rho(Y_n(x), Y_n(y)) : \rho(x, y) \leq M\} \to 0$ *in probability as* $n \to \infty$, *for every* $M > 0$.

(b) *The sequence of distributions of* $\rho(X_n(x_0), x_0)$, $n \geq 1$, *is relatively weakly compact.*

Assuming (a), (b), note that for every real valued Lipschitz $f \epsilon \mathcal{L}$ (see (3.27)) one has for all $M > 0$,

$$\sup_{\rho(x,y)\leq M} |Ef(X_n(x)) - Ef(X_n(y))| \leq \sup_{\rho(x,y)\leq M} E[\rho(X_n(x), X_n(y)) \wedge 2] \to 0 \tag{3.34}$$

as $n \to \infty$, by (a). Now,

$$\begin{aligned} &|Ef(X_{n+m}(x_0)) - Ef(X_n(x_0))| = |Ef(Y_{n+m}(x_0)) - Ef(Y_n(x_0))| \\ &\equiv |Ef(\alpha_1...\alpha_n\alpha_{n+1}...\alpha_{n+m}x_0) - Ef(\alpha_1...\alpha_n x_0)| \\ &\leq E[\rho(\alpha_1...\alpha_n\, \alpha_{n+1}...\alpha_{n+m}\, x_0,\, \alpha_1...\alpha_n\, x_0) \wedge 2]. \end{aligned} \tag{3.35}$$

For a given n and arbitrary $M > 0$, divide the sample space into the disjoint sets:

(i) $B_1 = \{\rho(\alpha_{n+1}...\alpha_{n+m}x_0, x_0) \geq M\}$

and

(ii) $B_2 = \{\rho(\alpha_{n+1}...\alpha_{n+m}x_0, x_0) < M\}$.

The last expectation in (3.35) is no more than $2P(B_1)$ on B_1. On B_2, denoting for the moment $X = \alpha_{n+1}...\alpha_{n+m}x_0$, one has

$\rho(\alpha_1...\alpha_n\, \alpha_{n+1}...\alpha_{n+m}x_0,\, \alpha_1...\alpha_n x_0) \equiv \rho(\alpha_1...\alpha_n X,\, \alpha_1...\alpha_n x_0)$

$\equiv \rho(Y_n(X), Y_n(x_0)) \leq \sup\{\rho(Y_n(x), Y_n(y)) : \rho(x, y) \leq M\} = Z$, say. The last inequality holds since, on B_2, $\rho(X, x_0) \leq M$. Now, for every $\delta > 0$, $E(Z \wedge 2) = E(Z \wedge 2 \cdot 1_{\{Z>\delta\}}) + E(Z \wedge 2 \cdot 1_{\{Z\leq\delta\}}) \leq 2P(Z > \delta) + \delta$. Hence,

$$\begin{aligned} |Ef(X_{n+m}(x_0)) - Ef(X_n(x_0))| &\leq 2P(\rho(\alpha_{n+1}...\alpha_{n+m}x_0, x_0) \geq M) \\ &+2\, P(\sup\{\rho(Y_n(x), Y_n(y)) : \rho(x, y) \leq M\} > \delta) + \delta. \end{aligned} \tag{3.36}$$

Fix $\varepsilon > 0$, and let $\delta = \varepsilon/3$. Choose $M = M_\varepsilon$ such that

$$P(\rho(\alpha_1...\alpha_m x_0,\, x_0) \geq M_\varepsilon) < \varepsilon/6 \qquad \text{for all } m = 1, 2, ... \tag{3.37}$$

This is possible by (b). Now (3.36), (3.37) and (a) imply that there exists n_ε such that for all $m = 1, 2, ...$, and $f \epsilon \mathcal{L}$,

$$|Ef(X_{n+m}(x_0)) - Ef(X_n(x_0))| < \varepsilon \qquad \text{for all } n \geq n_\varepsilon. \tag{3.38}$$

In other words, recalling the definition (3.27) of the bounded Lipschitzian distance $d_{BL}(\mu, \nu)$,

$$\sup_{m \geq 1} d_{BL}(p^{(n+m)}(x_0,\, dy),\, p^{(n)}(x_0,\, dy)) \to 0 \text{ as } n \to \infty. \tag{3.39}$$

This implies that $p^{(n)}(x_0,\, dy)$ is Cauchy in $(\mathcal{P}(S),\, d_{BL})$ and, therefore converges weakly to a probability measure $\pi(dy)$ as $n \to \infty$, by the *completeness* of $(\mathcal{P}(S),\, d_{BL})$ (Dudley, 1989, p. 317). Since $p(x,\, dy)$ has the *Feller property*, it follows that π is an invariant probability. In view of (a), $p^{(n)}(x,\, dy)$ converges weakly to π, for every $x \epsilon S$. This implies that π is the unique invariant probability.

It remains to prove (a), (b).

To prove (a), write L_n for the Lipschitz coefficient of α_n, i.e., $L_n = \sup\{\rho(\alpha_n x,\, \alpha_n y)/\rho(x, y) : x \neq y\}$, and as in (3.28),

$$\begin{aligned}\rho(Y_n(x),\, Y_n(y)) &= \rho(\alpha_1\alpha_2...\alpha_n x,\, \alpha_1\alpha_2...\alpha_n y)\\ &\leq L_1\rho(\alpha_2...\alpha_n x,\, \alpha_2...\alpha_n y) \leq ... \leq\, L_1L_2...L_n\rho(x, y)\\ &\leq L_1L_2...L_n M \text{ for all } (x, y) \text{ such that } \rho(x, y) \leq M.\end{aligned} \tag{3.40}$$

Take logarithms to get

$$\begin{aligned}&\sup_{\{(x,y):\rho(x,y)\leq M\}} \log\, \rho(Y_n(x),\, Y_n(y)) \leq \log\, L_1\\ &+ ... +\, \log\, L_n + \log\, M \to -\infty \text{ a.s. as } n \to \infty,\end{aligned} \tag{3.41}$$

by the strong law of large numbers and (3.26).

To prove (b), it is enough to prove that $\rho(\alpha_1...\alpha_n\, x_0,\, x_0)$ converges almost surely and, therefore in distribution, to some a.s. finite limit (Use the fact that $\rho(Y_n(x_0),\, x_0) = \rho(\alpha_1...\alpha_n\, x_0,\, x_0)$ and $\rho(X_n(x_0),\, x_0) = \rho(\alpha_n...\alpha_1 x_0,\, x_0)$ have the same distribution). For this use the triangle inequality to write

$$\begin{aligned}&|\rho(\alpha_1...\alpha_{n+m}x_0, x_0) - \rho(\alpha_1...\alpha_n x_0, x_0)| \leq \rho(\alpha_1...\alpha_{n+m}x_0, \alpha_1...\alpha_n x_0)\\ &\leq \sum_{j=1}^{m} \rho(\alpha_1...\alpha_{n+j}x_0, \alpha_1...\alpha_{n+j-1}x_0)\\ &\leq \sum_{j=1}^{m} L_1...L_{n+j-1}\rho(\alpha_{n+j}x_0, x_0).\end{aligned} \tag{3.42}$$

Let $c := -E \log L_1 (> 0)$. First assume c is finite and fix ε, $0 < \varepsilon < c$. It follows from (3.26) and the strong law of large numbers that there exists $n_1(\omega)$ such that, outside a P-null set N_1,

$$L_1 ... L_{n'} < e^{-n'(c-\varepsilon/2)} \text{ for all } n' \geq n_1(\cdot). \tag{3.43}$$

