Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2008, Vol. 44, No. 6, 1090–1095
DOI: 10.1214/07-AIHP145

ANNALES DE L'INSTITUT HENRI POINCARÉ
PROBABILITÉS ET STATISTIQUES
www.imstat.org/aihp

Quasi-compactness and mean ergodicity for Markov kernels acting on weighted supremum normed spaces

Loïc Hervé

I.R.M.A.R., UMR-CNRS 6625, Institut National des Sciences Appliquées de Rennes, 20, Avenue des Buttes de Couësmes CS 14 315, 35043 Rennes Cedex, France. E-mail: Loic.Herve@insa-rennes.fr

Received 12 December 2006; revised 29 June 2007; accepted 10 September 2007

Abstract. Let P be a Markov kernel on a measurable space E with countably generated σ-algebra, let $w : E \to [1, +\infty[$ such that $Pw \le Cw$ with $C \ge 0$, and let $\mathcal{B}_w$ be the space of measurable functions on E satisfying $\|f\|_w = \sup\{w(x)^{-1}|f(x)|,\ x \in E\} < +\infty$. We prove that P is quasi-compact on $(\mathcal{B}_w, \|\cdot\|_w)$ if and only if, for all $f \in \mathcal{B}_w$, $(\frac{1}{n}\sum_{k=1}^{n} P^k f)_n$ contains a subsequence converging in $\mathcal{B}_w$ to $\Pi f = \sum_{i=1}^{d} \mu_i(f) v_i$, where the v_i's are non-negative bounded measurable functions on E and the μ_i's are probability distributions on E. In particular, when the space of P-invariant functions in $\mathcal{B}_w$ is finite-dimensional, uniform ergodicity is equivalent to mean ergodicity.

Résumé. Soit P un noyau markovien sur un espace mesurable E muni d'une tribu à base dénombrable, soit $w : E \to [1, +\infty[$ tel que $Pw \le Cw$, avec $C \ge 0$, et soit $\mathcal{B}_w$ l'espace des fonctions f mesurables de E dans $\mathbb{C}$ telles que $\|f\|_w = \sup\{w(x)^{-1}|f(x)|,\ x \in E\} < +\infty$. Nous démontrons que P est quasi-compact sur $(\mathcal{B}_w, \|\cdot\|_w)$ si et seulement si, pour tout $f \in \mathcal{B}_w$, $(\frac{1}{n}\sum_{k=1}^{n} P^k f)_n$ contient une sous-suite convergeant dans $\mathcal{B}_w$ vers $\Pi f = \sum_{i=1}^{d} \mu_i(f) v_i$, où v_i est une fonction mesurable positive bornée sur E et μ_i une probabilité sur E. En particulier, quand le sous-espace de $\mathcal{B}_w$ constitué des fonctions P-invariantes est de dimension finie, la convergence uniforme des moyennes est équivalente à la convergence ponctuelle.

MSC: 37A30; 60J10

Keywords: Markov kernel; Quasi-compactness; Mean ergodicity; Geometrical ergodicity

1. Introduction

Let $(E, \mathcal{E})$ be a measurable space with countably generated σ-algebra, let $(\tilde{\mathcal{B}}, \|\cdot\|)$ denote the space of complex-valued bounded measurable functions on E, equipped with the supremum norm, and let P be a Markov kernel on $(E, \mathcal{E})$. Under some irreducibility conditions, P is quasi-compact on $\tilde{\mathcal{B}}$ if and only if P is mean ergodic with one-dimensional limit projection defined by the unique P-invariant distribution. This result was proved in [1] under the Harris condition (see also [11]), and in [8] under the ergodicity condition.[1] See also [6].

Now let $w : E \to [1, +\infty[$, and let $(\mathcal{B}_w, \|\cdot\|_w)$ denote the Banach space of complex-valued measurable functions on E satisfying $\|f\|_w := \sup\{w(x)^{-1}|f(x)|, x \in E\} < +\infty$. Assuming $Pw \le Cw$, with $C \in \mathbb{R}_+^*$, P acts continuously on $\mathcal{B}_w$. This work extends to $\mathcal{B}_w$ the equivalence between mean ergodicity with finite rank limit projection and quasi-compactness.

