ANNALES
POLONICI MATHEMATICI
105.3 (2012)

A criterion of asymptotic stability for Markov–Feller e-chains on Polish spaces

by Dawid Czapla (Katowice)

Abstract. Stettner [Bull. Polish Acad. Sci. Math. 42 (1994)] considered the asymptotic stability of Markov–Feller chains, provided the sequence of transition probabilities of the chain converges to an invariant probability measure in the weak sense and converges uniformly with respect to the initial state variable on compact sets. We extend those results to the setting of Polish spaces and relax the original assumptions. Finally, we present a class of Markov–Feller chains with a linear state space model which satisfy the assumptions of our main theorem.

1. Introduction. Let X be a Polish space. We consider a homogeneous X-valued Markov chain with transition semigroup $(P^n)_{n\in\mathbb{N}}$, where $P^1 = P$ is the transition kernel associated with the chain. A finite Borel measure π on X is said to be *invariant* for the chain if

$$\pi(A) = \int_X P(y,A)\,\pi(dy)$$

for each Borel set $A \subseteq X$.

We are concerned with sufficient conditions under which the Markov chain admits an invariant probability measure π and $(P^n(x,\cdot))_{n\in\mathbb{N}}$ converges to π both weakly and uniformly with respect to x on compact subsets of X, i.e. for each bounded, continuous function $f : X \to \mathbb{R}$ the convergence

$$\int_X f(y)\,P^n(x,dy) \to \int_X f(y)\,\pi(dy)$$

is uniform with respect to x on any compact subset of X. It is easily seen that, in particular, the latter condition implies the asymptotic stability of the chain, i.e. for each probability Borel measure μ on X and any bounded continuous function $f : X \to \mathbb{R}$ we have

$$\int_X f(y)\,\mu P^n(dy) \to \int_X f(y)\,\pi(dy),$$

2010 *Mathematics Subject Classification*: Primary 60J05; Secondary 37A30.
Key words and phrases: e-chains, asymptotic stability, invariant measure.

DOI: 10.4064/ap105-3-5

where $\mu P^n(A) = \int_X P^n(x,A)\,\mu(dx)$. This condition also guarantees the uniqueness of the invariant probability measure.

Motivated by the results of Stettner [9], we relax the assumptions of [9, Theorem 4.3] and give a new proof of that theorem based upon ideas found in [1, Theorem 2.1] due to Bessaih, Kapica and Szarek.

Recently, a lot of attention in the literature has been paid to various criteria for the existence of invariant measures and the asymptotic stability of Markov chains (initially, in the setting of locally compact and separable metric spaces and later, Polish spaces). Possible properties of a Markov chain (or equivalently, a Markov operator), which is typically assumed to be Feller, are usually related to the notion of *nonextensibility* (with respect to an appropriate norm in the space of measures), *semi-concentration* ([11, 7, 5, 12, 4]), *e-property*, *tightness* or *concentration at a point* ([9, 10, 1]). In the present paper we are mainly concerned with the last three properties of Markov chains.

The asymptotic stability problem consists, in general, in two issues: firstly, the existence of an invariant probability measure and, secondly, a kind of stability that provides the uniqueness of the invariant measure. Assuming the Markov chain is Feller and enjoys the tightness property, it does admit an invariant measure. Indeed, under these assumptions the invariant measure may be obtained as a weak limit of the Cesàro averages of $(P^n(x,\cdot))_{n\in\mathbb{N}}$ (see Proposition 5.3 of the Appendix or [11]). Therefore, we shall focus on the stability properties of Markov chains.

The outline of the paper is as follows. The first section deals with basic notions, notation and facts that we shall often refer to in this work. Among other things we formulate some sufficient conditions for tightness of a collection of measures.

We state our main result (which generalizes [9, Theorem 4.3]) in Section 2. At the end of that section, we derive a corollary from [10] and the main result. Then we specialise to the locally compact and separable case.

In Section 4, we give an example illustrating applications of the criterion for asymptotic stability formulated in Section 3. Appealing to results of Meyn and Tweedie [8], we consider a class of Markov chains with a linear state space model of $\mathbb{R}^d$ given by the recursive formula

$$\phi_{n+1} = F\phi_n + G\eta_{n+1} \quad (n \in \mathbb{N}),$$

where $(\eta_n)_{n\in\mathbb{N}}$ is a sequence of mutually independent and identically distributed random variables and F, G are real matrices. Assuming that the spectral radius of F is less than 1, it turns out that the model satisfies conditions (A1) and (A2) formulated in Section 3. Condition (A3) can be easily verified when the common distribution of variables η_n is given.

For the convenience of the reader and self-containedness, in the Appendix we give detailed proofs of some facts of importance from our perspective which are probably well-known to experts.

2. Preliminaries. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and let X be a Polish space (that is, a separable and complete metric space endowed with a metric ρ). Denote by $\mathcal{B}_X$ the σ-field of Borel subsets of X. We need the following notation:

- $\mathcal{M}(X)$ = the family of all finite, nonnegative and countably additive Borel measures on X;
- $\mathcal{M}_1(X)$ = the subset of $\mathcal{M}(X)$ consisting of probability measures;
- $\mathcal{M}_1^A(X)$ = the subset of $\mathcal{M}(X)$ consisting of probability measures supported on a given set $A \in \mathcal{B}_X$;
- $B(X)$ = the Banach space of all bounded, Borel, real-valued functions defined on X, equipped with the supremum norm;
- $C(X)$ = the subspace of $B(X)$ consisting of bounded continuous functions;
- $C_{bs}(X)$ = the subspace of $C(X)$ consisting of functions with bounded supports;
- $B(z, \delta) = \{x \in X : \rho(x, z) < \delta\}$ ($z \in X$ and $\delta > 0$).

Consider a time-homogeneous, X-valued Markov chain $(\xi_n)_{n\in\mathbb{N}_0}$ defined on Ω, where $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. Let P be the transition operator associated with that chain, that is, $P = P^1 : X \times \mathcal{B}_X \to [0, 1]$ determines the transition probability at the first step:

$$P(x, A) = \mathbb{P}(\xi_1 \in A \,|\, \xi_0 = x) = \mathbb{P}_x(\xi_1 \in A) \quad (x \in X \quad \text{and } A \in \mathcal{B}_X)$$

and satisfies the Chapman–Kolmogorov equations:

$$P^{n+m}(x, A) = \int_X P^m(y, A)\, P^n(x, dy) \quad (n, m \in \mathbb{N},\ x \in X \text{ and } A \in \mathcal{B}_X).$$

We associate with each transition probability P^n ($n \in \mathbb{N}$) a Markov operator $(\cdot)P^n$ acting on $\mathcal{M}(X)$ in the following manner:

$$\mu P^n(A) = \int_X P^n(x, A)\, \mu(dx) \quad (A \in \mathcal{B}_X \text{ and } \mu \in \mathcal{M}(X)).$$

The action of the dual operator $P^n(\cdot)$ on bounded Borel functions on X is defined by

$$P^n f(x) = \int_X f(y) P^n(x, dy) \quad (x \in X \text{ and } f \in B(X)).$$

We often write $\langle f, \mu \rangle$ instead of $\int_X f\, d\mu$.

Now, we list several definitions and facts that are essential for many proofs in the paper.

Definition 2.1. The chain $(\xi_n)_{n\in\mathbb{N}_0}$ is said to be:

(i) a *Feller chain* whenever $Pf \in C(X)$ for each $f \in C(X)$;
(ii) an *e-chain* if for each $f \in C_{bs}(X)$ the family $\{P^n f : n \in \mathbb{N}\}$ is equicontinuous at all points of X, that is,

$$(\forall x \in X)(\forall \varepsilon > 0)(\exists \delta > 0)(\forall n \in \mathbb{N})(\forall y \in B(x,\delta))(|P^n f(x) - P^n f(y)| < \varepsilon).$$

There are several definitions of e-process in the literature which are not equivalent in general. In this paper we adopt the definition from [10].

Definition 2.2. We say that a measure $\pi \in \mathcal{M}(X)$ is *invariant* for the chain $(\xi_n)_{n\in\mathbb{N}_0}$ (or the operator P) if

$$\pi(A) = \pi P(A) \quad \text{for each } A \in \mathcal{B}_X.$$

Definition 2.3. We say that a sequence $(\mu_n)_{n\in\mathbb{N}}$ of measures in $\mathcal{M}(X)$ *converges weakly* to a measure $\mu \in \mathcal{M}(X)$ whenever

$$\lim_{n\to\infty} \int_X f\, d\mu_n = \int_X f\, d\mu \quad \text{for each } f \in C(X).$$

Definition 2.4. The chain $(\xi_n)_{n\in\mathbb{N}_0}$ (or the operator P) is called *asymptotically stable* if it has an invariant probability measure π such that for each $\mu \in \mathcal{M}_1(X)$ the sequence $(\mu P^n)_{n\in\mathbb{N}}$ converges weakly to π.

In case a Markov process admits a unique invariant measure one may expect that it is also asymptotically stable.

Remark 2.1. If the chain $(\xi_n)_{n\in\mathbb{N}_0}$ is asymptotically stable then its invariant probability measure is unique.

Indeed, suppose π and π' are invariant probability measures for $(\xi_n)_{n\in\mathbb{N}_0}$. Then $\pi' P^n \xrightarrow{w} \pi$. Hence, for each $f \in C(X)$,

$$\Big|\int_X f(x)\, \pi'(dx) - \int_X f(x)\, \pi(dx)\Big| = \Big|\int_X f(x)\, \pi' P^n(dx) - \int_X f(x)\, \pi(dx)\Big| \to 0$$

as $n \to \infty$. Consequently, $\int_X f(x)\, \pi'(dx) = \int_X f(x)\, \pi(dx)$, and thus $\pi' = \pi$.

Since the ambient space X is Polish, the topology of weak convergence of measures in $\mathcal{M}_1(X)$ is metrisable by the Fortet–Mourier norm (see Section 1.5 of [6]). Hence one can equivalently define the asymptotic stability in terms of convergence with respect to that norm.

Now, we turn to the notion of tightness of measures, which plays a crucial role in further considerations.