In view of (3.33),

$$\sum_{k=1}^{\infty} P\left(\frac{\log^+ \rho(\alpha_k x_0, x_0)}{\varepsilon/2} > k\right) = \sum_{k=1}^{\infty} P\left(\frac{\log^+ \rho(\alpha_1 x_0, x_0)}{\varepsilon/2} > k\right) \leq \frac{E \log^+ \rho(\alpha_1 x_0, x_0)}{\varepsilon/2} < \infty. \tag{3.44}$$

Therefore, by the Borel-Cantelli Lemma, there exists $n_2(\omega)$ such that, outside a P-null set N_2,

$$\rho(\alpha_k x_0, x_0) \leq e^{k\varepsilon/2} \text{ for all } k \geq n_2(\cdot). \tag{3.45}$$

Using (3.43) and (3.45) in (3.41) one gets

$$\begin{aligned} &\sup_{m \geq 1} |\rho(\alpha_1 ... \alpha_n \alpha_{n+1} ... \alpha_{n+m} x_0, x_0) - \rho(\alpha_1 ... \alpha_n x_0, x_0)| \\ &\leq \sum_{j=1}^{\infty} e^{-(n+j-1)(c-\varepsilon/2)} e^{(n+j)\varepsilon/2} \\ &\leq e^{c-\varepsilon/2} e^{-n(c-\varepsilon)} \sum_{j=1}^{\infty} e^{-j(c-\varepsilon)} \\ &\qquad \text{for all } n \geq n_0(\cdot) := \max\{n_1(\cdot),\, n_2(\cdot)\}, \end{aligned} \tag{3.46}$$

outside the P-null set $N = N_1 \cup N_2$. Thus the left side of (3.46) goes to zero almost surely as $n \to \infty$, showing that $\rho(\alpha_1 ... \alpha_n x_0, x_0)(n \geq 1)$ is Cauchy and, therefore, converges almost surely as $n \to \infty$. The proof of (b) is now complete if $-E \log L_1 < \infty$. If $-E \log L_1 = \infty$, then (3.43), (3.46) hold for any c>0. This concludes the proof of the theorem for $r = 1$ (in (3.26)). Now, the extension to the case $r > 1$ follows exactly as in the proof of Theorem 3.2 (see Remark 3.3).

Remark 3.5 The theorem holds if $P(L_1 = 0) > 0$. The proof (for the case $E \log L_1 = -\infty$) includes this case. But the proof may be made simpler in this case by noting that for all sufficiently large n, say $n \geq n(\omega)$, $X_n(x) = X_n(y) \ \ \forall\, x, y$.

Remark 3.6 The moment condition (3.33) under which Theorem 3.3 is derived here is nearly optimal, as seen from its specialization to the affine linear case in the next section. Under the additional moment assumption $EL_1^1 < \infty$, the stability result is proved in Diaconis and Freedman (1999) (for the case $r = 1$), and a speed of convergence is derived. The authors also provide a number of important applications as well as an extensive list of references to the earlier literature.

Remark 3.7 Random iterates of linear and distance-diminishing maps have been studied extensively in psychology: the monograph by Norman (1972) is an introduction to the literature with a comprehensive list of references.

4 Applications

4.1 Linear time series models

For motivating the main result of this section (which is proved by appealing to Theorem 3.3), let us consider briefly the linear autoregressive model of order one (the $AR(1)$ model). Let b be a real number and $\{\varepsilon_n : n \geq 1\}$ an i.i.d. sequence of real-valued random variables defined on some probability space. Given an initial random variable X_0 independent of $\{\varepsilon_n\}$ define recursively the sequence of random variables $\{X_n : n \geq 0\}$ as follows:

$$X_{n+1} = bX_n + \varepsilon_{n+1}. \tag{4.1}$$

Then X_n is a Markov process with initial distribution given by the distribution of X_0, and the transition probability (of going from x to some Borel set C of R in one step)

$$p(x, C) = Q(C - bx)$$

where Q is the common distribution of ε_n. To be sure, in terms of our earlier notation,

$$\alpha_n(x) = bx + \varepsilon_n. \tag{4.2}$$

Note that from (4.1) we get

$$X_n = b^n X_0 + b^{n-1}\varepsilon_1 + ... + b\varepsilon_{n-1} + \varepsilon_n \quad (n \geq 1) \tag{4.3}$$

The distribution of X_n is the same as that of

$$Y_n = b^n X_0 + \varepsilon_1 + ... + b^{n-1}\varepsilon_n \tag{4.4}$$

Assume that

$$|b| < 1 \tag{4.5}$$

and $|\varepsilon_n| \leq c$ with probability one for some constant c. Then it follows from (4.4) that

$$Y_n \to \sum_{n=0}^{\infty} b^n \varepsilon_{n+1} \qquad a.s. \tag{4.6}$$

regardless of X_0. Let π be the distribution of the random variable on the right side of (4.6). Then Y_n converges in distribution to π as $n \to \infty$. As $X_n \stackrel{d}{=} Y_n$, X_n converges in distribution to π. Hence π is the unique invariant distribution for the Markov process $\{X_n\}$.

The assumption that the random variable ε_1 is bounded can be relaxed; indeed, it suffices to assume that

$$\sum_{n=1}^{\infty} P(|\varepsilon_1| > c\delta^n) < \infty \qquad \text{for some } \delta < \frac{1}{|b|} \text{ and for some } c > 0. \tag{4.7}$$

Since the sum on the left equals $\sum_{n=1}^{\infty} P(|\varepsilon_{n+1}| > c\delta^n)$, one has, by the Borel-Cantelli Lemma,

$$P(|\varepsilon_{n+1}| \leq c\delta^n \text{ for all but finitely many } n) = 1$$

This implies that, with probability 1, $|b^n \varepsilon_{n+1}| \le c(|b|\,\delta)^n$ for all but finitely many n. Since $|b|\,\delta < 1$, the series on the right side of (4.6) converges and is the limit of Y_n.

For the sake of completeness, we state a more definitive characterization of the process (4.1).

Theorem 4.1 *(a) Assume that ε_1 is nondegenerate and*

$$E(\log|\varepsilon_1|)^+ < \infty \tag{4.8}$$

Then the condition

$$|b| < 1 \tag{4.9}$$

is necessary and sufficient for the existence of a unique invariant probability π for $\{X_n\}_{n=1}^{\infty}$. In addition, if both (4.8) and (4.9) hold, then X_n converges in distribution to π as $n \to \infty$ no matter what X_0 is.

(b) If

$$E(\log|\varepsilon_1|)^+ = \infty \tag{4.10}$$

and $|b| \ge e^{-1}$, then $\{X_n\}_{n=0}^{\infty}$ does not converge in distribution (irrespective of X_0) and does not have any invariant probability.