[1] The equivalence between mean ergodicity and quasi-compactness is not mentioned in [1], but it is an easy consequence of Theorem II.2 in [1]. In [8] $\mathcal{E}$ is not supposed to be countably generated.

Theorem 1. *P is quasi-compact on $\mathcal{B}_w$ if and only if there exist $d \in \mathbb{N}^*$, linearly independent non-negative functions $v_1, \ldots, v_d$ in $\tilde{\mathcal{B}}$, and P-invariant distributions $\mu_1, \ldots, \mu_d$ on E satisfying $\mu_i(w) < +\infty$ such that, for all $f \in \mathcal{B}_w$, the sequence $(\frac{1}{n}\sum_{k=1}^{n} P^k f)_n$ contains a subsequence converging in $\mathcal{B}_w$ to $\sum_{i=1}^{d} \mu_i(f)v_i$.*

Observe that the naive idea which consists in applying the similarity transformation $\tilde{P} : f \mapsto w^{-1}P(wf)$ in order to deduce the theorem from [1,8] does not work because $\tilde{P}$ is not Markovian when $\|Pw\|_w > 1$ (i.e. when w is not sub-invariant). The proof of Theorem 1 is actually based on a recent work of Hennion [3], which gives criteria for quasi-compactness of kernels acting on $\mathcal{B}_w$, on spectral theory [2], and on positive operator theory [12,13]. As in [3], the above theorem does not require any irreducibility or aperiodicity conditions; in this sense, when applied with $w = 1_E$, it improves [1,8]. This theorem shows too that a quasi-compact Markov kernel on $\mathcal{B}_w$ is necessarily power-bounded. This fact was already proved in [4] (Section IV.3), together with the equivalence between quasi-compactness and uniform ergodicity, which also follows from [9].

The above theorem does not hold when $\mathcal{B}_w$ is replaced with continuous function spaces. For instance, if E is a compact metric space and P is uniquely ergodic on the space $\mathcal{C}(E)$ of all complex-valued continuous functions on E, then P is mean ergodic [7], but in general P is not quasi-compact on $\mathcal{C}(E)$ (consider irrational rotations of the circle).[2]

We shall present in Section 3 (Corollary 1) a direct application to w-geometrically ergodic Markov chains [10] whose transition probability is, by definition, quasi-compact on $\mathcal{B}_w$, with $\lambda = 1$ as a simple eigenvalue and the unique peripheral eigenvalue. Many examples of such Markov chains, with unbounded functions w, are presented in [10].

A simple example is provided by the linear model $X_n = \alpha X_{n-1} + \varepsilon_n$, with $\alpha \in]-1, 1[$, where $(\varepsilon_n)_{n\geq 1}$ is an i.i.d. sequence of real-valued random variables, independent of X_0, such that $m = \mathbb{E}[|\varepsilon_1|] < +\infty$. In this case the state space is $E = \mathbb{R}$ with its Lebesgue sets, and $P(x, A) = \mathbb{E}[1_A(\alpha x + \varepsilon_1)]$, which yields $Pf(x) = \mathbb{E}[f(\alpha x + \varepsilon_1)]$. Let $w(y) = 1 + |y|$ ($y \in \mathbb{R}$). Then, for any $x \in \mathbb{R}$, we have $Pw(x) = \mathbb{E}[w(\alpha x + \varepsilon_1)] \leq 1 + |\alpha||x| + m$, so $Pw \leq |\alpha|w + L$, with $L = 1 - |\alpha| + m$. From this inequality, called drift condition, one can deduce that, if ε_1 has an everywhere positive density, then $(X_n)_n$ is w-geometrically ergodic [10] (Section 15.5.2). Observe that w is not sub-invariant. Indeed, $Pw(0) = 1 + m > w(0)$, so $\|Pw\|_w > 1$. Obviously, this conclusion extends to any function $w(y) = a + b|y|$, with constants $a, b > 0$. Actually, in most of the examples of w-geometrically ergodic Markov chains, w is not sub-invariant when it is unbounded.