Definition 2.5. A measure $\mu \in \mathcal{M}(X)$ is said to be *tight* if for each $\varepsilon > 0$ there exists a compact set $K \subseteq X$ such that $\mu(X \setminus K) < \varepsilon$.

A family of measures $\mathcal{R} \subset \mathcal{M}(X)$ is *tight* if for each $\varepsilon > 0$ there exists a compact set $K \subseteq X$ such that $\mu(X \setminus K) < \varepsilon$ for all $\mu \in \mathcal{R}$.

Ulam's Theorem ([2, Theorem 1.4], [8, Theorem 1.17]) asserts that each finite Borel measure on a Polish space is tight.

We shall prove two useful conditions equivalent to tightness of a collection of Borel measures. To do that, we need to introduce one more piece of notation: for each $\delta > 0$ set

$$\mathcal{C}(X,\delta)=\Big\{A \subset X\colon A\subset\bigcup_{i=1}^{N}\operatorname{cl} B(x_i,\delta) \text{ for some } N\in\mathbb{N} \text{ and } x_1,\dots,x_N\in X\Big\}.$$

We note that the following fact remains valid if we drop the assumption of separability of X.

Proposition 2.1 ([11, Lemma 3.2]). *A collection $\mathcal{A} \subseteq \mathcal{M}_1(X)$ of measures is tight if and only if for each $\delta > 0$ there exists a set $A \in \mathcal{C}(X,\delta)$ such that*

$$\nu(A) > 1-\delta \quad \text{for each } \nu \in \mathcal{A}.$$

Proof. It is enough to prove that the condition given above is sufficient for the tightness of $\mathcal{A}$. Let $\varepsilon > 0$. By assumption, there is a sequence $(A_m)_{m\in\mathbb{N}}$ of sets such that

$$A_m \in \mathcal{C}(X,\varepsilon/2^m) \quad \text{and} \quad \nu(A_m) > 1-\varepsilon/2^m \quad (m\in\mathbb{N},\, \nu\in\mathcal{A}).$$

Plainly, for each $m \in \mathbb{N}$, we have $\operatorname{cl} A_m \in \mathcal{C}(X,\varepsilon/2^m)$. Set $K = \bigcap_{m=1}^{\infty} \operatorname{cl} A_m$. Since for every $m \in \mathbb{N}$ the set K has an $\varepsilon/2^m$-net, it is compact by [8, Theorem 5.4]. Let $\nu \in \mathcal{A}$. We infer that

$$\nu(X\setminus K) = \nu\Big(\bigcup_{m=1}^{\infty}(X\setminus \operatorname{cl} A_m)\Big) \le \sum_{m=1}^{\infty}\nu(X\setminus \operatorname{cl} A_m) = \sum_{m=1}^{\infty}(1-\nu(\operatorname{cl} A_m))$$
$$\le \sum_{m=1}^{\infty}\frac{\varepsilon}{2^m} = \varepsilon. \ \blacksquare$$

We define a *norm-like function* to be a Borel function $V\colon X \to [0,\infty)$ for which there exists an increasing sequence $(C_n)_{n\in\mathbb{N}}$ of compact subsets of X such that

$$\lim_{n\to\infty}\inf_{x\in X\setminus C_n} V(x) = \infty. \tag{2.1}$$

We can now state the following result ([8, p. 531]).

Proposition 2.2. *Let $\mathcal{A} \subseteq \mathcal{M}_1(X)$. If there exists a norm-like function $V\colon X \to [0,\infty)$ with*

$$\sup_{\nu\in\mathcal{A}}\int_X V(x)\,\nu(dx) < \infty, \tag{2.2}$$

then the family $\mathcal{A}$ is tight. If additionally X is separable and locally compact, then the converse also holds.

Proof. Let $V\colon X\to[0,\infty)$ be a norm-like function satisfying (2.2) and let $(C_n)_{n\in\mathbb{N}}$ be an increasing sequence of compact sets enjoying (2.1). We note that

$$\begin{aligned}(2.3)\qquad &\sup_{\nu\in\mathcal{A}}\int_X V(x)\,\nu(dx)\ge\sup_{\nu\in\mathcal{A}}\int_{X\setminus C_n}V(x)\,\nu(dx)\\ &\ge\sup_{\nu\in\mathcal{A}}\Big(\inf_{x\in X\setminus C_n}V(x)\nu(X\setminus C_n)\Big)=\inf_{x\in X\setminus C_n}V(x)\sup_{\nu\in\mathcal{A}}\nu(X\setminus C_n)\quad(n\in\mathbb{N}).\end{aligned}$$

The sequence $(\sup_{\nu\in\mathcal{A}}\nu(X\setminus C_n))_{n\in\mathbb{N}}$ is bounded and decreasing (as $C_n\subset C_{n+1}$, $n\in\mathbb{N}$), so it converges to a non-negative number, say p. Letting $n\to\infty$ in (2.3), we obtain

$$\limsup_{n\to\infty}\Big[\inf_{x\in X\setminus C_n}V(x)\sup_{\nu\in\mathcal{A}}\nu(X\setminus C_n)\Big]\le\sup_{\nu\in\mathcal{A}}\int_X V(x)\,\nu(dx),$$

thus

$$p=\lim_{n\to\infty}\sup_{\nu\in\mathcal{A}}\nu(X\setminus C_n)=0,$$

because if $p>0$, then by (2.1) the right-hand side of the last inequality would be infinite, contrary to (2.2).

Consequently, taking any $\varepsilon>0$ we may find $N\in\mathbb{N}$ with $\sup_{\nu\in\mathcal{A}}\nu(X\setminus C_N)<\varepsilon$. This proves the tightness of $\mathcal{A}$.

Conversely, suppose that X is locally compact and separable and the family $\mathcal{A}$ is tight. Then there is an increasing sequence $(K_n)_{n\in\mathbb{N}}$ of compact subsets of X such that

$$\sup_{\nu\in\mathcal{A}}\nu(X\setminus K_n)<1/2^n\quad(n\in\mathbb{N}).$$

The assumptions about X yield a sequence $(U_n)_{n\in\mathbb{N}}$ of open and relatively compact subsets such that

$$K_n\subset U_n\quad(n\in\mathbb{N})\quad\text{and}\quad X=\bigcup_{n=1}^{\infty}U_n.$$

Set $C_n=\operatorname{cl}U_n$ $(n\in\mathbb{N})$. Plainly,

$$\nu(X\setminus C_n)\le\nu(X\setminus K_n)<1/2^n\quad(\nu\in\mathcal{A},\,n\in\mathbb{N}).$$

Define $V\colon X\to[0,\infty)$ as follows:

$$V(x)=\sum_{n=1}^{\infty}1_{X\setminus C_n}(x)\quad(x\in X).$$

This is well-defined, since each $x\in X$ is in some C_n, thus $x\in C_m$ for $m\ge n$, and consequently $V(x)\le n<\infty$. Moreover, V is a norm-like function: if $n\in\mathbb{N}$ and $x\in X\setminus C_n$, then $x\in X\setminus C_m$ for each $m\le n$, so $V(x)\ge n$, and hence

$$\inf_{x \in X \setminus C_n} V(x) \geq n \quad (n \in \mathbb{N}).$$

Furthermore, V satisfies (2.2) because for $\nu \in \mathcal{A}$ we have

$$\int_X V(x)\,\nu(dx) = \sum_{n=1}^{\infty} \nu(X \setminus C_n) \leq \sum_{n=1}^{\infty} \frac{1}{2^n} = 1. \blacksquare$$

3. A criterion of asymptotic stability. Throughout this section we assume that the chain $(\xi_n)_{n \in \mathbb{N}_0}$ enjoys the Feller property. Let us introduce the following conditions:

(A1) The process $(\xi_n)_{n \in \mathbb{N}_0}$ is an e-chain.
(A2) The family of measures $\{P^n(x, \cdot) \colon n \in \mathbb{N}\}$ is tight for each $x \in X$.
(A3) There exists $z \in X$ with the following property: for any compact set $C \subset X$ and each $\delta > 0$ there is an $N \in \mathbb{N}$ such that $P^N(x, B(z, \delta)) > 0$ for all $x \in C$.
(A2)′ The family of measures $\{P^n(x, \cdot) : x \in C,\ n \in \mathbb{N}\}$ is tight for any compact set $C \subset X$.
(A3)′ There exists $z \in X$ with the following property: for any compact set $C \subset X$ and for each $\delta > 0$ there is an $N \in \mathbb{N}$ such that for some $\alpha > 0$, each measure $\mu \in \mathcal{M}_1^C(X)$ satisfies $\mu P^N(B(z, \delta)) \geq \alpha$.

We need the following lemma [9, Lemma 4.4]:

LEMMA 3.1. *Conditions* (A3) *and* (A3)′ *are equivalent.*

Proof. Plainly, it is sufficient to prove that (A3) implies (A3)′. Assume that $z \in X$ satisfies (A3). Let C be a compact subset of X and fix $\delta > 0$. There is a natural number N such that

$$P^N(x, B(z, \delta/2)) > 0 \quad (x \in C). \tag{3.1}$$

Set $U = B(z, \delta)$. By the Urysohn Lemma, there is a continuous function $f \colon X \to [0, 1]$ such that

$$f(\operatorname{cl} B(z, \delta/2)) = \{1\} \quad \text{and} \quad f(X \setminus B(z, \delta)) = \{0\}.$$

Notice it is enough to show that

$$\alpha = \inf_{x \in C} P^N(x, U) > 0,$$

because for each $\mu \in \mathcal{M}_1^C(X)$ we then have

$$\mu P^N(U) \geq \int_C P^N(x, U)\mu(dx) \geq \alpha \mu(C) = \alpha.$$

We argue by contradiction. Suppose that $\inf_{x\in C} P^N(x,U)=0$. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence of points of C such that

$$\lim_{n\to\infty} P^N(x_n,U)=0.$$

Since C is compact, we can assume that $x_n\to x_0\in C$. Observe that

$$P^N f(x_0)=0.$$

Indeed, for all $n\in\mathbb{N}$ we have

$$P^N(x_n,U)=\int_X 1_U(y)\,P^N(x_n,dy)\ge\int_X f(y)\,P^N(x_n,dy)=P^N f(x_n),$$

and since the chain has the Feller property and the function f is nonnegative we infer that

$$0\le P^N f(x_0)=\lim_{n\to\infty} P^N f(x_n)\le\limsup_{n\to\infty} P^N(x_n,U)=0.$$

Therefore,

$$\begin{aligned}P^N(x_0,B(z,\delta/2))&=\int_X 1_{B(z,\delta/2)}(y)\,P^N(x_0,dy)\le\int_X f(y)\,P^N(x_0,dy)\\&=P^N f(x_0)=0,\end{aligned}$$

which contradicts (3.1). ∎

Armed with this information, we are ready to prove a theorem which is a crucial ingredient in the proof of the main result (the idea is based on the proof of [1, Theorem 2.1]).

