



HAL
open science

Generalized eigenfunctions of Markov kernels and application to the convergence rate of discrete random walks

Denis Guibourg, Loïc Hervé, James Ledoux

► **To cite this version:**

Denis Guibourg, Loïc Hervé, James Ledoux. Generalized eigenfunctions of Markov kernels and application to the convergence rate of discrete random walks. 2012. hal-00705523v1

HAL Id: hal-00705523

<https://hal.science/hal-00705523v1>

Preprint submitted on 7 Jun 2012 (v1), last revised 5 Dec 2013 (v4)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Generalized eigenfunctions of Markov kernels and application to the convergence rate of discrete random walks

Denis GUIBOURG, Loïc HERVÉ, and James LEDOUX *

June 6, 2012

Abstract

Let $(X_n)_{n \in \mathbb{N}}$ be a Markov chain on a measurable space \mathbb{X} with transition kernel P and let $V : \mathbb{X} \rightarrow [1, +\infty)$. Under a weak drift condition, the size of generalized eigenfunctions of P is estimated, where P is here considered as a linear bounded operator on the weighted-supremum space \mathcal{B}_V associated with V . Then combining this result and quasi-compactness arguments enables us to derive upper bounds for the geometric rate of convergence of $(X_n)_{n \in \mathbb{N}}$ to its invariant probability measure in operator norm on \mathcal{B}_V . Applications to discrete Markov random walks are presented.

AMS subject classification : 60J10; 47B07

Keywords : Geometric ergodicity, quasi-compactness, Drift condition, Birth-and -Death Markov chains.

1 Introduction

Let $(\mathbb{X}, \mathcal{X})$ be a measurable space with a σ -field \mathcal{X} , and let $(X_n)_{n \geq 0}$ be a Markov chain with state space \mathbb{X} and transition kernels $\{P(x, \cdot) : x \in \mathbb{X}\}$. Let $V : \mathbb{X} \rightarrow [1, +\infty)$. Assume that $(X_n)_{n \geq 0}$ has an invariant probability measure π such that $\pi(V) := \int_{\mathbb{X}} V(x) \pi(dx) < \infty$. This paper is based on the connection between spectral properties of the Markov kernel P and the so-called V -geometric ergodicity [MT93] that is the following convergence property for some constants $c_\rho > 0$ and $\rho \in (0, 1)$:

$$\sup_{|f| \leq V} \sup_{x \in \mathbb{X}} \frac{|\mathbb{E}[f(X_n) | X_0 = x] - \pi(f)|}{V(x)} \leq c_\rho \rho^n. \quad (1)$$

Let us introduce the weighted-supremum Banach space $(\mathcal{B}_V, \|\cdot\|_V)$ composed of measurable functions $f : \mathbb{X} \rightarrow \mathbb{C}$ such that

$$\|f\|_V := \sup_{x \in \mathbb{X}} \frac{|f(x)|}{V(x)} < \infty.$$

*INSA de Rennes, IRMAR, F-35042, France; CNRS, UMR 6625, Rennes, F-35708, France; Université Européenne de Bretagne, France. {Denis.Guibourg,Loic.Herve,James.Ledoux}@insa-rennes.fr

Then (1) reads as $\|P^n f - \pi(f)1_{\mathbb{X}}\|_V \leq c_\rho \rho^n$ for any $f \in \mathcal{B}_V$ such that $\|f\|_V \leq 1$, and there is a great interest in obtaining upper bounds for *the convergence rate* $\rho_V(P)$ defined by

$$\rho_V(P) := \inf \left\{ \rho \in (0, 1), \sup_{\|f\|_V \leq 1} \|P^n f - \pi(f)1_{\mathbb{X}}\|_V = O(\rho^n) \right\}. \quad (2)$$

For irreducible and aperiodic discrete Markov chains, criteria for the V -geometrical ergodicity are well-known from the literature using either the equivalence between geometric ergodicity and V -geometric ergodicity of \mathbb{N} -valued Markov chains [HS92, Prop. 2.4], or the strong drift condition. For instance, when $\mathbb{X} := \mathbb{N}$ (with $V(n) \rightarrow +\infty$ as $n \rightarrow +\infty$), the strong drift condition is

$$PV \leq \varrho V + b 1_{\{0,1,\dots,n_0\}}$$

for some $\varrho < 1, b < \infty$ and $n_0 \in \mathbb{N}$ (see [MT93]). Estimating $\rho_V(P)$ from the parameters ϱ, b, n_0 is a difficult issue. This often leads to unsatisfactory bounds, except for stochastically monotone P (see [MT94, LT96, Bax05] and the references therein).