Proof. One may express X_n as

$$X_n = b^n X_0 + b^{n-1}\varepsilon_1 + b^{n-2}\varepsilon_2 + .. + b\varepsilon_{n-1} + \varepsilon_n = b^n X_0 + \sum_{j=1}^{n} b^{n-j}\varepsilon_j. \tag{4.11}$$

Now, recalling the notation $\underset{=}{d}$ for equality in distribution, note that for each $n \ge 1$,

$$\sum_{j=0}^{n-1} b^j \varepsilon_{j+1} \underset{=}{d} \sum_{j=1}^{n} b^{n-j}\varepsilon_j (n \ge 1). \tag{4.12}$$

To prove *sufficiency* in part (a), assume (4.8) and (4.9) hold. Then $b^n X_0 \to 0$ as $n \to \infty$. Therefore, we need to prove that

$$\sum_{j=0}^{n-1} b^j \varepsilon_{j+1} \ \mathit{converges\ in\ distribution\ as}\ n \to \infty. \tag{4.13}$$

Now choose and fix δ such that

$$1 < \delta < \frac{1}{|b|}. \tag{4.14}$$

Then $\delta\,|b| < 1$, and writing $Z = (\log^+ |\varepsilon_1|)/\log\,\delta$ one has (see the Lemma 4.1 below for the last step)

$$\begin{aligned}\sum_{j=0}^{\infty} P\left(\left|b^j\varepsilon_{j+1}\right| \geq (\delta\,|b|)^j\right) &= \sum_{j=0}^{\infty} P(|\varepsilon_{j+1}| \geq \delta^j)\\ &= \sum_{j=0}^{\infty} P(|\varepsilon_1| \geq \delta^j) = \sum_{j=0}^{\infty} P(\log|\varepsilon_1| \geq j\log\delta)\\ &= \sum_{j=1}^{\infty} P((\log|\varepsilon_1|)^+ \geq j\log\delta) + P(\log|\varepsilon_1| \geq 0)\\ &= \sum_{j=1}^{\infty} P(Z \geq j) + P(\log|\varepsilon_1| \geq 0)\\ &= E[Z] + P(\log|\varepsilon_1| \geq 0), \end{aligned} \tag{4.15}$$

where $[Z]$ is the integer part of Z. By (4.8), $E[Z] < \infty$. Hence, by the (*first*) Borel-Cantelli Lemma,

$$P(\left|b^j\varepsilon_{j+1}\right| < (\delta\,|b|)^j \quad \text{for all but finitely many } j) = 1. \tag{4.16}$$

Since $\delta\,|b| < 1$, this implies the existence of a (finite) random variable X such that

$$\sum_{j=0}^{n-1} b^j\varepsilon_{j+1} \to X \;\; a.s. \quad as\;\, n \to \infty. \tag{4.17}$$

Therefore, $\sum_{j=0}^{n-1} b^j\varepsilon_{j+1}$ converges in distribution to X. Therefore, by (4.12) (and the fact $b^nX_0 \to 0$), X_n converges in distribution to X. That is, if π denotes the distribution of the a.s. limit in (4.17), the distribution of X_n converges weakly to π, irrespective of X_0. Since the transition probability $p(x, dy)$ of the Markov process (which is the distribution of $bx+\varepsilon_1$) is a weakly continuous function of x (on $\mathbb{R}$ into $\mathcal{P}(\mathbb{R})$), it follows that π is invariant. π is obviously the *unique* invariant probability. For if π' is another invariant probability (different from π) then, letting X_0 have distribution π', the process $\{X_n\}_{n=0}^{\infty}$ is stationary with common distribution π' and, therefore, X_n trivially converges in distribution to π'. This is a contradiction, since we have shown that X_n converges in distribution to π, *irrespectively of* X_0.

To prove the *necessity* of (4.9), first let $|b| > 1$. Choose $\hat{\theta}$ such that $1 < \hat{\theta} < |b|$. Then

$$\begin{aligned}\sum_{j=0}^{\infty} P(\left|b^j\varepsilon_{j+1}\right| \geq \hat{\theta}^{\,j}) &= \sum_{j=0}^{\infty} P(|\varepsilon_{j+1}| \geq \left(\frac{\hat{\theta}}{|b|}\right)^j)\\ &= \sum_{j=0}^{\infty} P(|\varepsilon_1| \geq \left(\frac{\hat{\theta}}{|b|}\right)^j) = \infty, \end{aligned} \tag{4.18}$$

since $P(|\varepsilon_1| \geq \left(\frac{\hat{\theta}}{|b|}\right)^j) \uparrow P(|\varepsilon_1| > 0) > 0$ as $j \uparrow \infty$. But (4.18) implies, by the (*second*) Borel-Cantelli Lemma that

$$P(\left|b^j \varepsilon_{j+1}\right| \geq \hat{\theta}^j \quad \text{for infinitely many } j) = 1, \tag{4.19}$$

so that $\sum_{j=0}^{n-1} b^j \varepsilon_{j+1}$ diverges a.s. as $n \to \infty$. By a standard result on the series of independent random variables (see Loeve 1963, p. 251) it follows that $\sum_{j=0}^{n-1} b^j \varepsilon_{j+1}$ does not converge in distribution. By (4.12), therefore, $\sum_{j=1}^{n} b^{n-j} \varepsilon_j$ does not converge in distribution. If $X_0 = 0$, we see from (4.11) that $X_n \equiv \sum_{j=1}^{n} b^{n-j} \varepsilon_j$ does not converge in distribution.

Now, if $X_0 \neq 0$, then the same argument applies to the process $Y_n := X_n - X_0$, since $Y_{n+1} = bY_n + \eta_{n+1}$, with $\{\eta_n : n \geq 1\}$ i.i.d., $\eta_n := \varepsilon_n + (b-1)X_0$. Since $Y_0 = 0$, we see that Y_n does not converge in distribution. Therefore, $X_n \equiv Y_n + X_0$ does not converge in distribution.

To complete the proof of necessity in part (a), consider the case $|b| = 1$. Let $c > 0$ be such that $P(|\varepsilon_1| > c) > 0$. Then

$$\sum_{j=0}^{\infty} P(\left|b^j \varepsilon_{j+1}\right| > c) = \sum_{j=0}^{\infty} P(|\varepsilon_{j+1}| > c) = \sum_{j=0}^{\infty} P(|\varepsilon_1| > c) = \infty. \tag{4.20}$$

By the (second) Borel-Cantelli Lemma, $P(\sum_{j=0}^{n-1} b^j \varepsilon_{j+1}$ converges a.s.$) = 0$. Therefore $\sum_{j=0}^{n-1} b^j \varepsilon_{j+1}$ does not converge in distribution (Loeve, 1963, p. 217). Hence $\sum_{j=1}^{n} b^{n-j} \varepsilon_j$ does not converge in distribution. Since $b^n X_0 = X_0$ in case $b = +1$, and $b^n X_0 = \pm\, X_0$ in case $b = -1$, $X_n \equiv b^n X_0 + \sum_{j=1}^{n-1} b^{n-j} \varepsilon_j$ does not converge in distribution.

(b) Assume (4.10) and let $1 \geq |b| \geq e^{-1}$. Then

$$\begin{aligned}
&\sum_{j=0}^{\infty} P(\left|b^j \varepsilon_{j+1}\right| \geq |be|^j) = \sum_{j=0}^{\infty} P(|\varepsilon_{j+1}| \geq e^j)) = \sum_{j=0}^{\infty} P(|\varepsilon_1| \geq e^j) \\
&= \sum_{j=0}^{\infty} P(\log|\varepsilon_1| \geq j) = \sum_{j=1}^{\infty} P((\log|\varepsilon_1|)^+ \geq j) + P(\log|\varepsilon_1| \geq 0) \\
&= E[(\log|\varepsilon_1|)^+] + P(\log|\varepsilon_1| \geq 0) \\
&\geq E(\log|\varepsilon_1|)^+ - 1 + P(\log|\varepsilon_1| \geq 0) = \infty.
\end{aligned} \tag{4.21}$$

We have used Lemma 4.1 in the next to the last step.