THEOREM 3.1. *Suppose that conditions* (A1)–(A3) *are satisfied. Then, for each $f\in C(X)$ and any $x_1,x_2\in X$,*

$$\lim_{n\to\infty}|P^n f(x_1)-P^n f(x_2)|=0. \tag{3.2}$$

Proof. Firstly, it suffices to show that (3.2) is satisfied for each $f\in C_{bs}(X)$. Indeed, take $f\in C(X)$, $f\ne 0$, $x_1,x_2\in X$ and $\varepsilon>0$. By (A2) we find a compact set $C\subset X$ such that

$$P^n(x_i,X\setminus C)\le\frac{\varepsilon}{3\|f\|}\quad(n\in\mathbb{N},\ i=1,2).$$

Let $B\subset X$ be an open ball containing C. The Urysohn Lemma yields a continuous function $\varphi\colon X\to[0,1]$ with $\varphi|_C=1$ and $\varphi|_{X\setminus B}=0$. Set $f_0=f\varphi$. Then $f_0\in C_{bs}(X)$ and $\|f-f_0\|=\|f\|\,\|1-\varphi\|\le\|f\|$. Now, assuming that (3.2) holds for each function in $C_{bs}(X)$, we may choose $n_0\in\mathbb{N}$ such that

$$|P^n f_0(x_1)-P^n f_0(x_2)|\le\varepsilon/3\quad(n\ge n_0).$$

Hence, for each $n \ge n_0$ we have

$$
\begin{aligned}
|P^n f(x_1) - P^n f(x_2)| &\le |P^n f(x_1) - P^n f_0(x_1)| + |P^n f_0(x_1) - P^n f_0(x_2)| \\
&\quad + |P^n f_0(x_2) - P^n f(x_2)| \\
&\le \int_{X\setminus C} |(f - f_0)(y)|\, P^n(x_1, dy) + \varepsilon/3 \\
&\quad + \int_{X\setminus C} |(f - f_0)(y)|\, P^n(x_2, dy) \\
&\le \|f\| P^n(x_1, X \setminus C) + \varepsilon/3 + \|f\| P^n(x_2, X \setminus C) \\
&\le \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon.
\end{aligned}
$$

Secondly, we prove that (3.2) holds for each function in $C_{bs}(X)$. Fix $f \in C_{bs}(X)$, $f \ne 0$, $x_1, x_2 \in X$ and $\varepsilon \in (0, 1/2)$. By (A1), there is $\delta > 0$ such that

$$
|P^n f(x) - P^n f(y)| < \varepsilon/2 \quad (x, y \in B(z, \delta),\ n \in \mathbb{N}). \tag{3.3}
$$

Making use of (A2), we choose a compact set $K \subset X$ such that

$$
P^n(x_i, K) > 1 - \frac{\varepsilon^2}{4\|f\|} \quad (n \in \mathbb{N},\ i = 1, 2). \tag{3.4}
$$

There is no loss of generality in assuming that $x_1, x_2 \in K$. By the Feller property, each sequence of measures in $\{P(x, \cdot) : x \in K\}$ clearly admits a weakly convergent subsequence, whence by the Prokhorov Theorem ([2, Appendix III, Theorems 6, 7]) this family is tight. Therefore, there exists a compact set $L \subset X$ containing K such that

$$
\inf_{x \in K} P(x, L) \ge 3/4.
$$

Consequently, for each $\mu \in \mathcal{M}_1^K(X)$ we have

$$
\mu P(L) \ge \int_K P(x, L)\, \mu(dx) \ge 3/4 > 1/2. \tag{3.5}
$$

Applying condition (A3)$'$ to the set L and the number δ, we find a natural number N and a constant $\alpha \in (0, 1)$ such that

$$
\nu P^N(B(z, \delta)) \ge \alpha \quad (\nu \in \mathcal{M}_1^L(X)). \tag{3.6}
$$

Set

$$
\gamma = \alpha\varepsilon/2 \quad \text{and} \quad k = \min\{l \in \mathbb{N} : 4(1 - \gamma)^l \|f\| < \varepsilon\}.
$$

Clearly, $\gamma \in (0, 1/4)$. Let $m = N + 2$. We shall prove inductively that for each $l \in \{1, \ldots, k\}$ there exist $\nu_1^i, \ldots, \nu_l^i \in \mathcal{M}_1^{B(z,\delta)}(X)$ and $\mu_l^i \in \mathcal{M}_1(X)$ $(i = 1, 2)$ such that

$$P^{N+(l-1)m}(x_i,\cdot) = \gamma\,\nu_1^i P^{(l-1)m} + \gamma(1-\gamma)\,\nu_2^i P^{(l-2)m} + \cdots + \gamma(1-\gamma)^{l-2}\,\nu_{l-1}^i P^m + \gamma(1-\gamma)^{l-1}\,\nu_l^i + (1-\gamma)^l\,\mu_l^i. \tag{3.7}$$

Let $l=1$. Define

$$\nu_1^i(A) = \frac{P^N(x_i, A\cap B(z,\delta))}{P^N(x_i, B(z,\delta))} \qquad (A\in\mathcal{B}_X,\ i=1,2),$$
$$\mu_1^i(A) = \frac{1}{1-\gamma}(P^N(x_i,A) - \gamma\nu_1^i(A)) \quad (A\in\mathcal{B}_X,\ i=1,2).$$

Certainly, $\nu_1^i \in \mathcal{M}_1^{B(z,\delta)}(X)$ $(i=1,2)$. By (3.6) we infer that

$$P^N(x_i, B(z,\delta)) = \delta_{x_i}P^N(B(z,\delta)) \ge \alpha = 2\gamma/\varepsilon > \gamma \quad (i=1,2),$$

whence, by the definition of ν_1^i,

$$P^N(x_i,A) \ge P^N(x_i, A\cap B(z,\delta)) = P^N(x_i,B(z,\delta))\nu_1^i(A) \ge \gamma\,\nu_1^i(A)$$

for each $A\in\mathcal{B}_X$ $(i=1,2)$. The last inequality guarantees that μ_1^i is non-negative, so $\mu_1^i\in\mathcal{M}_1(X)$ $(i=1,2)$. It remains to note that the above defined measures satisfy (3.7):

$$\gamma\,\nu_1^i + (1-\gamma)\,\mu_1^i = \gamma\,\nu_1^i + P^N(x_i,\cdot) - \gamma\,\nu_1^i = P^N(x_i,\cdot) \quad (i=1,2).$$

Let now $l\in\{1,\dots,k-1\}$ and assume inductively that there exist measures $\nu_1^i,\dots,\nu_l^i \in \mathcal{M}_1^{B(z,\delta)}(X)$ and $\mu_l^i\in\mathcal{M}_1(X)$ $(i=1,2)$ such that (3.7) holds. By the definition of k we infer that

$$4(1-\gamma)^l\|f\| \ge \varepsilon. \tag{3.8}$$

We now prove

$$\mu_l^i P(K) > \varepsilon \quad (i=1,2). \tag{3.9}$$

Indeed, it follows from (3.4) and the inductive assumption that

$$\begin{aligned}
1-\frac{\varepsilon^2}{4\|f\|} &< P^{1+N+(l-1)m}(x_i,K) = P^{N+(l-1)m}P(x_i,K)\\
&= (\delta_{x_i}P^{N+(l-1)m})P(K) = (\gamma\,\nu_1^i P^{(l-1)m} + \gamma(1-\gamma)\,\nu_2^i P^{(l-2)m} + \cdots\\
&\qquad + \gamma(1-\gamma)^{l-2}\,\nu_{l-1}^i P^m + \gamma(1-\gamma)^{l-1}\,\nu_l^i + (1-\gamma)^l\,\mu_l^i)P(K)\\
&\le \gamma + \gamma(1-\gamma) + \cdots + \gamma(1-\gamma)^{l-1} + (1-\gamma)^l\,(\mu_l^iP)(K)\\
&= \gamma + \gamma\sum_{j=1}^{l-1}(1-\gamma)^j + (1-\gamma)^l(\mu_l^iP)(K)\\
&= \gamma + \gamma(1-\gamma)\frac{1-(1-\gamma)^{l-1}}{1-(1-\gamma)} + (1-\gamma)^l(\mu_l^iP)(K)\\
&= \gamma + (1-\gamma) - (1-\gamma)^l + (1-\gamma)^l(\mu_l^iP)(K)\\
&= 1-(1-\gamma)^l + (1-\gamma)^l(\mu_l^iP)(K) \quad (i=1,2),
\end{aligned}$$

whence by (3.8),

$$\mu_l^i P(K) > \frac{(1-\gamma)^l - \frac{\varepsilon^2}{4\|f\|}}{(1-\gamma)^l} = 1 - \frac{\varepsilon^2}{4\|f\|(1-\gamma)^l} > 1 - \frac{\varepsilon^2}{\varepsilon} > \varepsilon$$

for each $i = 1, 2$, proving (3.9).