This work presents a new procedure to study the convergence rate $\rho_V(P)$ under the following weak drift condition

$$\exists N \in \mathbb{N}^*, \exists d \in (0, +\infty), \exists \delta \in (0, 1), \quad P^N V \leq \delta^N V + d 1_{\mathbb{X}}. \quad (\mathbf{WD})$$

The V -geometrical ergodicity clearly implies **(WD)**. Conversely, such a condition with $N = 1$ was introduced in [MT93, Lem. 15.2.8] as an alternative to the drift condition [MT93, (V4)] to obtain the V -geometrical ergodicity under suitable assumption on V . Note that, under Condition **(WD)**, the following real number $\delta_V(P)$ is well defined:

$$\delta_V(P) := \inf \left\{ \delta \in [0, 1) : \exists N \in \mathbb{N}^*, \exists d \in (0, +\infty), P^N V \leq \delta^N V + d 1_{\mathbb{X}} \right\}.$$

Our procedure is based on quasi-compactness arguments and on the next statement (Theorem 1) which is the main result of the paper. Under Condition **(WD)**, for any $\lambda \in \mathbb{C}$ such that $\delta \leq |\lambda| \leq 1$, and for any $p \in \mathbb{N}^*$, the following property holds with $\beta(\lambda) := \ln |\lambda| / \ln \delta$:

$$f \in \mathcal{B}_V \cap \text{Ker}(P - \lambda I)^p \Rightarrow \exists c \in (0, +\infty), |f| \leq c (\ln V)^{\frac{p(p-1)}{2}} V^{\beta(\lambda)}. \quad (3)$$

In particular, if λ is an eigenvalue such that $|\lambda| = 1$, then any associated eigen-function f is bounded on \mathbb{X} . By contrast, if $|\lambda|$ is close to $\delta_V(P)$, then $|f| \leq c V^{\beta(\lambda)}$ with $\beta(\lambda)$ close to 1.

When the Markov kernel P has an invariant probability distribution, the connection between the V -geometric ergodicity and the quasi-compactness of P is recalled in Subsection 2.1 (Proposition 1). Namely, P is V -geometrically ergodic if and only if P is a power-bounded quasi-compact operator on \mathcal{B}_V for which $\lambda = 1$ is a simple eigenvalue and the unique eigenvalue of modulus one. In this case, if $r_{ess}(P)$ denotes the essential spectral radius of P on \mathcal{B}_V (see (7)) and if \mathcal{V} denotes the set of eigenvalues λ of P such that $r_{ess}(P) < |\lambda| < 1$, then the convergence rate $\rho_V(P)$ is given by:

$$\rho_V(P) = r_{ess}(P) \text{ if } \mathcal{V} = \emptyset \quad \text{and} \quad \rho_V(P) = \max\{|\lambda|, \lambda \in \mathcal{V}\} \text{ if } \mathcal{V} \neq \emptyset. \quad (4)$$

Property (4) is relevant to study the convergence rate $\rho_V(P)$ provided that, first an accurate bound of $r_{ess}(P)$ is known, second the above set \mathcal{V} is available. Bounds of $r_{ess}(P)$

related to drift conditions can be found in [Wu04] and [HL12] under various assumptions (see Subsection 2.1). In view of our applications, let us just recall that $r_{ess}(P) = \delta_V(P)$ in case $\mathbb{X} := \mathbb{N}$ and $\lim_n V(n) = +\infty$. However, even if the state space is discrete, finding the above set \mathcal{V} is difficult. Our main result (3) may be of great interest to solve this issue. In Section 3, this is illustrated in the case $\mathbb{X} := \mathbb{N}$, in particular, for birth-and-death Markov kernel P defined by $P(0,0) := a$ and $P(0,1) := 1 - a$ for some $a \in (0,1)$ and by

$$\forall n \geq 1, P(n, n-1) := p, \quad P(n, n) := r \in [0,1), \quad P(n, n+1) := q, \quad (5)$$

where $p, q, r \in [0,1]$ are such that $p + r + q = 1$, $p > q > 0$. Setting $\hat{\gamma} := \sqrt{p/q}$ and assuming that $\sum_{n \geq 0} P(0,n) (\hat{\gamma})^n < \infty$, the convergence rate $\rho_V(P)$ with respect to $V := (\hat{\gamma}^n)_{n \in \mathbb{N}}$ is computed in Subsection 3.2. When $r := 0$, such a result has been obtained for $a < p$ in [RT99] and [Bax05, Ex. 8.4] using Kendall's theorem, and for $a \geq p$ in [LT96] using the stochastic monotony of P . Our method gives a unified and simpler computation of $\rho_V(P)$ which moreover encompasses the case $r \neq 0$. Another example concerning a random walk on $\mathbb{X} := \mathbb{N}$ with unbounded increments is investigated in Subsection 3.3.