Hence, by the (second) Borel-Cantelli Lemma,

$$P(\left|b^j \varepsilon_{j+1}\right| \geq |be|^j \quad \text{for infinitely many } j) = 1 \tag{4.22}$$

Hence $\sum_{j=1}^{n} b^{n-j} \varepsilon_j$ does not converge in distribution. Therefore, X_n does not converge in distribution, no matter what X_0 may be. It follows that $\{X_n\}_{n=0}^{\infty}$ does not have any invariant probability.

Remark 4.1 In proving that X_n does not have any invariant probability in case $|b| \geq 1$ we did not require the hypothesis $E(\log|\varepsilon_1|)^+ < \infty$. The hypothesis '$\varepsilon_1$ is nondegenerate' in the theorem is also not quite needed, except for the fact that if $P(\varepsilon_n = c) = 1$ and $b \neq 1$, then $X_n = b^n X_0 + c(b^n - 1)/(b-1)$, and $-c/(b-1) = z$, say is a fixed point, i.e., δ_z is an invariant probability. But even in this last case, if $|b| > 1$ then X_n does not converge if $X_0 \neq z$. If $b = +1$ and $\varepsilon_n = 0$ a.s., then x is a fixed point for every x, so that δ_x is an invariant probability for every $x \,\epsilon\, \mathbb{R}$. If $b = -1$ and $\varepsilon_n = c$ a.s., then $x = \frac{c}{2}$ is a fixed point and $\{x, -x+c\}$ is a periodic orbit for every $x \neq \frac{c}{2}$; it follows in this case that $\frac{1}{2}\delta_x + \frac{1}{2}\delta_{-x+c}$ are invariant probabilities for every $x \,\epsilon\, \mathbb{R}$. These are the only possible extremal invariant probabilities for the cases $|b| \geq 1$. In particular, if (4.8) holds (irrespective of whether ε_1 is degenerate or not), then (4.9) is necessary and sufficient for stability in distribution of $\{X_n\}$.

The following lemma has been used in the proof.

Lemma 4.1 *Let Z be a nonnegative random variable. Then*

$$\sum_{j=1}^{\infty} P(Z \geq j) = E[Z], \tag{4.23}$$

where $[Z]$ is the integer part of Z.

Proof.

$$\begin{aligned} E[Z] &= 0.P(0 \leq Z < 1) + 1.P(1 \leq Z < 2) + 2.P(2 \leq Z < 3) \\ &\quad +3.P(3 \leq Z < 4) + ... + j.P(j \leq Z < j+1) + ... \\ &= \{P(Z \geq 1) - P(Z \geq 2)\} + 2\{P(Z \geq 2) \\ &\quad -P(Z \geq 3)\} + 3\{P(Z \geq 3) \\ &\quad -P(Z \geq 4) + ... + j\{P(Z \geq j) - P(Z \geq j+1)\} + ... \\ &= P(Z \geq 1) + P(Z \geq 2) + P(Z \geq 3) + ... + P(Z \geq j) + ... \end{aligned} \tag{4.24}$$

A significant generalization of part (a) of Theorem 4.1 is the following Theorem 4.2. Although it holds with respect to other norms also, for specificity we take the *norm of a matrix A* to be defined by

$$\|A\| = \sup_{|x|=1} |Ax| \tag{4.25}$$

where $|x|$ denotes the Euclidean norm of the vector x.

Theorem 4.2 *Let $S = \mathbb{R}^k$, and $(A_i, B_i)(i \geq 1)$ be an i.i.d. sequence with A_i's random $(k \times k)$ matrices and B_i's random vectors (k-dimensional). If, for some $r \geq 1$,*

$$-\infty \leq E \log \|A_1....A_r\| < 0, \qquad E \log^+ |B_1| < \infty, \tag{4.26}$$

and

$$E \log^+ \|A_1\| < \infty, \tag{4.27}$$

then the Markov process defined recursively by $X_{n+1} = A_{n+1}X_n + B_{n+1} (n \geq 0)$ is stable in distribution.

Proof. To apply Theorem 3.3 with $\alpha_i(x) = A_i x + B_i$, and Euclidean distance for P, note that

$$\begin{aligned}\rho(\alpha_r ... \alpha_1 x,\, \alpha_r ... \alpha_1 y) &= \rho(A_r ... A_1 x,\, A_r ... A_1 y) \\ &= |A_r A_{r-1} ... A_1 (x-y)| \leq \|A_r ... A_1\| \cdot |x-y| \,. \quad (4.28)\end{aligned}$$

Thus the first condition in (4.26) implies (3.26) in this case. To verify (3.33) take $x_0 = 0$. If $r = 1$ then $\rho(\alpha_1(0),\, 0) = |B_1|$, and the second relation in (4.26) coincides with (3.33). In this case (4.27) follows from the first relation in (4.26): $E \log^+ \|A_1\|$ is finite because $E \log \|A_1\|$ exists and is negative. For $r > 1$,

$$\rho(\alpha_r ... \alpha_1 0, 0) = |A_r ... A_2 B_1 + A_r ... A_3 B_2 + ... + A_r B_{r-1} + B_r|$$

$$\leq \sum_{j=2}^{r} \|A_r\| ... \|A_j\| \cdot |B_{j-1}| + |B_r| \,,$$

$$\log^+ \rho(\alpha_r ... \alpha_1(0), 0) \leq \sum_{j=2}^{r} \log(\|A_r\| ... \|A_j\| \cdot |B_{j-1}| + 1)$$

$$+ \log(\,|B_r| + 1),$$

$$\log(\|A_r\| ... \|A_j\| \cdot |B_{j-1}| + 1) \leq \log(\|A_r\| + 1) + ... + \log(\|A_j\| + 1)$$

$$+ \log(|B_{j-1}| + 1). \quad (4.29)$$

Now $E \log(\|A_1\|+1) \leq E \log^+ \|A_1\|+1$, and $E \log(|B_1|+1) \leq E \log^+ |B_1|+1$ are finite by (4.26), (4.27). Hence (4.29) implies $E \log^+ \rho(\alpha_r ... \alpha_1(0), 0) < \infty$, i.e., the hypotheses of Theorem 3.3 are satisfied.

For the case $r = 1$, Theorem 4.2 is contained in a result of Brandt (1986). For $r > 1$, the above result was obtained by Berger (1992) using a different method.

In dimension $k = 2$, by an appropriate choice of a finite set of affine contractions $A_i x + B_i$ chosen with probabilities $p_i (i = 1, 2, ..., m)$ one is often able to approximately reproduce (and encode) a *fractal image* by the (empirical plot of the) invariant probability π of Theorem 4.2 (see Diaconis and Shashahani, 1986; Barnsley and Elton, 1988).

4.2 Non-linear autoregressive processes

There is a substantial and growing literature on non-linear time series models (see Tong, 1990; Granger and Terasvirta, 1993; Diks, 1999). A stability property of the general model

$$X_{n+1} = f(X_n, \varepsilon_{n+1})$$

(where ε_n is an i.i.d. sequence) can be derived from contraction assumptions [see Lasota and Mackey (1989) and Theorem 3.3 above; as an application of Corollary 3.3, we consider a simple first order non-linear autoregressive model (NLAR(1)].