Define

$$\widetilde{\mu}_l^i(A) = \frac{\mu_l^i P(A \cap K)}{\mu_l^i P(K)} \quad (A \in \mathcal{B}_X,\ i = 1, 2).$$

Clearly, $\widetilde{\mu}_l^i \in \mathcal{M}_1^K(X)$ $(i = 1, 2)$. The definition together with (3.9) yield

(3.10)
$$\mu_l^i P(A) \ge \mu_l^i P(A \cap K) = \mu_l^i P(K)\, \widetilde{\mu}_l^i(A) \ge \varepsilon \widetilde{\mu}_l^i(A) \quad (A \in \mathcal{B}_X,\ i = 1, 2).$$

The measures $\widetilde{\mu}_l^i$ $(i = 1, 2)$ satisfy (3.5) as they belong to $\mathcal{M}_1^K(X)$. Therefore,

$$\widetilde{\mu}_l^i P(L) > 1/2 \quad (i = 1, 2),$$

whence by (3.10),

(3.11) $$\mu_l^i P^2(L) = (\mu_l^i P)P(L) \ge \varepsilon \widetilde{\mu}_l^i P(L) > \varepsilon/2 \quad (i = 1, 2).$$

Define

$$\overline{\mu}_l^i(A) = \frac{\mu_l^i P^2(A \cap L)}{\mu_l^i P^2(L)} \quad (A \in \mathcal{B}_X,\ i = 1, 2).$$

Then $\overline{\mu}_l^i \in \mathcal{M}_1^L(X)$ $(i = 1, 2)$, so by (3.6) we get

(3.12) $$\overline{\mu}_l^i P^N(B(z, \delta)) \ge \alpha \quad (i = 1, 2).$$

Appealing to the definition of $\overline{\mu}_l^i$ and (3.11) we infer that

(3.13) $$\begin{aligned} \mu_l^i P^2(A) \ge \mu_l^i P^2(A \cap L) = \mu_l^i &= P^2(L)\overline{\mu}_l^i(A) \\ &\ge (\varepsilon/2)\overline{\mu}_l^i(A) \quad (A \in \mathcal{B}_X,\ i = 1, 2). \end{aligned}$$

Now, applying both (3.13) and (3.12) we observe that

(3.14) $$\begin{aligned} \mu_l^i P^{N+2}(B(z, \delta)) = (\mu_l^i P^2)P^N(B(z, \delta)) &\ge (\varepsilon/2)\overline{\mu}_l^i P^N(B(z, \delta)) \\ &\ge \varepsilon\alpha/2 = \gamma \quad (i = 1, 2). \end{aligned}$$

We are now ready to construct the relevant measures. Set

$$\nu_{l+1}^i(A) = \frac{\mu_l^i P^{N+2}(A \cap B(z, \delta))}{\mu_l^i P^{N+2}(B(z, \delta))} \qquad (A \in \mathcal{B}_X,\ i = 1, 2),$$

$$\mu_{l+1}^i(A) = \frac{1}{1-\gamma}(\mu_l^i P^{N+2}(A) - \gamma \nu_{l+1}^i(A)) \quad (A \in \mathcal{B}_X,\ i = 1, 2).$$

Certainly, $\nu_{l+1}^i \in \mathcal{M}_1^{B(z,\delta)}(X)$ $(i = 1, 2)$. Taking into account (3.14) together with the definition of ν_{l+1}^i we find that

$$\mu_l^i P^{N+2}(A) \ge \mu_l^i P^{N+2}(A \cap B(z, \delta)) = \mu_l^i P^{N+2}(B(z, \delta)) \cdot \nu_{l+1}^i(A) \ge \gamma \nu_{l+1}^i(A)$$

for $A \in \mathcal{B}_X$ and $i = 1, 2$. Therefore, μ^i_{l+1} is non-negative and belongs to $\mathcal{M}_1(X)$. Lastly, we prove that the measures defined above satisfy decomposition (3.7):

$$\begin{aligned} P^{N+lm}(x_i,\cdot) &= \delta_{x_i} P^{N+(l-1)m} P^m \\ &= (\gamma\,\nu^i_1 P^{(l-1)m} + \gamma(1-\gamma)\,\nu^i_2 P^{(l-2)m} \\ &\quad + \cdots + \gamma(1-\gamma)^{l-2}\,\nu^i_{l-1} P^m + \gamma(1-\gamma)^{l-1}\,\nu^i_l + (1-\gamma)^l\,\mu^i_l) P^m \\ &= \gamma\,\nu^i_1 P^{lm} + \gamma(1-\gamma)\,\nu^i_2 P^{(l-1)m} + \cdots + \gamma(1-\gamma)^{l-2}\,\nu^i_{l-1} P^{2m} \\ &\quad + \gamma(1-\gamma)^{l-1}\,\nu^i_l P^m + (1-\gamma)^l\,\mu^i_l P^m \qquad (i = 1, 2), \end{aligned}$$

and by the definition of μ^i_{l+1} the last summand above is of the form

$$\begin{aligned} (1-\gamma)^l \mu^i_l P^m = (1-\gamma)^l \mu^i_l P^{N+2} &= (1-\gamma)^l [(1-\gamma)\mu^i_{l+1} + \gamma\nu^i_{l+1}] \\ &= \gamma(1-\gamma)^l \nu^i_{l+1} + (1-\gamma)^{l+1} \mu^i_{l+1} \qquad (i = 1, 2). \end{aligned}$$

This proves the inductive step.

Set $n_0 = N + (k-1)m$ and fix $n \ge n_0$. By (3.7) we obtain

$$\begin{aligned} (3.15)\qquad & |P^n f(x_1) - P^n f(x_2)| = |P^{n_0} P^{n-n_0} f(x_1) - P^{n_0} P^{n-n_0} f(x_2)| \\ &= |\langle P^{n-n_0} f, P^{n_0}(x_1,\cdot) - P^{n_0}(x_2,\cdot)\rangle| \\ &\le \gamma |\langle P^{n-n_0} f, (\nu^1_1 - \nu^2_1) P^{(k-1)m}\rangle| + \gamma(1-\gamma) |\langle P^{n-n_0} f, (\nu^1_2 - \nu^2_2) P^{(k-2)m}\rangle| \\ &\quad + \cdots + \gamma(1-\gamma)^{k-1} |\langle P^{n-n_0} f, \nu^1_k - \nu^2_k\rangle| + (1-\gamma)^k |\langle P^{n-n_0} f, \mu^1_k - \mu^2_k\rangle|. \end{aligned}$$

Notice that for any $j \in \mathbb{N}$ and $l \in \{1, \ldots, k\}$,

$$(3.16)\qquad |\langle P^j f, \nu^1_l - \nu^2_l\rangle| \le \varepsilon/2.$$

Indeed, condition (3.3) together with $\nu^1_l(B(z,\delta)) = \nu^2_l(B(z,\delta)) = 1$ ensure that for each $l \in \{1, \ldots, k\}$ we have

$$\begin{aligned} |\langle P^j f, \nu^1_l - \nu^2_l\rangle| &= \Big| \int_{B(z,\delta)} P^j f(x)\,\nu^1_l(dx) - \int_{B(z,\delta)} P^j f(y)\,\nu^2_l(dy)\Big| \\ &= \Big| \int_{B(z,\delta)} \Big(\int_{B(z,\delta)} (P^j f(x) - P^j f(y))\,\nu^1_l(dx)\Big)\,\nu^2_l(dy)\Big| \\ &\le \int_{B(z,\delta)} \Big(\int_{B(z,\delta)} |P^j f(x) - P^j f(y)|\,\nu^1_l(dx)\Big)\,\nu^2_l(dy) \le \varepsilon/2. \end{aligned}$$

Moreover,

$$(3.17)\qquad |\langle P^{n-n_0} f, \mu^1_k - \mu^2_k\rangle| \le \|f\| \mu^1_k(X) + \|f\| \mu^2_K(X) = 2\|f\|.$$

Finally, the estimates (3.15)–(3.17) yield

$$|P^n(x_1) - P^n(x_2)| \le \frac{\varepsilon}{2}(\gamma + \gamma(1-\gamma) + \cdots + \gamma(1-\gamma)^{k-1}) + 2\|f\|(1-\gamma)^k$$
$$= \frac{\varepsilon}{2}(1-(1-\gamma)^k) + 2\|f\|(1-\gamma)^k < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \ \blacksquare$$

It follows from the theorem above that (A2)$'$ can be derived from the other conditions. To prove this we need an additional lemma. Recall that $\mathcal{C}(X,\delta)$ denotes the family of all subsets of X which can be covered by a finite number of closed balls of radius δ.

LEMMA 3.2. *For every compact set $K \subset X$ and for each $\delta > 0$ there exists a continuous function $f : X \to [0,1]$ such that* $\operatorname{supp} f \in \mathcal{C}(X,\delta)$ *and* $f \ge 1_K$.

Proof. Define $\tau\colon [0,\infty) \to [0,1]$ and $f\colon X \to [0,1]$ by

$$\tau(t) = \max\{1-t, 0\} \qquad (t \ge 0),$$
$$f(x) = \tau\left(\frac{3}{\delta}\operatorname{dist}(x,K)\right) \qquad (x \in X).$$

Plainly, τ and f are continuous and $f \ge 1_K$ as $f(x) = \tau(0) = 1$ for each $x \in K$.

Notice that

$$\operatorname{supp} f = \operatorname{cl}\{x \in X : f(x) \neq 0\} = \operatorname{cl}\{x \in X : \operatorname{dist}(x,K) < \delta/3\}$$
$$\subset \{x \in X : \operatorname{dist}(x,K) < \delta/2\}.$$

Since K is compact, we may find an integer N and $x_1,\ldots,x_N \in X$ such that

$$K \subset \bigcup_{j=1}^{N} B(x_j, \delta/2).$$

It remains to prove that $\operatorname{supp} f \subset \bigcup_{j=1}^N B(x_j,\delta)$. Fix $x \in \operatorname{supp} f$. Then $\operatorname{dist}(x,K) < \delta/2$, so there exists $z \in K$ such that $\rho(x,z) < \delta/2$. Clearly, $z \in B(x_i,\delta/2)$ for some $i \in \{1,\ldots,N\}$, so

$$\rho(x,x_i) \le \rho(x,z) + \rho(z,x_i) < \delta. \ \blacksquare$$

COROLLARY 3.1. *Suppose that conditions* (A1)–(A3) *hold. Then* (A2)$'$ *holds as well.*

Proof. Let C be a non-empty compact subset of X. To prove the tightness of $\{P^n(x,\cdot) : n \in \mathbb{N}, x \in C\}$ it is enough to show that each sequence in this family admits a tight subsequence. Indeed, let $(m_n)_{n\in\mathbb{N}}$ be a strictly increasing sequence of natural numbers, and let $(x_n)_{n\in\mathbb{N}}$ be a sequence of points in C. Suppose that $(k_n)_{n\in\mathbb{N}}$ is a strictly increasing sequence of natural numbers such that $\{P^{m_{k_n}}(x_{k_n},\cdot) : n \in \mathbb{N}\}$ is tight. Then by the Prokhorov Theorem a subsequence of $(P^{m_{k_n}}(x_{k_n},\cdot))_{n\in\mathbb{N}}$ is weakly convergent. Thus we have shown that each sequence in the family under consideration admits a weakly convergent subsequence. Applying [2, Theorem 7, Appendix II], we infer that the family is tight.