Finally we shall see in Corollary 2 that, in the special case of denumerable Markov chains, the above theorem enables us to obtain an elementary proof of the above mentioned well-known fact that geometric ergodicity is equivalent to some drift condition.

2. Proof of Theorem 1

Proof of $\Rightarrow$. Suppose P is quasi-compact on $\mathcal{B}_w$. It is proved in [4] (Section IV.3) that $(\frac{1}{n}\sum_{k=1}^{n} P^k)_n$ converges in the operator norm topology to a finite dimensional projection Π of the form: $\Pi f = \sum_{i=1}^{d} \phi_i(f) f_i$, where the f_i's are linearly independent functions in $\tilde{\mathcal{B}}$ and the ϕ_i's are bounded complex measures on E such that $|\phi_i|(w) < +\infty$, with $|\phi_i|$ the total variation of ϕ_i. It remains to prove that one can choose f_i and ϕ_i such that $f_i \geq 0$ and ϕ_i is a probability measure on E. Notice that $\Pi(\mathcal{B}_w) \subset \tilde{\mathcal{B}}$, $\Pi \geq 0$ and $\Pi 1_E = 1_E$.

Let $\mathcal{B}_{\mathbb{R}}$ be the subspace of $\mathcal{B}_w$ composed of real-valued functions. Then $\Pi(\mathcal{B}_{\mathbb{R}})$ is a Banach lattice which is isomorphic to $\mathbb{R}^d$ with the preservation of the order relation [13]. Consequently there exist non-negative functions $g_1, \ldots, g_d$ in $\Pi(\mathcal{B}_w)$ and positive linear form $e_1^*, \ldots, e_d^*$ on $\Pi(\mathcal{B}_w)$ such that $g = \sum_{i=1}^{d} e_i^*(g) g_i$ for all $g \in \Pi(\mathcal{B}_w)$. Let $\psi_j = e_j^* \circ \Pi$. The ψ_j's are positive continuous linear forms on $\mathcal{B}_w$, and $\psi_j = \sum_{i=1}^{d} e_j^*(f_i)\phi_i$. Thus the ψ_j's are positive bounded measures on E such that $\psi_j(w) < +\infty$. Set $\mu_j = \frac{1}{\psi_j(E)}\psi_j$ and $v_j = \psi_j(E) g_j$. Then $\Pi f = \sum_{i=1}^{d} \psi_i(f) g_i = \sum_{i=1}^{d} \mu_i(f) v_i$, and the μ_i's are P-invariant (use $\Pi P = \Pi$). □

[2] Also consider $E = [0, 1]$ and $Pf(x) = \frac{1}{2}[f(\frac{x}{2}) + f(\frac{x+1}{2})]$. P is quasi-compact on the space of Lipschitz functions on $[0, 1]$, so P is mean ergodic on the space of continuous functions on $[0, 1]$, but is not quasi-compact on this space: indeed, for $|z| < 1$, $f_z = \sum_{n\geq 1} z^{n-1}\cos(2^n\pi\cdot)$ is a continuous function satisfying $Pf_z = zf_z$.

Proof of $\Leftarrow$. We shall denote by (ME) the mean ergodicity (subsequential) condition of Theorem 1. We set $\Pi f = \sum_{i=1}^{d} \mu_i(f) v_i$. If T is a continuous linear operator on $\mathcal{B}_w$, we denote by $\|T\|_w$ its operator norm, and by $r(T)$ its spectral radius. We denote by I the identity operator on $\mathcal{B}_w$. Given $a \in \mathbb{C}$ and $\rho > 0$, we set $D(a,\rho) = \{z\colon z \in \mathbb{C}, |z-a| \le \rho\}$.