Let $S = R$, and $f : R \to [a, b]$ be a (bounded) measurable function. Consider the Markov process

$$X_{n+1} = f(X_n) + \varepsilon_{n+1}$$

where (A.1) $\{\varepsilon_n\}$ is an i.i.d. sequence of random variables whose common distribution has a strictly positive continuous density $\hat{\phi}$ with respect to the Lebesgue measure.

Let X_0 be a random variable (independent of $\{\varepsilon_n\}$). The transition probability of the Markov process X_n has the density

$$p(x, y) = \hat{\phi}(y - f(x))$$

Note that if we define $\psi(y)$ as:

$$\psi(y) \equiv \min\{\hat{\phi}(y - z) : a \leq z \leq b\} > 0,$$

it follows that for all $x \epsilon R$

$$p(x, y) \equiv \hat{\phi}(y - f(x)) \geq \psi(y)$$

Since (3.21) holds with λ having density ψ, the Doeblin-stability result Corollary 3.3 applies directly.

4.3 An estimation problem

Consider a Markov chain X_n with a unique stationary distribution π. Some of the celebrated results on ergodicity and the strong law of large numbers hold for π-almost every initial condition. However, even with $[0, 1]$ as the state space, the invariant distribution π may be hard to compute explicitly when the laws of motion are allowed to be non-linear, and its support may be difficult to determine or may be a set of zero Lebesgue measure. Moreover, in many economic models, the initial condition may be historically given, and there may be little justification in assuming that it belongs to the support of π.

Consider then a random dynamical system with state space $[c, d]$ (without loss of generality for what follows choose $c > 0$). Assume Γ consists of a family of monotone maps from S to S, and the splitting condition (H) holds. The process starts with a given x. There is, by Corollary 3.1, a unique invariant distribution π of the random dynamical system, and (3.14) holds. Suppose we want to estimate the equilibrium mean $\int_S y\pi(dy)$ by *sample means* $\frac{1}{n}\sum\limits_{j=0}^{n-1} X_j$. We say that the estimator $\frac{1}{n}\sum\limits_{j=0}^{n-1} X_j$ is $\sqrt{n}$*-consistent* if

$$\frac{1}{n}\sum_{j=0}^{n-1} X_j = \int y\pi(dy) + O_p(n^{-1/2}) \tag{4.30}$$

where $O_p(n^{-1/2})$ is a random sequence ε_n such that $\left|\varepsilon_n/n^{-1/2}\right|$ is bounded in probability. Thus, if the estimator is $\sqrt{n}$-consistent, the fluctuations of the empirical (or sample-) mean around the equilibrium mean is $O_p(n^{-1/2})$. We shall outline the main steps in the verification of (4.30) in our context. For any bounded (Borel) measurable f on $[c, d]$, define the transition operator T as:

$$Tf(x) = \int_S f(y)p(x, dy)$$

By using the estimate (3.14), one can show that (see Bhattacharya, and Majumdar, 2001, pp. 217–219) if

$f(z) = z - \int y\pi(dy)$

then

$$\sup_x \sum_{n=m+1}^{\infty} |T^n f(x)| \le (d-c) \sum_{n=m+1}^{\infty} (1-\delta)^{[n/N]} \to 0 \qquad as\ m \to \infty.$$

Hence, $g = -\sum\limits_{n=0}^{\infty} T^n f$ [where T^0 is the identity operator I] is well-defined, and g and Tg are bounded functions. Also, $(T - I)g = -\sum\limits_{n=1}^{\infty} T^n f + \sum\limits_{n=0}^{\infty} T^n f \equiv f$. Hence,

$$\begin{aligned}\sum_{j=0}^{n-1} f(X_j) &= \sum_{j=0}^{n-1} (T-I)g(X_j) \\ &= \sum_{j=0}^{n-1} ((Tg)(X_j) - g(X_j)) \\ &= \sum_{j=1}^{n} [(Tg)(X_{j-1}) - g(X_j)] + g(X_n) - g(X_0).\end{aligned}$$

By the Markov property and the definition of Tg it follows that

$$E((Tg)(X_{j-1}) - g(X_j)\,|\mathcal{F}_{j-1}) = 0.$$

where$\mathcal{F}_r$ is the σ-field generated by $\{X_j : 0 \le j \le r\}$. Hence, $(Tg)(X_{j-1}) - g(X_j)(j \ge 1)$ is a martingale difference sequence, and are uncorrelated, so that

$$E\left[\sum_{j=1}^{n} (Tg(X_{j-1}) - g(X_j))\right]^2 = \sum_{j=1}^{n} E((Tg)(X_{j-1}) - g(X_j))^2 \qquad (4.31)$$

Given the boundedness of g and Tg, the right side is bounded by $n.K$ for some constant K. It follows that

$$\frac{1}{n} E\left(\sum_{j=0}^{n-1} f(X_j)\right)^2 \le \eta' \qquad for\, all\, n,$$

where η' is a constant that does not depend on X_0. Thus,

$$E\left(\frac{1}{n}\sum_{j=0}^{n-1} X_j - \int y\pi(dy)\right)^2 \leq \eta'/n$$

which implies,

$$\frac{1}{n}\sum_{j=0}^{n-1} X_j = \int y\pi(dy) + O_p(n^{-1/2})$$

For a unified exposition of $\sqrt{n}$-consistent estimation and other results see Athreya and Majumdar (2002).

4.4 Stability of invariant distributions in models of economic growth

Models of descriptive as well as optimal growth under uncertainty have lead to random dynamical systems that are stable in distribution. We look at a “canonical” example and show how Corollay 3.1 can be applied. A complete list of references to the earlier literature which owes much to the pioneering efforts of Brock and Mirman, is in Majumdar, Mitra and Nyarko (1989).

As a matter of notation, for any function h on S into S, we write $h^{(n)}$ for the nth iterate of h, i.e., $h^{(n)}$ is the n-fold composition $h \circ h \circ ... \circ h$.

Consider the case where $S = R_+$; and $\Gamma = \{F_1, F_2, ..., F_i, ..., F_N\}$ where the distinct laws of motion F_i satisfy:

F.1. *F_i is strictly increasing, continuous, and there is some $r_i > 0$ such that $F_i(x) > x$ on $(0, r_i)$ and $F_i(x) < x$ for $x > r_i$.*

Note that $F_i(r_i) = r_i$ for all $i = 1, ..., N$. Next, assume:

F.2. $r_i \neq r_j$ for $i \neq j$.

In other words, the unique positive fixed points r_i of distinct laws of motion are all distinct. We choose the indices $i = 1, 2, ..., N$ so that

$$r_1 < r_2 < .... < r_N$$

Let Prob $(\alpha_n = F_i) = p_i > 0 (1 \leq i \leq N)$.

Consider the Markov process $\{X_n(x)\}$ with the state space $(0, \infty)$. If $y \geq r_1$, then $F_i(y) \geq F_i(r_1) > r_1$ for $i = 2, ...N$, and $F_1(r_1) = r_1$, so that $X_n(x) \geq r_1$ for all $n \geq 0$ if $x \geq r_1$. Similarly, if $y \leq r_N$, then $F_i(y) \leq F_i(r_N) < r_N$ for $i = 1, ..., N-1$ and $F_N(r_N) = r_N$, so that $X_n(x) \leq r_N$ for all $n \geq 0$ if $x \leq r_N$. Hence, if the initial state x is in $[r_1, r_N]$, then the process $\{X_n(x) : n \geq 0\}$ remains in $[r_1, r_N]$ forever. We shall presently see that for a long run analysis we can consider $[r_1, r_N]$ as the effective state space.