Let $(P^{m_n}(x_n,\cdot))_{n\in\mathbb{N}}$ be a subsequence of $\{P^n(x,\cdot) : n\in\mathbb{N},\, x\in C\}$. Fix $\delta>0$ and let $x_0\in C$. It follows from (A2) that there exists a compact set $K\subset X$ such that

$$P^n(x_0,K)\ge 1-\delta/2 \quad (n\in\mathbb{N}). \tag{3.18}$$

Appealing to Lemma 3.2, we find a function $f\colon X\to[0,1]$ with

$$f\ge 1_K \quad\text{and}\quad \operatorname{supp} f\in\mathcal{C}(X,\delta).$$

Certainly, $f\in C_{bs}(X)$. Making use of (A1) we deduce that the family $\{P^n f : n\in\mathbb{N}\}$ is equicontinuous at all points of X. Moreover, it is uniformly bounded with constant 1. Hence, by the Arzelà–Ascoli Theorem, we may extract a subsequence $(P^{m_{k_n}}f)_{n\in\mathbb{N}}$ which is uniformly convergent to a continuous function φ on C. Applying Theorem 3.1 we conclude that

$$\lim_{n\to\infty}|P^n f(x)-P^n f(y)|=0 \quad (x,y\in X).$$

Thus, for all $x,y\in C$ we get

$$|\varphi(x)-\varphi(y)|=\lim_{n\to\infty}|P^{m_{k_n}}f(x)-P^{m_{k_n}}f(y)|=0,$$

so φ is constant, say $\varphi(x)=m$ for each $x\in C$. Since the sequence $(P^{m_{k_n}}f)_{n\in\mathbb{N}}$ is uniformly convergent on C, we may pick $n_0\in\mathbb{N}$ such that

$$|P^{m_{k_n}}f(x)-m|<\delta/4 \quad (n\ge n_0 \text{ and } x\in C).$$

Consequently,

$$|P^{m_{k_n}}f(x)-P^{m_{k_n}}f(y)|<|P^{m_{k_n}}f(x)-m|+|m-P^{m_{k_n}}f(y)|<\delta/2 \tag{3.19}$$

for $n\ge n_0$ and $x,y\in C$. Set $A=\operatorname{supp} f$. Let $n\ge n_0$. By (3.19) applied to x_0, x_{k_n} and (3.18) we obtain

$$\begin{aligned}
P^{m_{k_n}}(x_{k_n},A)
&=\int_X 1_A(y)\,P^{m_{k_n}}(x_{k_n},dy)\ge\int_X f(y)\,P^{m_{k_n}}(x_{k_n},dy)=P^{m_{k_n}}f(x_{k_n})\\
&=P^{m_{k_n}}f(x_0)+(P^{m_{k_n}}f(x_{k_n})-P^{m_{k_n}}f(x_0))>P^{m_{k_n}}f(x_0)-\delta/2\\
&\ge P^{m_{k_n}}1_K(x_0)-\delta/2>1-\delta.
\end{aligned}$$

Since each measure $P^{m_{k_i}}(x_{k_i},\cdot)$ ($i\in\{1,\dots,n_0-1\}$) is tight in view of the Ulam Theorem ([2, Theorem 1.4] or [8, Theorem 1.17]), Proposition 2.1 yields a set $B\in\mathcal{C}(X,\delta)$ such that

$$P^{m_{k_i}}(x_{k_i},B)>1-\delta \quad (i\in\{1,\dots,n_0-1\}).$$

Therefore, letting $D=A\cup B$ we see that $D\in\mathcal{C}(X,\delta)$ and

$$P^{m_{k_n}}(x_{k_n},D)>1-\delta \quad (n\in\mathbb{N}).$$

We have proved that for any $\delta > 0$ there exists a set $D \in \mathcal{C}(X, \delta)$ such that $P^{m_{k_n}}(x_{k_n}, D) > 1 - \delta$ for each $n \in \mathbb{N}$. Now Proposition 2.1 implies that $\{P^{m_{k_n}}(x_{k_n}, \cdot) : n \in \mathbb{N}\}$ is tight. ∎

Observe that condition (A2) implies the boundedness in probability on average at each point of X. In particular, it ensures the existence of an invariant probability measure (see Proposition 5.3 or e.g. [11]).

Now we are prepared to prove the main result of this section (cf. [9, Theorem 4.3]).

THEOREM 3.2. *Suppose that conditions* (A1)–(A3) *are satisfied. Then the chain* $(\xi_n)_{n\in\mathbb{N}_0}$ *is asymptotically stable with a unique invariant probability measure* π. *Furthermore, the convergence* $P^n(x, \cdot) \xrightarrow{w} \pi$ *is uniform with respect to* x *on compact subsets of* X.

Proof. Let π be an arbitrary invariant probability measure for $(\xi_n)_{n\in\mathbb{N}_0}$ (see Proposition 5.3). We shall deduce the asymptotic stability of $(\xi_n)_{n\in\mathbb{N}_0}$ by using Theorem 3.1 and the Lebesgue Dominated Convergence Theorem. Indeed, let $\mu \in \mathcal{M}_1(X)$ and $f \in C(X)$. Then

$$\begin{aligned}
\Big|\int_X f(x)\,\mu P^n(dx) - \int_X f(y)\,\pi(dy)\Big| &= \Big|\int_X f(x)\,\mu P^n(dx) - \int_X f(y)\,\pi P^n(dy)\Big| \\
&= \Big|\int_X P^n f(x)\,\mu(dx) - \int_X P^n f(y)\,\pi(dy)\Big| \\
&= \Big|\int_X \Big(\int_X P^n f(x)\,\mu(dx) - P^n f(y)\Big)\,\nu(dy)\Big| \\
&= \Big|\int_X \Big(\int_X (P^n f(x) - P^n f(y))\,\mu(dx)\Big)\,\nu(dy)\Big| \\
&\le \int_X \Big(\int_X |P^n f(x) - P^n f(y)|\,\mu(dx)\Big)\,\nu(dy) \to 0 \quad \text{as } n \to \infty.
\end{aligned}$$

Consequently, $\mu P^n \xrightarrow{w} \pi$.

One should keep in mind that π is a unique probability invariant measure for $(\xi_n)_{n\in\mathbb{N}_0}$ (see Remark 2.1).

Set $\pi(f) = \int_X f(x)\,\pi(dx)$. In particular, for each $x \in X$ we have $\delta_x P^n \xrightarrow{w} \pi$ (where δ_x is the Dirac measure supported at x), so

$$\lim_{n\to\infty} |P^n f(x) - \pi(f)| = 0 \quad (f \in C(X)). \tag{3.20}$$

We shall prove that for each $f \in C(X)$ the sequence $(P^n f)_{n\in\mathbb{N}}$ converges to $\pi(f)$ uniformly on compact subsets of X. Indeed, fix $f \in C(X)$, $f \neq 0$ and a compact set $C \subset X$. Assume to the contrary that $(P^n f)_{n\in\mathbb{N}}$ does not converge uniformly to $\pi(f)$ on C. Then we may find $\varepsilon > 0$, a strictly increasing sequence $(k_n)_{n\in\mathbb{N}}$ of integers and a sequence $(x_n)_{n\in\mathbb{N}}$ of points of C such that

$$|P^{k_n} f(x_n) - \pi(f)| \geq \varepsilon \quad (n \in \mathbb{N}). \tag{3.21}$$

Since C is compact, we can assume that $x_n \to x_0 \in C$. It follows from (A2)$'$ that there is a compact set $K \subset X$ such that

$$\sup_{x\in C} P^n(x, X\setminus K) \leq \frac{\varepsilon}{8\|f\|} \text{ and } P^n(x_0, X\setminus K) \leq \frac{\varepsilon}{8\|f\|} \quad (n \in \mathbb{N}). \tag{3.22}$$

Let $B \subset X$ be an open ball containing K. The Urysohn Lemma yields a continuous function $\varphi\colon X \to [0,1]$ such that $\varphi|_K = 1$ and $\varphi|_{X\setminus B} = 0$. Set $f_0 = f\varphi$. Then $f_0 \in C_{bs}(X)$ and $\|f - f_0\| \leq \|f\|$. Applying (A1) to f_0 and x_0 we find $\delta > 0$ such that

$$|P^n f_0(x) - P^n f_0(x_0)| \leq \varepsilon/8 \quad (x \in B(x_0,\delta),\, n \in \mathbb{N}). \tag{3.23}$$

Pick $N \in \mathbb{N}$ such that $x_n \in B(x_0,\delta)$ $(n \geq N)$. Appealing to (3.23) with $x = x_n$ $(n \geq N)$ and making use of (3.20), we choose $n_0 \geq N$ such that

$$|P^n f(x_0) - \pi(f)| \leq \varepsilon/8 \quad (n \geq n_0). \tag{3.24}$$

Let $n \geq n_0$. Then, applying (3.23), (3.24) and (3.22) successively we obtain

$$\begin{aligned} |P^{k_n} f(x_n) - \pi(f)| &\leq |P^{k_n} f(x_n) - P^{k_n} f_0(x_n)| + |P^{k_n} f_0(x_n) - P^{k_n} f_0(x_0)| \\ &\quad + |P^{k_n} f_0(x_0) - P^{k_n} f(x_0)| + |P^{k_n} f(x_0) - \pi(f)| \\ &\leq \int_{X\setminus K} |(f-f_0)(y)|\, P^{k_n}(x_n, dy) + \varepsilon/8 \\ &\quad + \int_{X\setminus K} |(f-f_0)(y)|\, P^{k_n}(x_0, dy) + \varepsilon/8 \\ &\leq \varepsilon/4 + \|f\| P^{k_n}(x_n, X\setminus K) + \|f\| P^{k_n}(x_0, X\setminus K) \\ &\leq \varepsilon/4 + 2\|f\| \cdot \frac{\varepsilon}{8\|f\|} = \frac{\varepsilon}{2} < \varepsilon, \end{aligned}$$

contrary to (3.21), whence the proof is complete. ■

Consider the following two conditions:

(A1)$'$ For each $f \in C(X)$ the family $\{P^n f : n \in \mathbb{N}\}$ is equicontinuous at all points of X.