Since $P1_E = 1_E$, we have $r(P) \ge 1$. Besides, there exists $n_k \nearrow +\infty$ such that $\sup_k \|n_k^{-1} \sum_{j=1}^{n_k} P^j w\|_w < +\infty$, thus $\sup_k n_k^{-1} \|P^{n_k} w\|_w < +\infty$. Since $\|P^n\|_w = \|P^n w\|_w$, one gets $r(P) = \lim_n \|P^n\|_w^{1/n} = 1$. In particular this yields $\sum_{n \ge 0} 2^{-(n+1)} \|P^n\|_w < +\infty$, so we can define the following bounded operator on $\mathcal{B}_w$, which is obviously Markovian:

$$Q = \sum_{n \ge 0} 2^{-(n+1)} P^n = (2I - P)^{-1}.$$

Proposition 1. *Q is quasi-compact on $\mathcal{B}_w$.*

Proof. Let $\nu = \frac{1}{d} \sum_{i=1}^{d} \mu_i$. Since the σ-algebra $\mathcal{E}$ is countably generated, there exist a non-negative measurable function α on $(E \times E, \mathcal{E} \otimes \mathcal{E})$ and a positive kernel S on E such that we have $Q(x, \mathrm{d}y) = \alpha(x,y)\,\mathrm{d}\nu(y) + S(x, \mathrm{d}y)$, with $S(x,\cdot) \perp \nu$, for each $x \in E$ [11]. For $p \in \mathbb{N}^*$, set $\alpha_p = \min\{\alpha, p\}$, and

$$T_p(x,\mathrm{d}y) = \alpha_p(x,y)\,\mathrm{d}\nu(y), \qquad S_p(x,\mathrm{d}y) = Q(x,\mathrm{d}y) - T_p(x,\mathrm{d}y).$$

If $f \in \mathcal{B}_w$, then $|T_p f| \le \|f\|_w T_p w \le p\nu(w)\|f\|_w$, so $T_p(\mathcal{B}_w) \subset \tilde{\mathcal{B}}$. Besides T_p acts continuously on $\mathcal{B}_w$, and so is S_p. In order to apply [3], observe that, for each $p \in \mathbb{N}^*$, the functions $\alpha_p^{(w)}(x,\cdot) = w(x)^{-1}\alpha_p(x,\cdot)w(\cdot)$, $x \in E$, are uniformly ν-integrable (use $\alpha_p^{(w)}(x,y) \le p w(y)$, $\nu(w) < +\infty$ and Lebesgue's theorem).

Finally, since $Q = \phi(P)$ with $\phi(z) = \sum_{n \ge 0} 2^{-(n+1)} z^n$ and ϕ is analytic on $D(0, \frac{3}{2})$, the spectral mapping theorem [2] yields $r(Q) = \phi(r(P)) = \phi(1) = 1$. Proposition 1 then follows from [3] and [4] (Section IV) via the following lemma. □

Lemma 1. *There exists $p \ge 1$ such that $r(S_p) < 1$.*

Proof. Suppose that $r(S_p) = 1$ for all $p \ge 1$. Since $S_p \ge 0$, there exists a positive continuous linear form, η_p, on $\mathcal{B}_w$ such that $\eta_p = \eta_p \circ S_p$ and $\eta_p(w) = 1$, see [12], p. 267. Let $\tilde{P}, \tilde{Q}, \tilde{T}_p, \tilde{S}_p, \tilde{\eta}_p$ be the restriction to $\tilde{\mathcal{B}}$ of P, Q, T_p, S_p, η_p. Since $\eta_p = \eta_p \circ S_p \le \eta_p \circ Q$ and $(\eta_p \circ Q - \eta_p)(1_E) = 0$, we have $\tilde{\eta}_p = \tilde{\eta}_p \circ \tilde{Q}$, thus $\tilde{\eta}_p \circ \tilde{P} = \tilde{\eta}_p$. Moreover we have:

(a) $\tilde{\eta}_p \ne 0$. Indeed, if $\tilde{\eta}_p = 0$, then, from $\eta_p \circ Q = \eta_p \circ T_p + \eta_p \circ S_p$ and $T_p(\mathcal{B}_w) \subset \tilde{\mathcal{B}}$, one would get $\eta_p \circ Q = \eta_p \circ S_p = \eta_p$, thus $\eta_p \circ P = \eta_p$. Then, by (ME), $\eta_p = \sum_{i=1}^{d} \eta_p(v_i)\mu_i$ would be a positive measure on E such that $\eta_p(\tilde{\mathcal{B}}) = \{0\}$, so $\eta_p = 0$, which is impossible.