We shall first indicate that on the state space $[r_1, r_N]$ the splitting condition (H) is satisfied. If $x \geq r_1$, $F_1(x) \leq x$, $F_1^{(2)}(x) \leq F_1(x)$ etc. The limit of this decreasing sequence $F_1^{(n)}(x)$ must be a fixed point of F_1, and therefore must be r_1. Similarly, if $x \leq r_N$, then $F_N^n(x)$ increases to r_N. In particular,

$$\lim_{n\to\infty} F_1^{(n)}(r_N) = r_1, \qquad \lim_{n\to\infty} F_N^{(n)}(r_1) = r_N.$$

Thus, there must be a positive integer n_0 such that

$$F_1^{(n_0)}(r_N) < F_N^{(n_0)}(r_1).$$

This means that if $z_0 \,\epsilon\, [F_1^{(n_0)}(r_N), F_N^{(n_0)}(r_1)]$, then

$$\begin{aligned} \Pr\text{ob}(X_{n_0}(x) \le z_0 \;\; \forall x \epsilon [r_1, r_N]) \\ \ge \Pr ob(\alpha_n = F_1 \, for \, 1 \le n \le n_0) = p_1^{n_0} > 0 \\ \Pr\text{ob}(X_{n_0}(x) \ge z_0 \;\; \forall x \epsilon [r_1, r_n]) \\ \ge \Pr ob(\alpha_n = F_N \, for \, 1 \le n \le n_0) = p_N^{n_0} > 0 \end{aligned}$$

Hence, considering $[r_1, r_N]$ as the state space, and using Theorem 3.1, there is a unique invariant probability π with the stability property holding for all initial $x \epsilon [r_1, r_N]$.

Now, fix the initial state $x \epsilon (0, r_1)$, and define $m(x) = \min_{i=1,\ldots,N} F_i(x)$.

One can verify that (i) m is continuous; (ii) m is strictly increasing; (iii) $m(r_1) = r_1$ and $m(x) > x$ for $x \epsilon (0, r_1)$, and $m(x) < x$ for $x > r_1$. Let $x \,\epsilon\, (0, r_1)$. Clearly $m^{(n)}(x)$ increases with n, and $m^{(n)}(x) \le r_1$. The limit of the sequence $m^{(n)}(x)$ must be a fixed point of m and is, therefore, r_1. Since $F_i(r_1) > r_1$ for $i = 2, ..., N$, there exists some $\varepsilon > 0$ such that $F_i(y) > r_1 (2 \le i \le N)$ for all $y \epsilon [r_1 - \varepsilon, r_1]$. Clearly there is some n_ε such that $m^{n_\varepsilon}(x) \ge r_1 - \varepsilon$. If $\tau_1 = \inf\{n \ge 1 : X_n(x) > r_1\}$ then it follows that for all $k \ge 1$

$$\Pr\text{ob}(\tau_1 > n_\varepsilon + k) \le p_1^k.$$

Since p_1^k goes to zero as $k \to \infty$, it follows that τ_1 is finite almost surely. Also, $X_{\tau_1}(x) \le r_N$, since for $y \le r_1$, (i) $F_i(y) < F_i(r_N)$ for all i and (ii) $F_i(r_N) < r_N$ for $i = 1, 2, ..., N-1$ and $F_N(r_N) = r_N$. (In a single period it is not possible to go from a state less than r_1 to one larger than r_N). By the strong Markov property, and our earlier result, $X_{\tau+m}(x)$ converges in distribution to π as $m \to \infty$ for all $x \epsilon (0, r_1)$. Similarly, one can check that as $n \to \infty$, $X_n(x)$ converges in distribution to π for all $x > r_N$.

Note that the result sketched above is derived without the restrictive "ordering" assumption F.2 in Bhattacharya and Majumdar (2001). The assumption that Γ is finite can be dispensed with if one has additional structures in the model.

4.5 Random iterations of quadratic maps

4.5.1 Qualitative analysis The quadratic maps $F_\theta(x) = \theta x(1-x)$, $x \epsilon [0, 1]$, with the parameter θ ranging over $[0, 4]$, constitute one of the most important families of dynamical systems. The apparent simplicity of this family of maps belies its richness and complexity. Few other families have provided as much insight into chaotic phenomena as it has (see, e.g., Collet and Eckmann, 1980; Davaney, 1989; Day, 1994; May, 1976; de Melo and Van Strien, 1993). This family of maps originally arose in population biology and, for θ in the chaotic range, these maps may even model certain forms of turbulence (Ruelle, 1989). In the context of economics, this entire family of maps has been shown to arise as optimal programs in dynamic optimization models (Majumdar and Mitra, 2000). Since physical systems are often stochastically excited (Eckmann and Ruelle, 1985), and since uncertainty is inherent in most economic systems, in many applications a randomly perturbed system is a more appropriate model than a deterministic map F_θ. This provides a motivation for our review in this section of the process (1.2) with $\alpha_n = F_{\varepsilon_n}(\cdot)$, where $\{\varepsilon_n : n \geq 1\}$ is an i.i.d. sequence with values in the parameter space $[0, 4]$. That is, we consider here

$$X_n = F_{\varepsilon_n} .... F_{\varepsilon_1} X_0 \ (n \geq 1), \tag{4.32}$$

with X_0 independent of $\{\varepsilon_n : n \geq 1\}$ and taking values in $[0, 1]$. Not unexpectedly, these randomly perturbed dynamical systems comprise a mathematically rich class of Markov processes.

In our investigation into the stability in distribution of the process X_n in (4.32), we take the state space to be $S = (0, 1)$. As $F_\theta(0) = 0$ for all $\theta \epsilon [0, 4]$, it follows that $\delta_{\{0\}}$ is an invariant probability of X_n if the state space is taken to be $[0, 1]$. Our main interest here is to seek coinditions for the existence of an invariant probability π other than $\delta_{\{0\}}$ such that $X_n(x)$ converges in distribution to π for every $x \epsilon (0, 1)$. The end point 1 needs to be excluded also to ensure that $X_n(x)$ does not leave the state space.