(A3)$''$ For any point z and each $\delta > 0$ there exist $\alpha > 0$ and $N \in \mathbb{N}$ such that $P^N(x, B(z,\delta)) > \alpha$ for all x.

The conditions above are stronger than (A1) and (A3), respectively. Observe that (A1)$'$ and (A3)$''$ together imply (A2). Indeed, if $z \in X$, $\delta > 0$ are arbitrary and $\alpha > 0$, $N \in \mathbb{N}$ satisfy (A3)$''$, then $P^n(x, B(z,\delta)) > \alpha$ for all

$n \geq N$, because for each $k \in \mathbb{N}$ we have

$$P^{N+k}(x, B(z,\delta)) \geq \int_X P^N(y, B(z,\delta))\, P^k(x,dy) \geq \alpha P^k(x,X) = \alpha.$$

Consequently, for any integer $n > N$ and for each $x \in X$ we have

$$\frac{1}{n}\sum_{i=1}^{n} P^i(x, B(z,\delta)) \geq \frac{1}{n}\sum_{i=N+1}^{n} P^i(x, B(z,\delta)) \geq \alpha\frac{n-N}{n} = \alpha\Big(1 - \frac{N}{n}\Big).$$

In the proof of Proposition 2.1 of [10] it is shown this implies the tightness of $\{P^n(z,\cdot) : n \in \mathbb{N}\}$, hence condition (A2) holds. This leads to the following conclusion:

Corollary 3.2. *Assume that conditions* (A1)$'$ *and* (A3)$''$ *are satisfied. Then the chain* $(\xi_n)_{n\in\mathbb{N}_0}$ *is asymptotically stable with a unique invariant probability measure* π *and* $P^n(x,\cdot) \xrightarrow{w} \pi$ *uniformly with respect to* $x \in X$ *on compact sets of* X.

To end this section, we consider the case when X is a locally compact and separable metric space. Of course, we can treat this as a special case of the previous considerations as every locally compact, separable metric space is homeomorphic to a Polish space.

We introduce one more condition:

(A1)^C For any $f \in C(X)$ with compact support the family $\{P^n f : n \in \mathbb{N}\}$ is uniformly equicontinuous on each compact subset of X, that is, if $K \subset X$ is compact, then

$$(\forall \varepsilon > 0)(\exists \delta > 0)(\forall n \in \mathbb{N})(\forall x, y \in K)(y \in B(x,\delta) \Rightarrow |P^n f(x) - P^n f(y)| < \varepsilon).$$

The uniform equicontinuity condition is equivalent to the pointwise equicontinuity on compact sets (cf. Proposition 5.1), so in general, condition (A1)^C is seemingly stronger than (A1). It turns out that for locally compact and separable metric spaces these conditions are equivalent:

Proposition 3.1. *Let* X *be a locally compact, separable space. Then conditions* (A1) *and* (A1)^C *are equivalent.*

Proof. Each locally compact and separable metric space (which is also σ-compact under these assumptions) admits a metric which is equivalent to the original one and has the additional property that each closed and bounded set is compact ([13]). Consequently, when X is endowed with such a metric, bounded supports of functions are already compact. In this case it is sufficient to see that the implication (A1)^C $\Rightarrow$ (A1) can be derived from the fact that every compact set in X is contained in an open relatively compact set. The opposite implication is obvious (cf. Proposition 5.1). ■

Let us recall a simple consequence of the Arzelà–Ascoli theorem:

Remark 3.1. Suppose there exists a unique invariant probability measure π for the chain $(\xi_n)_{n\in\mathbb{N}_0}$ such that $P^n(x,\cdot)\xrightarrow{w}\pi$ uniformly with respect to $x\in X$ on compact sets. Then for each $f\in C(X)$ the family $\{P^n f: n\in\mathbb{N}\}$ is uniformly equicontinuous on each compact subset of X.

In fact, Proposition 3.1 implies that in case X is locally compact and separable, the above assumptions can be weakened by (A1). Thus, Remark 3.1 together with Theorem 3.2 yield the following corollary:

Corollary 3.3. *Let X be a locally compact, separable space. Suppose the conditions* (A2) *and* (A3) *are satisfied. Then* (A1) (*or, equivalently,* $(\mathrm{A1})^C$) *holds if and only if the chain $(\xi_n)_{n\in\mathbb{N}_0}$ admits a unique invariant probability measure π such that $P^n(x,\cdot)\xrightarrow{w}\pi$ uniformly with respect to $x\in X$ on compact sets.*

4. An example of application. Let $\phi_0\colon \Omega\to\mathbb{R}^d$ be a random variable with a given distribution. Assume that $(\eta_n)_{n\in\mathbb{N}}$ is a sequence of mutually independent and identically distributed random variables taking values in $\mathbb{R}^p$. Let Γ be the probability distribution of η_n ($n\in\mathbb{N}$) and suppose that Γ has finite expectation and finite variance. Moreover, assume that the sequence $(\eta_n)_{n\in\mathbb{N}}$ is independent of ϕ_0. Additionally, fix matrices $F\in\mathbb{R}^d_d$ and $G\in\mathbb{R}^p_d$. For each $n\in\mathbb{N}_0$ we define a random variable ϕ_n as follows:

$$\phi_{n+1}=F\phi_n+G\eta_{n+1}. \tag{4.1}$$

From (4.1) it follows easily that

$$\phi_n=F^n\phi_0+\sum_{i=0}^{n-1}F^iG\eta_{n-i}\quad (n\in\mathbb{N}). \tag{4.2}$$

Under these assumptions, the sequence $(\phi_n)_{n\in\mathbb{N}_0}$ is a time-homogeneous Markov chain. The model defined in this way is called a *linear state space model* and denoted by $\mathrm{LSS}(F,G)$ (see [8]).

Throughout this section P stands for the stochastic kernel of the model $\mathrm{LSS}(F,G)$, that is, the stochastic kernel of the chain $(\phi_n)_{n\in\mathbb{N}_0}$.

For $f\colon\mathbb{R}^d\to\mathbb{R}^m$ we denote by Df its Jacobian matrix (or gradient if $m=1$). Thus, $Df(x)=\big[\frac{\partial f_i}{\partial x_j}(x)\colon i\le m,\ j\le d\big]$ ($x\in\mathbb{R}^d$).

The next remark describes the Jacobian matrix of the expectation of the chain $(f(\phi_n))_{n\in\mathbb{N}}$ at time n (cf. [8, Theorem 7.5.1]).

Remark 4.1. Let $f\colon\mathbb{R}^d\to\mathbb{R}^m$ be a Borel function which has all its first-order partial derivatives bounded. Then for each $n\in\mathbb{N}$ the function $\mathbb{R}^d\ni x\mapsto\mathbb{E}_x(f(\phi_n))\in\mathbb{R}^m$ is differentiable and

$$D\mathbb{E}_x(f(\phi_n))=F^n\mathbb{E}_x(Df(\phi_n))\quad (x\in\mathbb{R}^d).$$

Indeed, applying the Lebesgue Dominated Convergence Theorem, we obtain

$$\begin{aligned} D\mathbb{E}_x(f(\phi_n)) &= D\mathbb{E}_x\Big(f\Big(F^n\phi_0 + \sum_{i=0}^{n-1} F^iG\eta_{n-i}\Big)\Big) \\ &= D\mathbb{E}\Big(f\Big(F^n x + \sum_{i=0}^{n-1} F^iG\eta_{n-i}\Big)\Big) \\ &= F^n\mathbb{E}\Big(Df\Big(F^n x + \sum_{i=0}^{n-1} F^iG\eta_{n-i}\Big)\Big) \\ &= F^n\mathbb{E}_x(Df(\phi_n)) \quad (x \in \mathbb{R}^d). \end{aligned}$$

The next theorem gives a sufficient condition for the model $\mathrm{LSS}(F,G)$ to have the e-property (this is, in fact, a sufficient condition for $(\phi_n)_{n\in\mathbb{N}_0}$ to satisfy (A1); consult also [8, Theorem 7.5.3]).

THEOREM 4.1. *Let* $\|\cdot\|$ *be the matrix norm arising from the underlying operator norm. Suppose that* $\sup_{n\in\mathbb{N}} \|F^n\| < \infty$. *Then the model* $\mathrm{LSS}(F,G)$ *satisfies* (A1).

Proof. As $\mathbb{R}^d$ is locally compact and separable, Proposition 3.1 shows that it is sufficient to prove that for any $f \in C(\mathbb{R}^d)$ with compact support the family $\{P^n f : n \in \mathbb{N}\}$ is equicontinuous on compact sets.