(b) $\forall f \in \tilde{\mathcal{B}}$, $\eta_p(f) = \sum_{i=1}^{d} \eta_p(v_i)\mu_i(f)$. This follows from $\tilde{\eta}_p \circ \tilde{P} = \tilde{\eta}_p$ and (ME).

Now, from (a) and (b), there exist $j \in \{1,\dots,d\}$ and $p_k \nearrow +\infty$ such that we have $\eta_{p_k}(v_j) \ne 0$. Besides $\eta_{p_k}(v_j)\mu_j(T_{p_k}1_E) \le \eta_{p_k}(T_{p_k}1_E) = \eta_{p_k}(Q1_E - S_{p_k}1_E) = 0$, thus $\mu_j(T_{p_k}1_E) = 0$. When $k \to +\infty$, this gives $\iint \alpha(x,y)\,\mathrm{d}\nu(y)\,\mathrm{d}\mu_j(x) = 0$, hence $\int \alpha(x_0,y)\,\mathrm{d}\nu(y) = 0$ for a $x_0 \in E$. So $Q(x_0,\cdot) = S(x_0,\cdot) \perp \nu$: there exists $A \in \mathcal{E}$ such that $Q(x_0, A) = 0$ and $\nu(A) = 1$.

But: $Q(x_0,A) = 0 \Rightarrow \forall n \ge 1, P^n 1_A(x_0) = 0 \Rightarrow \sum_{i=1}^{d} \mu_i(A) v_i(x_0) = 0$ (by condition (ME)). While: $\nu(A) = \frac{1}{d}\sum_{i=1}^{d} \mu_i(A) = 1 \Rightarrow \mu_i(A) = 1, i = 1,\dots,d$.

Thus $\sum_{i=1}^{d} v_i(x_0) = 0$: this is impossible because (ME) gives $1_E = \sum_{i=1}^{d} v_i$. □

We shall denote by $\sigma(Q)$ and $\sigma(P)$ the spectrum of Q and P when acting on $\mathcal{B}_w$.

Lemma 2. *We have $\sigma(Q) \setminus \{1\} \subset D(\frac{2}{3}, \frac{1}{3}) \cap D(0, 1-\varepsilon)$ for a certain $\varepsilon \in]0,1[$.*

Proof. We have $Q = \phi(P)$ with $\phi(z) = \frac{1}{2-z}$, thus $\sigma(Q) = \phi(\sigma(P))$ [2]. Since $r(P) = 1$, we get $\sigma(Q) \subset \phi(D(0,1)) = D(\frac{2}{3}, \frac{1}{3})$. So $\lambda = 1$ is the unique peripheral spectral value of Q, and Lemma 2 then follows from Proposition 1. □

Lemma 3. $\lambda = 1$ *is a first order pole for* P, *with a corresponding finite-rank residue.*