In the first specific study of the process (4.32), Bhattacharya and Rao (1993) considered a two-point distribution of ε_n. Let Q denote the distribution of ε_n, and let $Q(\{\theta_1\}) = \gamma$, $Q(\{\theta_2\}) = 1 - \gamma$ $(0 < \gamma < 1)$. If $1 < \theta_1 < \theta_2 \leq 2$, then, denoting by $p_\theta = 1 - \frac{1}{\theta}$ the non-zero fixed point of $F_\theta (1 < \theta \leq 4)$, the interval $[p_{\theta_1}, p_{\theta_2}]$ is invariant under both maps $F_{\theta_1}, F_{\theta_2} : F_{\theta_i}(x) \,\epsilon\, [p_{\theta_1}, p_{\theta_2}]$ if $x \,\epsilon\, [p_{\theta_1}, p_{\theta_2}]$. Now F_{θ_i} is monotone increasing on $I = [p_{\theta_1}, p_{\theta_2}] \subset (0, \frac{1}{2}]$. Therefore, using the fact that p_{θ_i} is an attractive fixed point of $F_{\theta_i}(i = 1, 2)$ it is simple to check that the splitting condition $(\mathbf{H})$ holds and Corollary 3.1 applies, when the state space is restricted to $[p_{\theta_1}, p_{\theta_2}]$. Hence there exists a unique invariant probability π for this process on I. Also, one can show that if $x \,\epsilon\, (0, 1) \backslash [p_{\theta_1}, p_{\theta_2}]$, then there exists an a.s. finite integer $n_x(\cdot)$ such that $X_n(x) \,\epsilon\, [p_{\theta_1}, p_{\theta_2}]$ for all $n \geq n_x(\cdot)$. It follows that π is the unique invariant probability for X_n on the state space $S = (0, 1)$ and X_n is stable in distribution on S. This technique of finding $\theta_1 < \theta_2$ such that (1) both $F_{\theta_i}(i = 1, 2)$ are invariant on a subinterval I of $(0, 1/2]$ (so that F_{θ_i} are monotone increasing on I) or of $[\frac{1}{2}, 1)$ (so that F_{θ_i} are monotone decreasing on I), and (2) the splitting condition $(\mathbf{H})$ holds, applies also to those pairs (θ_1, θ_2) satisfying $2 < \theta_1 \leq 3 < \theta_2 \leq 1 + \sqrt{5}$, $\theta_1 \epsilon [\frac{8}{\theta_2(4-\theta_2)}, \ \theta_2)$. These results of Bhattacharya and Rao (1993) on stability (applying Corollary 3.1) are extended in Bhattacharya and Majumdar (1999) to include certain pairs (θ_1, θ_2) satisfying $3 < \theta_1 < \theta_2 < 1 + \sqrt{5}$. In the latter case, F_{θ_i} have attractive two-period orbits $(i = 1, 2)$, and each map leaves an appropriate interval $I \subset [\frac{1}{2}, 1)$ invariant. Here X_n does not satisfy the splitting condition $(\mathbf{H})$ on I. However, on each of two disjoint subintervals I_1, I_2 of I, the skeleton process $\{X_{2n} : n = 0, 1, ...\}$ satisfies the hypothesis of Corollary 3.1. Hence X_{2n} has two distinct ergodic invariant probabilities π_1, π_2, say, with supports contained in I_1, I_2, respectively. Thus the process X_n on I moves cyclically between I_1 and I_2, yielding a unique invariant probability $\pi = \frac{1}{2}\pi_1 + \frac{1}{2}\pi_2$, and stability in distribution holds on the state space $S = (0,1)$ in the average sense: $\frac{1}{n}\sum_{m=1}^{n} p^{(m)}(x,dy)$ converges weakly to π as $n \to \infty$, for every $x\epsilon S = (0,1)$.

The following theorem extends the above results on stability in distribution to a much broader class of distributions, using the technique of splitting of Corollary 3.1.

Theorem 4.3 *(Bhattacharya and Waymire, 1999). Let Q be a probability measure on $(1,4)$, and let θ_1, θ_2 be the infimum and the supremum, respectively, of the support of Q, with $1 < \theta_1 < \theta_2 < 4$. Assume*

(a) F_{θ_i} has an attractive periodic orbit of period m_i, and let q_i be a point of this orbit $(i = 1,2)$;

(b) there is an interval I contained either in $(0, \frac{1}{2}]$ or $[\frac{1}{2}, 1)$ which is left invariant by maps $F_{\theta_i}^m$ for some common multiple m of $m_i (i = 1,2)$;

(c) $q_i \epsilon I$, and $F_{\theta_i}^{km} I \to \{q_i\}$ as $k \to \infty$ $(i = 1,2)$;

(d) for every $x\epsilon(0,1)$, $\Pr ob(\tau_i < \infty | X_0 = x) = 1$ where $\tau_i = \inf\{n \geq 1 : X_{nm}\epsilon I\}(i = 1,2)$.

Then the Markov process X_n has a unique invariant probability, on $S = (0,1)$, and it is stable in distribution on the average.

Proof. Consider the skeleton Markov process $\{X_{km} : k = 0,1,2,...\}$. By assumptions (a), (b), its state space may be restricted to I, on which it is defined by random iterations of the i.i.d. monotone maps $\beta_k := F_{\varepsilon_{km}} ... F_{\varepsilon_{(k-1)m+1}} (k = 1,2,...)$. By assumption (c), the splitting condition $(\mathbf{H})$ of Corollary 3.1 holds for the process $\{X_{km} : k = 0,1,2,...\}$, with any x_0 lying in the interior of the interval joining q_1 and q_2. Hence this skeleton process $\{X_{km} : k = 0,1,...\}$ has a unique invariant probability π on I, and it is stable in distribution on I. Finally, assumption (d) ensures that π is the unique invariant probability of the Markov process $\{X_n : n = 0,1,2,...\}$ on $S = (0,1)$ and that it is stable in distribution on the average (see Remark 3.3).

Using techniques similar to those for Lipschitz maps in Section 3.3 (Theorem 3.2), Carlsson (2002) has recently shown that if Q has a two-point support $\{\theta_1, \theta_2\}$ and $1 < \theta_1 < \theta_2 \leq 3$, then the Markov process $\{X_n : n = 0,1,...\}$ has a unique invariant probability on $S = (0,1)$ and that it is stable in distribution.

Judging by the rather severe restrictions imposed on Q above to ensure the existence of a unique invariant probability on $S = (0,1)$, one may expect nonuniqueness to occur in many instances. So far the only concrete examples of nonuniqueness are due to Athreya and Dai (2002), who take a Q with support $\{\theta_1, \theta_2\}$ satisfying $\frac{1}{\theta_1} + \frac{1}{\theta_2} = 1$, $F_{\theta_i} p_{\theta_i} = p_{\theta_i} (i = 1,2)$ and $F_{\theta_1} p_{\theta_2} = p_{\theta_1}$, $F_{\theta_2} p_{\theta_1} = p_{\theta_2}$, so that $\{p_{\theta_1}, p_{\theta_2}\}$ is left invariant by both maps $F_{\theta_1}, F_{\theta_2}$. It follows that there is an invariant probability π with support $\{p_{\theta_1}, p_{\theta_2}\}$. It is shown in Athreya and Dai (2002) that there is at least one more invariant probability on $S = (0,1)$ for every θ_2 in the range $(3, \theta_2^0]$ where $\theta_2^0 = 3.67...$is the solution of $x^3(4-x) = 16$.

In a somewhat different, but related, direction Athreya and Dai (2000) proved that for an invariant probability to exist on $S = (0,1)$ one must have

$$E \ \log \ \varepsilon_1 > 0. \tag{4.33}$$

In other words, if $E \log \varepsilon_1 \leq 0$ then X_n converges in distribution (and in probability) to 0. To prove this use the recursion $X_{n+1} = \varepsilon_{n+1} X_n(1 - X_n)$ to express $\log X_{n+1}$ as $\log X_{n+1} = \log \varepsilon_{n+1} + \log X_n + \log(1 - X_n)$. If there exists an invariant probability π then letting X_0 have distribution π, the process $\{X_n : n = 0, 1, ...\}$ is stationary. Assume, for simplicity, that $E \log X_0$ is finite. Then taking expectations of both sides in the last equation, and using $E \log X_{n+1} = E \log X_n$, in the last equation, one arrives at $E \log \varepsilon_1 \equiv E \log \varepsilon_{n+1} = -E \log(1 - X_n) \equiv -E \log(1 - X_0) > 0$, proving (4.33) under the assumption $E \log X_0$ is finite. The proof without this finiteness assumption may be given by a truncation argument. Athreya and Dai (2000) also show that (4.33) and the condition

$$E\,|\log(4 - \varepsilon_1)| < \infty \tag{4.34}$$

together are *sufficient* for the existence of at least one invariant probability on $S = (0, 1)$. Of course, (4.33) and (4.34) together do not imply the existence of a unique invariant probability, as demonstrated, e.g., by the counter example discussed earlier. The following result proviodes a broad criterion for stability when the distribution Q of ε_n has a density (component) w.r.t. Lebesgue measure on $(0, 4)$.