Fix $f \in \mathcal{C}_c^\infty(\mathbb{R}^d)$. Then $\sup_{x\in\mathbb{R}^d} \|Df(x)\|_d \le M$ for some $M > 0$. Remark 4.1 ensures that for each $n \in \mathbb{N}$ we have

$$\begin{aligned} \|DP^n f(x)\|_d &= \Big\|D\int_{\mathbb{R}^d} f(y)\,\mathbb{P}_x(\phi_n \in dy)\Big\|_d = \Big\|D\int_{\mathbb{R}^d} f(\phi_n)\,d\mathbb{P}_x\Big\|_d \\ &= \|D\mathbb{E}_x f(\phi_n)\|_d = \|F^n\mathbb{E}_x(Df(\phi_n))\|_d \\ &\le \|F^n\|\,\|\mathbb{E}_x(Df(\phi_n))\|_d \le M\sup_{k\in\mathbb{N}}\|F^k\|. \end{aligned}$$

Consequently, the family $\{P^n f : n \in \mathbb{N}\}$ is uniformly equicontinuous (cf. Proposition 5.2). Since $\mathcal{C}_c^\infty$ functions are dense in the space of bounded continuous functions with compact supports, the conclusion follows. ∎

Let A be a $d\times d$ matrix. Denote by $\sigma(A)$ its spectrum, that is, $\sigma(A) = \{\lambda \in \mathbb{C} : \det(A - \lambda I) = 0\}$ (where I stands for the identity matrix). Recall that $\rho(F) = \max\{|\lambda| : \lambda \in \sigma(F)\}$) is called the *spectral radius* of A. The following fact is known in the theory of linear models (see [3] for its proof):

THEOREM 4.2. *Suppose* $\rho(A) > 0$. *Then for each matrix norm* $\|\cdot\|$ *the sequence* $(n^{-1}\ln(\|A^n\|))_{n\in\mathbb{N}}$ *converges to* $\ln(\rho(A))$.

Of course, for any real a, b with $a < \rho(A) < b$ there exist $c > 1$ and an integer n_0 such that for each $n \ge n_0$,

$$c^{-1}a^n \le \|A^n\| \le cb^n.$$

Now, if we assume that $0 < \rho(A) < 1$, then we may choose b such that $0 < \rho(A) < b < 1$ and by the above observation, we obtain $\|A^n\| \to 0$. Hence, $\sup_{n\in\mathbb{N}} \|A^n\|$ is finite. Although not required here, it is worth noting that the condition $\rho(A) > 0$ holds even in case A is non-singular.

Assuming that $0 < \rho(F) < 1$, the model $\mathrm{LSS}(F, G)$ satisfies (A2), i.e. the family $\{P^n(x, \cdot) : n \in \mathbb{N}\}$ is tight for each $x \in X$. Let us quote the following theorem from [8, Theorem 12.5.9].

THEOREM 4.3. *Suppose* $0 < \rho(F) < 1$. *Then the model* $\mathrm{LSS}(F, G)$ *satisfies* (A2).

Proof. Taking into account Proposition 2.2, it suffices to prove that there exists a norm-like function $V : \mathbb{R}^d \to [0, \infty)$ satisfying (2.2) with respect to $\mathcal{A} = \{P^n(x, \cdot) : n \in \mathbb{N}\}$.

Let $\|\cdot\|$ be a matrix norm arising from the underlying operator norm. It is well-known that it is a Banach-algebra norm, that is, $\|\cdot\|$ is submultiplicative. As $\rho(F^T) = \rho(F) < 1$, Theorem 4.2 implies that there exist constants $b_1, b_2 < 1$ and $c_1, c_2 \in \mathbb{R}$ such that for almost all $n \in \mathbb{N}$ we have $\|F^n\| \le c_1 b_1^n$ and $\|(F^T)^n\| \le c_2 b_2^n$. Thus, for almost all $n \in \mathbb{N}$ we obtain

$$\|(F^T)^n F^n\| \le \|(F^T)^n\| \, \|F^n\| \le c_1 c_2 (b_1 b_2)^n.$$

In particular, the matrix

$$M = I + \sum_{n=1}^{\infty} (F^T)^n F^n$$

exists. Notice that

$$\begin{aligned} x^T M x &= x^T x + \sum_{n=1}^{\infty} x^T (F^T)^n F^n x = \|x\|^2 + \sum_{n=1}^{\infty} (F^n x)^T (F^n x) \\ &= \|x\|^2 + \sum_{n=1}^{\infty} \|F^n x\|^2 \quad (x \in \mathbb{R}^d). \end{aligned} \tag{4.3}$$

Hence, the matrix M is positive-definite. Define

$$|x|_M^2 = x^T M x \quad (x \in \mathbb{R}^d).$$

It follows from (4.3) that

$$\begin{aligned} |Fx|_M^2 = \|Fx\|^2 + \sum_{n=1}^{\infty} \|F^{n+1} x\|^2 &= \sum_{n=1}^{\infty} \|F^n x\|^2 \\ &< x^T M x = |x|_M^2 \quad (x \in \mathbb{R}^d \setminus \{0\}). \end{aligned}$$

Thus, we may choose $\alpha < 1$ with

$$|Fx|_M^2 \le \alpha |x|_M^2 \quad (x \in \mathbb{R}^d). \tag{4.4}$$

Define $V:\mathbb{R}^d\to[0,\infty)$ by

$$V(x)=|x-a|_M^2\quad(x\in\mathbb{R}^d)\quad\text{where}\quad a=\Big(I+\sum_{n=1}^{\infty}F^n\Big)G\mathbb{E}\eta_1.$$

Since $a=Fa-G\,\mathbb{E}\eta_1=Fa-G\,\mathbb{E}(\eta_{n+1})$ $(n\in\mathbb{N})$, we conclude that

$$\begin{aligned}(4.5)\qquad V(\phi_{n+1})&=|F\phi_n+G\eta_{n+1}-a|_M^2\\ &=|F\phi_n+G\eta_{n+1}-(Fa+G\,\mathbb{E}\eta_{n+1})|_M^2\\ &=((F\phi_n)^T-(Fa)^T+(G\eta_{n+1})^T-(G\,\mathbb{E}\eta_{n+1})^T)\\ &\quad\times M(F\phi_n-Fa+G\eta_{n+1}-G\,\mathbb{E}\eta_{n+1})\\ &=((\phi_n-a)^TF^T+(\eta_{n+1}-\mathbb{E}\eta_{n+1})^TG^T)M(F(\phi_n-a)+G(\eta_{n+1}-\mathbb{E}\eta_{n+1}))\\ &=|F(\eta_n-a)|_M^2+|G(\eta_{n+1}-\mathbb{E}\eta_{n+1})|_M^2+(\eta_{n+1}-\mathbb{E}\eta_{n+1})^TG^TMF(\eta_n-a)\\ &\quad+(\eta_n-a)^TF^TMG(\eta_{n+1}-\mathbb{E}\eta_{n+1})\quad(n\in\mathbb{N}).\end{aligned}$$

Making use of (4.2) we infer that for each $n\in\mathbb{N}$ and $i\le n$ the random variables η_{n+1} and ϕ_i are mutually independent. Consequently, taking into account (4.5) and applying (4.4) to an arbitrary point $x\in\mathbb{R}^d$ we obtain

$$\begin{aligned}\mathbb{E}_x(V(\phi_{n+1})\,|\,\phi_0,\ldots,\phi_n)&=|F_n(\phi_n-a)|_M^2+\mathbb{E}_x(|G(\eta_{n+1}-\mathbb{E}_x\eta_{n+1})|_M^2)\\ &\le\alpha V(\phi_n)+\mathbb{E}_x(|G(\eta_{n+1}-\mathbb{E}_x\eta_{n+1})|_M^2)\quad(n\in\mathbb{N}).\end{aligned}$$

So,

$$\begin{aligned}\mathbb{E}_xV(\phi_{n+1})&\le\alpha\mathbb{E}_xV(\phi_n)+\mathbb{E}_x(|G(\eta_{n+1}-\mathbb{E}_x\eta_{n+1})|_M^2)\\ &=\alpha\mathbb{E}_xV(\phi_n)+\mathbb{E}(|G(\eta_1-\mathbb{E}_x\eta_1)|_M^2)\quad(n\in\mathbb{N}).\end{aligned}$$

Letting $\beta=\mathbb{E}(|G(\eta_1-\mathbb{E}_x\eta_1)|_M^2)$ we have

$$\mathbb{E}_xV(\phi_{n+1})\le\alpha^n\mathbb{E}_xV(\phi_1)+\beta\sum_{k=0}^{n-1}\alpha^k\quad(n\in\mathbb{N})$$

and the right-hand side tends to a real number as $n\to\infty$. Consequently,

$$\sup_{n\in\mathbb{N}}\int_X V(y)\,P^n(x,dy)=\sup_{n\in\mathbb{N}}\mathbb{E}_xV(\phi_n)<\infty,$$

so condition (2.2) holds.

It remains to show that V is a norm-like function. For this purpose we define

$$C_n=\{x\in\mathbb{R}^d:V(x)\le n\}\quad(n\in\mathbb{N}).$$

For each $n\in\mathbb{N}$ the set C_n is closed (as V is continuous) and bounded.

Indeed, let $x \in C_n$. Then, by (4.3),

$$\begin{aligned}\|x-a\|^2 &\le \|x-a\|^2 + \sum_{n=1}^{\infty} \|F^n(x-a)\|^2 = (x-a)^T M(x-a) \\ &= |x-a|_M^2 = V(x) \le n.\end{aligned}$$

Thus, $(C_n)_{n\in\mathbb{N}}$ is an increasing sequence of compact subsets of $\mathbb{R}^d$. Moreover, $V(x) \ge n$ for each $n \in \mathbb{N}$ and $x \in X \setminus C_n$, so

$$\lim_{n\to\infty} \inf_{x\in X\setminus C_n} V(x) = \infty. \ \blacksquare$$

Let us determine the explicit form of the kernel P^n $(n \in \mathbb{N})$ for the model LSS(F,G). We have

$$\begin{aligned}P^n(x,A) &= \mathbb{P}_x(\phi_n \in A) = \mathbb{P}\Big(F^n\phi_0 + \sum_{i=0}^{n-1} F^i G\eta_{n-i} \in A \;\Big|\; \phi_0 = x\Big) \\ &= \frac{\mathbb{P}(F^n x + \sum_{i=0}^{n-1} F^i G\eta_{n-i} \in A, \phi_0 = x)}{\mathbb{P}(\phi_0 = x)} = \mathbb{P}\Big(F^n x + \sum_{i=0}^{n-1} F^i G\eta_{n-i} \in A\Big) \\ &= \mathbb{P}\Big(\sum_{i=0}^{n-1} F^i G\eta_{n-i} \in A - F^n x\Big) \quad (x \in \mathbb{R}^d,\ A \in \mathcal{B}_{\mathbb{R}^d},\ n \in \mathbb{N}).\end{aligned}$$

Condition (A3) is easy to verify in case the distribution Γ is given. The form of the kernel P^n enables us to see that the model LSS(F,G) has the Feller property. In fact, for each $f \in C(X)$ we have

$$Pf(x) = \int_{\mathbb{R}^d} f(y)\,\mathbb{P}(Fx + G\eta_1 \in dy) = \int_{\mathbb{R}^d} f(Fx+y)\,\mathbb{P}(G\eta_1 \in dy) \quad (x \in \mathbb{R}^d).$$

By continuity of $x \mapsto f(Fx+y)$ $(y \in \mathbb{R}^d)$ and the Lebesgue Dominated Convergence Theorem we conclude that Pf is continuous.