2 Quasi-compactness on \mathcal{B}_V and V -geometric ergodicity

We assume that P satisfies **(WD)**. Then P continuously acts on \mathcal{B}_V , and iterating **(WD)** shows that P is power-bounded on \mathcal{B}_V , namely $\sup_{n \geq 1} \|P^n\|_V < \infty$, where $\|\cdot\|_V$ also stands for the operator norm on \mathcal{B}_V . Thus we have $r(P) := \lim_n \|P^n\|_V^{1/n} = 1$ since P is Markov.

2.1 From quasi-compactness on \mathcal{B}_V to V -geometrical ergodicity

Let I denote the identity operator on \mathcal{B}_V . Recall that P is said to be quasi-compact on \mathcal{B}_V if there exist $r_0 \in (0,1)$ and $m \in \mathbb{N}^*$, $\lambda_i \in \mathbb{C}$, $p_i \in \mathbb{N}^*$ ($i = 1, \dots, m$) such that:

$$\mathcal{B}_V = \bigoplus_{i=1}^m \text{Ker}(P - \lambda_i I)^{p_i} \oplus H, \quad (6a)$$

where the λ_i 's are such that

$$|\lambda_i| \geq r_0 \quad \text{and} \quad 1 \leq \dim \text{Ker}(P - \lambda_i I)^{p_i} < \infty, \quad (6b)$$

and H is a closed P -invariant subspace such that

$$\sup_{h \in H, \|h\|_V \leq 1} \|P^n h\|_V = O(r_0^n). \quad (6c)$$

Concerning the essential spectral radius of P , denoted by $r_{ess}(P)$, here it is enough to have in mind that, if P is quasi-compact on \mathcal{B}_V , then we have (see for instance [Hen93])

$$r_{ess}(P) := \inf \{r_0 \in (0,1) \text{ such that (6a) (6b) (6c) hold}\}. \quad (7)$$

As mentioned in Introduction, the essential spectral radius of Markov kernels acting on \mathcal{B}_V is studied in [Wu04, HL12]. For instance, under Condition **(WD)**, the following result is

proved in [HL12, Th. 1]: if P^ℓ is compact from \mathcal{B}_0 to \mathcal{B}_V for some $\ell \geq 1$, where $(\mathcal{B}_0, \|\cdot\|_0)$ is the Banach space composed of bounded measurable functions $f : \mathbb{X} \rightarrow \mathbb{C}$ equipped with the supremum norm $\|f\|_0 := \sup_{x \in \mathbb{X}} |f(x)|$, then P is quasi-compact on \mathcal{B}_V with

$$r_{ess}(P) \leq \delta_V(P).$$

Moreover, equality $r_{ess}(P) = \delta_V(P)$ holds in many situations, in particular in the discrete state case with $V(n) \rightarrow \infty$ (see Proposition 2).

Next we recall a result which makes explicit the relationship between the quasi-compactness of P and the V -geometric ergodicity of the Markov chain $(X_n)_{n \in \mathbb{N}}$ with transition kernel P . Moreover, we provide an explicit formula for $\rho_V(P)$ in terms of the spectral elements of P . Note that for any $r_0 \in (r_{ess}(P), 1)$, the set of all the eigenvalues of λ of P such that $r_0 \leq |\lambda| \leq 1$ is finite (use (7)).

Proposition 1 *Let P be a transition kernel which has an invariant probability measure π such that $\pi(V) < \infty$. The two following assertions are equivalent:*

- (a) *P is V -geometrically ergodic.*
- (b) *P is a power-bounded quasi-compact operator on \mathcal{B}_V , for which $\lambda = 1$ is a simple eigenvalue (i.e. $\text{Ker}(P - I) = \mathbb{C} \cdot 1_{\mathbb{X}}$) and the unique eigenvalue of modulus one.*

Under any of these conditions, we have $\rho_V(P) \geq r_{ess}(P)$. In fact, for $r_0 \in (r_{ess}(P), 1)$, denoting the set of all the eigenvalues λ of P such that $r_0 \leq |\lambda| < 1$ by \mathcal{V}_{r_0} , we have:

- *either $\rho_V(P) \leq r_0$ when $\mathcal{V}_{r_0} = \emptyset$,*
- *or $\rho_V(P) = \max\{|\lambda|, \lambda \in \mathcal{V}_{r_0}\}$ when $\mathcal{V}_{r_0} \neq \emptyset$.*

Moreover, if $\mathcal{V}_{r_0} = \emptyset$ for all $r_0 \in (r_{ess}(P), 1)$, then $\rho_V(P) = r_{ess}(P)$.