Proof. Set $\psi(z) = 2 - \frac{1}{z}$, $z \in \mathbb{C}^*$. Lemma 2 yields $0 \notin \sigma(Q)$, so Q is invertible on $\mathcal{B}_w$, ψ is analytic on a neighborhood of $\sigma(Q)$, and $P = 2I - Q^{-1} = \psi(Q)$. Thus $\sigma(P) = \psi(\sigma(Q))$, and $\sigma(P) \setminus \{1\} = \psi(\sigma(Q) \setminus \{1\}) \subset \psi(D(\frac{2}{3}, \frac{1}{3})) \cap \psi(D(0, 1-\varepsilon)) = D(0,1) \cap D(2, \frac{1}{1-\varepsilon})^c$. Thus $\lambda = 1$ is an isolated point in $\sigma(P)$. Let A_P and A_Q be the residue of the resolvent functions of P and Q at $\lambda = 1$. Let χ be an analytic function on a neighborhood of $\sigma(P)$ such that $\chi(V_0) = \{0\}$ and $\chi(V_1) = \{1\}$, where V_0 and V_1 are disjoint neighborhoods of the sets $\sigma(P) \setminus \{1\}$ and $\{1\}$, respectively. We know that $A_P = \chi(P)$ [2], thus $A_P = \chi(\psi(Q))$. Besides $W_0 = \psi^{-1}(V_0)$ and $W_1 = \psi^{-1}(V_1)$ are disjoint neighborhoods of respectively $\sigma(Q) \setminus \{1\}$ and $\{1\}$, and $\chi \circ \psi$ is an analytic function on $W_0 \cup W_1$ such that $\chi \circ \psi(W_0) = \{0\}$, $\chi \circ \psi(W_1) = \{1\}$. Thus $A_Q = \chi \circ \psi(Q)$, so $A_P = A_Q$. Since the Markov kernel Q is quasi-compact on $\mathcal{B}_w$ (Proposition 1) and Q is power-bounded [4] (Theorem IV.3(i)), $\lambda = 1$ is a first order pole for Q, and $A_Q(\mathcal{B}_w) = \mathrm{Ker}(Q - I)$ is finite-dimensional by [2] (Theorem VIII.8.3 and Corollary VIII.8.4). By the definition of Q as a series, $Pf = f$ implies $Qf = f$ ($f \in \mathcal{B}_w$), and the converse holds by using $P = 2I - Q^{-1}$. Finally $A_P(\mathcal{B}_w) = A_Q(\mathcal{B}_w) = \mathrm{Ker}(Q - I) = \mathrm{Ker}(P - I)$ is finite-dimensional, so $\lambda = 1$ is a first order pole for P (use the arguments of [2], Theorem VII.4.5). □

Lemma 4. $\{\lambda \in \sigma(P), |\lambda| = 1\}$ *is composed of a finite number of first order poles.*

Proof. From Lemma 3 and a classical result concerning the peripheral spectrum of positive operators on Banach lattice [13] (Theorem 5.5, p. 331), the set of peripheral spectral values of P is composed of a finite number of poles for P. Using the Laurent expansions, Lemma 3 implies that they are first order poles. □

Lemma 5. *For any peripheral pole* λ *of* P, *we have* $\dim \mathrm{Ker}(P - \lambda I) \le \dim \mathrm{Ker}(P - I) < +\infty$.

Proof. We have $\dim \mathrm{Ker}(P - I) < +\infty$ by (ME). Let $\lambda_1 = 1, \lambda_2, \ldots, \lambda_m$ be the peripheral poles of P. The previous results show that $\mathcal{B}_w = \mathrm{Ker}(P - I) \oplus F \oplus H$, where $F = \bigoplus_{i=2}^m \mathrm{Ker}(P - \lambda_i I)$, and H is a P-invariant closed subspace of $\mathcal{B}_w$ such that $r(P_{|H}) < 1$, with $P_{|H}$ the restriction of P to H. Thus $(\frac{1}{n} \sum_{k=1}^n P^k)_n$ converges in the operator norm topology to the projection onto $\mathrm{Ker}(P - I)$. Then Lemma 5 follows from [9] (Theorem 2). □

The quasi-compactness of P on $\mathcal{B}_w$ follows from Lemmas 4 and 5. □

3. Applications to geometrically ergodic Markov chains

Let $(X_n)_{n \ge 0}$ be a Markov chain with state space E and transition probability P. Recall that $(X_n)_{n \ge 0}$ is said to be w-geometrically ergodic if there exist an invariant distribution ν on E such that $\nu(w) < +\infty$, and some constants $r < 1$ and $D \in \mathbb{R}_+$ such that for every $f \in \mathcal{B}_w$ we have

$$\left\| P^n f - \nu(f) 1_E \right\|_w \le D r^n \|f\|_w.$$

Corollary 1. *Assume that* $(X_n)_{n \ge 0}$ *is an aperiodic positive Harris Markov chain with stationary distribution* ν. *Then* $(X_n)_{n \ge 0}$ *is* w-*geometrically ergodic if and only if one of the two next conditions holds*:

(a) $\forall f \in \mathcal{B}_w$, $P^n f \to \nu(f) 1_E$ *in* $\mathcal{B}_w$ *when* $n \to +\infty$.
(b) *For all* $f \in \mathcal{B}_w$, $(\frac{1}{n} \sum_{k=1}^n P^k f)_n$ *contains a subsequence converging in* $\mathcal{B}_w$ *to* $\nu(f) 1_E$.

Corollary 1 is an easy consequence of Theorem 1. (When (b) is assumed, the aperiodicity condition ensures that $\lambda = 1$ is the unique peripheral eigenvalue of P.)

The reader will find in [10] many examples of geometrically ergodic Markov chains. Geometric ergodicity with a bounded function w corresponds to an aperiodic Markov chain satisfying Doeblin's condition.

When w is unbounded and $(X_n)_{n\ge 0}$ is aperiodic and ψ-irreducible w.r.t. to some σ-finite positive measure ψ on E, w-geometric ergodicity is equivalent to the following drift condition [10] (Chapter 16): there exist $\rho < 1$, $L > 0$, and a petite set A in E such that $Pw_0 \le \rho w_0 + L1_A$, where w_0 is a function on E such that $d^{-1}w \le w_0 \le dw$ for some constant $d > 0$. Corollary 1 sheds new light on this fact, at least for countable Markov chains, and as an illustration, let us present a simple proof of the well-known next statement proved in [5].

Corollary 2. *Let $(X_n)_{n\ge 0}$ be an aperiodic and irreducible Markov chain with state space $E = \mathbb{N}$, and suppose $\lim_k w(k) = +\infty$. Then $(X_n)_{n\ge 0}$ is w-geometrically ergodic if and only if there exist $\rho < 1$ and $C > 0$ such that $P^n w \le C\rho^n w + C$ for all $n \ge 1$.*

By using the basic arguments of [10] (Section 16.1.1), one can easily see that the condition in Corollary 2 is equivalent to: $\exists \rho < 1, \exists L > 0$, $Pw_0 \le \rho w_0 + L$, with w_0 equivalent to w.

Proof of Corollary 2. If $(X_n)_{n\ge 0}$ is w-geometrically ergodic, then $P^n w \le Dr^n w + \nu(w)$. Conversely, suppose $P^n w \le C\rho^n w + C$ with $\rho < 1$, $C > 0$, independent of n. Then we have $\sup_{n\ge 1} \|P^n\|_w \le 2C$, and there exists an invariant distribution ν such that $\nu(w) < +\infty$.[3] Set $\Pi_n = \frac{1}{n}\sum_{k=1}^n P^k$, and let $\ell^1(\nu)$ be the space of $\mathbb{C}$-valued sequences $(x(n))_{n\in\mathbb{N}}$ such that $\sum_n \nu(n)|x(n)| < +\infty$. P is a contraction of $\ell^1(\nu)$, so for any $f \in \ell^1(\nu)$, $(\Pi_n f)_n$ converges in $\ell^1(\nu)$, use e.g. [2] (Section VIII.5). The limit $\alpha = \lim_n \Pi_n f$ is P-invariant, and by irreducibility, it is constant: $\forall i \in \mathbb{N},\ \alpha(i) = \nu(f)$. Thus $\lim_n \Pi_n f(i) = \nu(f)$ for all $i \in \mathbb{N}$.