Taking into account all the above considerations we can state the following theorem:

THEOREM 4.4. *Assume that* $0 < \rho(F) < 1$ *and suppose that there exists a point* $z \in \mathbb{R}^d$ *with the following property: for any compact set* $C \subset X$ *and for each* $\delta > 0$ *there is an integer* $N \in \mathbb{N}$ *such that*

$$\mathbb{P}\Big(\sum_{i=0}^{N-1} F^i G\eta_{N-i} \in B(z - F^N x, \delta)\Big) > 0 \quad \textit{for all } x \in C.$$

Then the model LSS(F,G) *is asymptotically stable with a unique invariant probability measure* π. *Moreover, the convergence* $P^n(x,\cdot) \xrightarrow{w} \pi$ *is uniform with respect to* x *on compact subsets of* $\mathbb{R}^d$.

5. Appendix. Here we list some standard facts that have been used in this paper. First, we show the equivalence of uniform equicontinuity with pointwise equicontinuity on compact sets (this has been applied in the proof of Proposition 3.1).

PROPOSITION 5.1. *Let (K, ρ) be a compact metric space and $\mathcal{F}$ a family of real-valued functions on K. Then $\mathcal{F}$ is equicontinuous at all points of K if and only if it is uniformly equicontinuous on K, i.e.*

$$(\forall \varepsilon > 0)(\exists \delta > 0)(\forall f \in \mathcal{F})(\forall x, y \in K)(\rho(x, y) < \delta \Rightarrow |f(x) - f(y)| < \varepsilon).$$

Proof. Suppose $\mathcal{F}$ is equicontinuous at all points of K. Fix $\varepsilon > 0$. Then, for each $z \in K$ we may find $\delta(z) > 0$ such that

$$|f(x) - f(z)| < \varepsilon/3 \quad (f \in \mathcal{F},\ x \in B(z, \delta(z))). \tag{5.1}$$

By compactness there exist $z_1, \ldots, z_N \in K$ such that

$$K = \bigcup_{i=1}^{N} B\big(z_i, \delta(z_i)/3\big).$$

Set $\delta = \frac{1}{3} \min\{\delta(z_1), \ldots, \delta(z_N)\}$. Let $x, y \in K$ with $\rho(x, y) < \delta$. Then $x \in B(z_k, \delta(z_k)/3)$ and $y \in B(z_l, \delta(z_l)/3)$ for some $k, l \in \{1, \ldots, N\}$. We may assume that $\delta(z_l) < \delta(z_k)$. Then

$$\rho(z_k, z_l) \le \rho(z_k, x) + \rho(x, y) + \rho(y, z_l) < \frac{\delta(z_k)}{3} + \delta + \frac{\delta(z_l)}{3} < \delta(z_k).$$

Therefore, for each $f \in \mathcal{F}$ by (5.1) we obtain

$$\begin{aligned} |f(x) - f(y)| &\le |f(x) - f(z_k)| + |f(z_k) - f(z_l)| + |f(z_l) - f(y)| \\ &< \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon. \ \blacksquare \end{aligned}$$

The following criterion of equicontinuity has been used in the proof of Theorem 4.1.

PROPOSITION 5.2. *Let $\{f_n : n \in \mathbb{N}\}$ be a collection of real-valued continuous functions on $\mathbb{R}^d$ with bounded first-order partial derivatives. Suppose there exists a constant $M > 0$ such that*

$$\|Df_n(x)\| \le M \quad (x \in \mathbb{R}^d,\ n \in \mathbb{N}).$$

Then the functions f_n $(n \in \mathbb{N})$ are Lipschitz continuous with constant M and, in particular, the family $\{f_n : n \in \mathbb{N}\}$ is uniformly equicontinuous.

Proof. Fix $n \in \mathbb{N}$ and $x, y \in \mathbb{R}^d$. Define a function $\varphi : [0, 1] \to \mathbb{R}$ by

$$\varphi(t) = f_n(tx + (1 - t)y) \quad (t \in [0, 1]).$$

Then $\varphi(1) = f_n(x)$ and $\varphi(0) = f_n(y)$. Certainly, the function φ is continuous and differentiable in the interval $(0,1)$. The Lagrange Mean Value Theorem implies that there exists $s \in (0,1)$ such that

$$\frac{\varphi(1)-\varphi(0)}{1-0} = \varphi'(s).$$

The Cauchy–Schwarz inequality yields, for each $t \in (0,1)$,

$$\begin{aligned} |\varphi'(t)| &= \left| \left\langle Df_n(tx+(1-t)y) \,\middle|\, \frac{d}{dt}(tx+(1-t)y) \right\rangle \right| \\ &\le \|Df_n(tx+(1-t)y)\| \left\| \frac{d}{dt}(tx+(1-t)y) \right\| \le M\|x-y\|. \end{aligned}$$

Consequently,

$$|f_n(x)-f_n(y)| = |\varphi(1)-\varphi(0)| = |\varphi'(s)| \le M\|x-y\|.$$

The uniform equicontinuity of $\{f_n : n \in \mathbb{N}\}$ is now obvious. ■

Finally, we give a simple criterion for the existence of an invariant probability measure for Markov–Feller chains which are bounded in probability. This fact has been applied in the proof of Theorem 3.2.

PROPOSITION 5.3. *Let $(\xi_n)_{n\in\mathbb{N}_0}$ be a Markov chain with transition kernel P. Suppose $(\xi_n)_{n\in\mathbb{N}_0}$ has the Feller property and is bounded in probability on average at a point $x \in X$, that is, the family $\{n^{-1}\sum_{k=1}^n P^k(x,\cdot) : n \in \mathbb{N}\}$ of measures is tight. Then there exists an invariant probability measure for the chain.*

Proof. Set $\overline{P}_n = n^{-1}\sum_{k=1}^n P^k$ $(n \in \mathbb{N})$. Since the family $\{\overline{P}_n(x,\cdot) : n \in \mathbb{N}\}$ is tight, it contains a sequence $(\overline{P}_{k_n}(x,\cdot))_{n\in\mathbb{N}}$ converging weakly to a probability measure π ([2, Appendix III, Theorem 6]). The Feller property of $(\xi_n)_{n\in\mathbb{N}_0}$ implies that for each $f \in C(X)$ we have

$$\begin{aligned} \langle f, \pi P\rangle &= \langle Pf, \pi\rangle = \lim_{n\to\infty}\langle Pf, \overline{P}_{k_n}(x,\cdot)\rangle = \lim_{n\to\infty}\frac{1}{k_n}\sum_{i=1}^{k_n}\langle Pf, P^i(x,\cdot)\rangle \\ &= \lim_{n\to\infty}\frac{1}{k_n}\sum_{i=1}^{k_n}\langle f, P^{i+1}(x,\cdot)\rangle = \lim_{n\to\infty}\frac{1}{k_n}\sum_{i=2}^{k_n+1}\langle f, P^i(x,\cdot)\rangle \\ &= \lim_{n\to\infty}\frac{1}{k_n}\sum_{i=1}^{k_n}\langle f, P^i(x,\cdot)\rangle - \lim_{n\to\infty}\frac{1}{k_n}\langle f, P(x,\cdot)\rangle \\ &\quad + \lim_{n\to\infty}\frac{1}{k_n}\langle f, P^{k_n+1}(x,\cdot)\rangle \\ &= \lim_{n\to\infty}\langle f, \overline{P}_{k_n}(x,\cdot)\rangle = \langle f, \pi\rangle. \end{aligned}$$

Consequently, $\pi P = \pi$. Thus, π is an invariant measure. ■

References

[1] H. Bessaih, R. Kapica and T. Szarek, *The stability of stochastic shell models*, submitted, 2010.

[2] P. Billingsley, *Convergence of Probability Measures*, Wiley, New York, 1999.

[3] P. Caines, *Linear Stochastic Systems*, Wiley, New York, 1988.

[4] K. Horbacz, *Invariant measures for random dynamical systems*, Dissertationes Math. 451 (2008), 68 pp.

[5] A. Lasota, *From fractals to stochastic differential equations*, in: Chaos—The Interplay Between Stochastic and Deterministic Behaviour, Springer, Berlin, 1995, 235–255.

[6] A. Lasota, *Układy dynamiczne na miarach*, Wydawnictwo UŚ, 2008.

[7] A. Lasota and J. Yorke, *Lower bound technique for Markov operators and iterated function systems*, Random Comput. Dynam. 2 (1994), 41–77.

[8] S. P. Meyn and R. L. Tweedie, *Markov Chains and Stochastic Stability*, Springer, London, 1993.

[9] Ł. Stettner, *Remarks on ergodic conditions for Markov processes on Polish spaces*, Bull. Polish Acad. Sci. Math. 42 (1994), 103–114.

[10] T. Szarek, *Feller processes on nonlocally compact spaces*, Ann. Probab. 34 (2006), 1849–1863.

[11] T. Szarek, *The stability of Markov operators on Polish spaces*, Studia Math. 143 (2000), 145–152.

[12] T. Szarek, *Invariant measures for Markov operators with applications to function systems*, Studia Math. 154 (2003), 207–222.

[13] R. Williamson and L. Janos, *Constructing metrics with the Heine Borel property*, Proc. Amer. Math. Soc. 100 (1987), 567–573.

Dawid Czapla
Institute of Mathematics
University of Silesia
Bankowa 14
40-007 Katowice, Poland
E-mail: dczapla@us.edu.pl

Received 2.11.2011
and in final form 7.3.2012 (2591)