Theorem 4.4 *(Bhattacharya and Majumdar, 2002). Assume (4.33), (4.34) hold, and that Q has a nonzero absolutely continuous component whose density is bounded away from zero on some nondegenerate open interval in* $(1, 4)$. *Then the Markov process* $\{X_n : n = 0, 1, 2, ...\}$ *on* $S = (0, 1)$ *in (4.32) has a unique invariant probability* π *and* $\frac{1}{N} \sum_{n=1}^{N} p^{(n)}(x, dy)$ *converges to* π *in total variation distance as* $N \to \infty$.

The invariant probability π in Theorem 4.4 is absolutely continuous with respect to Lebesgue measure, at least if Q is absolutely continuous. For discrete Q, however, very little is known about the nature of the invariant distribution, even under stabilitiy.

For the case of a two-point support of Q, say $\{\theta_1, \theta_2\}$, $1 < \theta_1 < \theta_2 \leq 2$, one has a unique invariant distribution π with support $S(\pi)$ contained in $[p_{\theta_1}, p_{\theta_2}]$ and an interesting result due to Bhattacharya and Rao (1993) is the following:

if $-1/2 + \frac{1}{2}(\sqrt{17}) < \theta_1 < 2$ *and* $\theta_2 = 2$, $S(\pi)$ *is a Cantor set (closed, nowhere dense and does not have any isolated point) of Lebesgue measure zero.*

References

Athreya, K.B., Dai, J.J.: Random logistic maps-I. Journal of Theoretical Probability **13**, 595–608 (2000)

Athreya, K.B., Dai, J.J.: On the nonuniqueness of the invariant probability for i.i.d. random logistic maps. Annals of Probability **30**, 437–442 (2002)

Athreya, K., Majumdar, M.: Estimating the stationary distribution of a Markov chain. Economic Theory (forthcoming) (2002)

Barnsley, M., Elton, J.: A new class of Markov processes for image encoding. Advances in Applied Probability **20**, 14–32 (1988)

Berger, M.: Random affine iterated function systems: mixing and encoding. In: Liggett, T., Newman, C.M., Pitt, L. (eds.) Diffusion processes and related problems in analysis, Vol. II. Progress in probability, vol. 27, pp. 315–346. Boston: Birkhäuser 1992

Bhattacharya, R.N., Lee, O.: Asymptotics of a class of Markov processes which are not in general irreducible. Annals of Probability **16**, 1333–1347 (1988). Correction, ibid **25**, 1541–1543 (1997)

Bhattacharya, R.N., Rao , B.V.: Random iterations of two quadratic maps. In: Cambanis, S., Ghosh, J.K., Karandikar, R.L., Sen, P.K. (eds.) Stochastic processes: A Festschrift in Honour of Gopinath Kallianpur, pp. 13–21. Berlin Heidelberg New York: Springer 1993

Bhattacharya, R.N., Waymire, E.C.: An approach to the existence of unique invariant probabilities for Markov processes. In: Colloquium for limit theorems in probability and statistics. Budapest: Janos Bolyai Soc. 1999

Bhattacharya, R., Majumdar, M.: On a theorem of Dubins and Freedman. Journal of Theoretical Probability **12**, 1067–1087 (1999)

Bhattacharya, R.N., Majumdar, M.: On a class of stable random dynamical systems: theory and applications. Journal of Economic Theory **96**, 208–229 (2001)

Bhattacharya, R.N., Majumdar, M.: Stability in distribution of randomly perturbed quadratic maps as Markov processes. CAE Working Paper 02–03, Cornell University (2002)

Bhattacharya, R.N., Waymire, E.C.: Stochastic processes with applications. New York: Wiley 1990

Billingsley, P.: Convergence of probability measures. New York: Wiley 1968

Brandt, A.: The stochastic equation $Y_{n+1} = A_n Y_n + B_n$ with stationary coefficients. Advances in Applied Probability **18**, 211–220 (1986)

Carlsson, N.: Contractive condition for iterated function system. Journal of Theoretical Probability (2002)

Collet, P., Eckmann, J.P.: Iterated random maps on the interval as dynamical systems. Boston: Birkhäuser 1980

Day, R.: Complex economic dynamics. Cambridge, MA.: M.I.T. Press 1994

de Melo, W., Van Strien, S.: One dimensional dynamics. Berlin Heidelberg New York: Springer 1993

Devaney, R.L.: An introduction to chaotic dynamical systems, 2nd edn. New York: Addison-Wesley 1989

Diaconis, P., Freedman, D.: Iterated random functions. SIAM Review **41**, 45–76 (1999)

Diaconis, P., Shashahani, M.: Products of random matrices and computer image generation. Contemp. Mathematics **50**, 173–182 (1986)

Diks, C.: Nonlinear time series analysis. Singapore: World Scientific 1999

Doob, J.L.: Stochastic processes. New York: Wiley 1953

Dubins, L.E., Freedman, D.: Invariant probability measures for certain Markov processes. Annals of Mathematical Statistics **37**, 837–848 (1966)

Dudley, R.M.: Real analysis and probability. Pacific Grove, CA: Wadsworth and Brooks-Cole 1989

Eckman, J-P., Ruelle, D.: Ergodic theory of chaos and strange attractors. Review of Modern Physics **57**, 617–656 (1985)

Granger, C.W.J., Terasvirta, T.: Modelling nonlinear economic relationships. Oxford: Oxford University Press 1993

Green, J., Majumdar, M.: The nature of stochastic equilibria. Econometrica **43**, 647–660 (1975)

Kiefer, Y.: Ergodic theory of random transformations. Boston: Birkhäuser 1986

Lasota, A., Mackey, M.C.: Stochastic perturbation of dynamical systems: the weak convergence of measures. Journal of Mathematical Analysis and Applications **138**, 232–248 (1989)

Loeve, M.: Probability theory, 3 edn. Princeton, NJ: Van Nostrand 1963

Majumdar, M., Mitra, T., Nyarko, Y.: Dynamic optimization under uncertainty; non-convex feasible set. In: Feiwel, G.R. (ed.) Joan Robinson and modern economic theory, pp. 545–590. London: MacMillan 1989

Majumdar, M., Mitra, T.: Robust ergodic chaos in discounted dynamic optimization models. Economic Theory **4**, 677–688 (1994)

Majumdar, M., Mitra, T.: Robust ergodic chaos in discounted dynamic optimization models. In: Majumdar, M., et al. (eds.) Optimization and chaos, pp. 240–257. Berlin Heidelberg New York: Springer 2000

May, R.M.: Simple mathematical models with very complicated dynamics. Nature **261**, 459–467 (1976)

Norman, M.F.: Markov processes and learning models. New York: Academic Press 1972

Ruelle, D.: Chaotic evolution and strange attractors. Cambridge: Cambridge University Press 1989

Tong, H.: Nonlinear time series: a dynamical systems approach. Oxford: Oxford University Press 1990