Now let $f \in \mathcal{B}_w$, and for convenience assume $\|f\|_w = 1$ (i.e. $|f| \le w$). We have

$$\forall i \in \mathbb{N}, \quad \left|P^k f(i) - \nu(f)\right| \le P^k w(i) + \nu\big(|f|\big) \le C\rho^k w(i) + C + \nu(w).$$

Let $\varepsilon > 0$. Then there exist $i_0 \ge 1$, $N_0 \ge 1$ such that $w(i)^{-1}|P^k f(i) - \nu(f)| \le \varepsilon$ for all $i > i_0$ and $k > N_0$. By using the fact that $\sup_{k\ge 1} \|P^k w\|_w < +\infty$ and

$$\Pi_n f(i) - \nu(f) = \frac{1}{n}\sum_{k=0}^{N_0}\big(P^k f(i) - \nu(f)\big) + \frac{1}{n}\sum_{k=N_0+1}^{n}\big(P^k f(i) - \nu(f)\big),$$

we easily deduce that there exists $N_1 \ge N_0$ such that $w(i)^{-1}|\Pi_n f(i) - \nu(f)| \le 2\varepsilon$ for all $i > i_0$ and $n > N_1$. Finally let $N_2 \ge N_1$ be such that $w(i)^{-1}|\Pi_n f(i) - \nu(f)| \le 2\varepsilon$ for all $i = 0, \ldots, i_0$ and $n > N_2$. Then $\|\Pi_n f - \nu(f)\|_w \le 2\varepsilon$ for all $n > N_2$, and Corollary 1 then applies. □

Acknowledgments

The author is grateful to Michael Lin for his helpful advice about this work. Also, the author thanks the referee for many very helpful comments which allowed to greatly enhance the content and the presentation of this paper.

References

[1] A. Brunel and D. Revuz. Quelques applications probabilistes de la quasi-compacité. *Ann. Inst. H. Poincaré, Sect. B (N.S.)* **10** (1974) 301–337. MR0373008
[2] N. Dunford and J. T. Schwartz. *Linear Operators. Part. I: General Theory*. Wiley, New York, 1958. MR1009162
[3] H. Hennion. Quasi-compactness and absolutely continuous kernels. *Probab. Theory Related Fields*. **139** (2007) 451–471. MR2322704

[3]This is a classical fact: consider the distributions $\mu_n(A) = \frac{1}{n}\sum_{k=1}^n (P^k 1_A)(x_0)$ ($x_0 \in E$ is fixed). From $P^n w \le C\rho^n w + C$, we easily obtain $\sup_{n\ge 1} \mu_n(w) \le 2Cw(x_0) < +\infty$, so $(\mu_n)_n$ is tight (use $\lim_k w(k) = +\infty$), and one can select a subsequence converging to an invariant distribution ν such that $\nu(w) < +\infty$.

[4] H. Hennion. Quasi-compactness and absolutely continuous kernels. Applications to Markov chains (2006). Available at ArXiv:math.PR/0606680.
[5] A. Hordijk and F. M. Spieksma. On ergodicity and recurrence properties of a Markov chain with an application to an open Jackson network. *Adv. in Appl. Probab.* **24** (1992) 343–376. MR1167263
[6] S. Horowitz. Transition probabilities and contractions of L_∞. *Z. Wahrsch. Verw. Gebiete* **24** (1972) 263–274. MR0331516
[7] U. Krengel. *Ergodic Theorems*. de Gruyter Studies in Mathematics, de Gruyter, Berlin, 1985. MR0797411
[8] M. Lin. Quasi-compactness and uniform ergodicity of Markov operators. *Ann. Inst. H. Poincaré, Sect. B (N.S.)* **11** (1975) 345–354. MR0402007
[9] M. Lin. Quasi-compactness and uniform ergodicity of positive operators. *Israel J. Math.* **29** (1978) 309–311. MR0493502
[10] S. P. Meyn and R. L. Tweedie. *Markov Chains and Stochastic Stability*. Springer, London, 1993. MR1287609
[11] D. Revuz. *Markov Chains*. North-Holland, Amsterdam, 1975. MR0758799
[12] H. H. Schaefer. *Topological Vector Spaces*. Springer, New York, 1971. MR0342978
[13] H. H. Schaefer. *Banach Lattices and Positive Operators*. Springer, New York, 1974. MR0423039