The V -geometrical ergodicity of P obviously implies that P is quasi-compact on \mathcal{B}_V with $\rho_V(P) \geq r_{ess}(P)$ (see e.g. [KM03]). This follows from (7) using $H := \{f \in \mathcal{B}_V : \pi(f) = 0\}$ in (6a)-(6c). The property that P has a spectral gap on \mathcal{B}_V in the recent paper [KM11] corresponds here to the quasi-compactness of P (which is a classical terminology in spectral theory). The spectral gap in [KM11] corresponds to the value $1 - \rho_V(P)$. Then, [KM11, Prop. 1.1]) is a reformulation, under ψ -irreducibility and aperiodicity assumptions, of the equivalence of properties (a) and (b) in Proposition 1 (see also [KM11, Lem. 2.1]). Details on the proof of Proposition 1 are provided in [GHL11]. For general quasi-compact Markov kernels on \mathcal{B}_V , the result [Wu04, Th. 4.6] also provides interesting additional material on peripheral eigen-elements.

2.2 Size of generalized eigenfunctions of P

Theorem 1 *Assume that the weak drift condition (WD) holds true. If $\lambda \in \mathbb{C}$ is such that $\delta \leq |\lambda| \leq 1$, with δ given in (WD), and if $f \in \mathcal{B}_V \cap \text{Ker}(P - \lambda I)^p$ for some $p \in \mathbb{N}^*$, then there exists $c \in (0, +\infty)$ such that*

$$|f| \leq c V^{\frac{\ln|\lambda|}{\ln\delta}} (\ln V)^{\frac{p(p-1)}{2}}.$$

The proof of Theorem 1 is based on the following lemma.

Lemma 1 *Let $\lambda \in \mathbb{C}$ be such that $\delta \leq |\lambda| \leq 1$. Then*

$$\forall f \in \mathcal{B}_V, \exists c \in (0, +\infty), \forall x \in \mathbb{X}, \quad |\lambda|^{-n(x)} |(P^{n(x)} f)(x)| \leq c V(x)^{\frac{\ln |\lambda|}{\ln \delta}} \quad (8)$$

with, for any $x \in \mathbb{X}$, $n(x) := \lfloor \frac{-\ln V(x)}{\ln \delta} \rfloor$, where $\lfloor \cdot \rfloor$ denotes the integer part function.

Proof. First note that the iteration of **(WD)** gives

$$\forall k \geq 1, \quad P^{kN} V \leq \delta^{kN} V + d \left(\sum_{j=0}^{k-1} \delta^{jN} \right) 1_{\mathbb{X}} \leq \delta^{kN} V + \frac{d}{1 - \delta^N} 1_{\mathbb{X}}. \quad (9)$$

Let $g \in \mathcal{B}_V$ and $x \in \mathbb{X}$. Using (9), the positivity of P and $|g| \leq \|g\|_V V$, we obtain with $b := d/(1 - \delta^N)$:

$$\forall k \geq 1, \quad |(P^{kN} g)(x)| \leq (P^{kN} |g|)(x) \leq \|g\|_V (P^{kN} V)(x) \leq \|g\|_V (\delta^{kN} V(x) + b). \quad (10)$$

The previous inequality is also fulfilled with $k = 0$. Next, let $f \in \mathcal{B}_V$ and $n \in \mathbb{N}$. Writing $n = kN + r$, with $k \in \mathbb{N}$ and $r \in \{0, 1, \dots, N-1\}$, and applying (10) to $g := P^r f$, we obtain with $\xi := \max_{0 \leq \ell \leq N-1} \|P^\ell f\|_V$ (use $P^n f = P^{kN} (P^r f)$):

$$|(P^n f)(x)| \leq \xi [\delta^{kN} V(x) + b] \leq \xi [\delta^{-r} (\delta^n V(x) + b)] \leq \xi \delta^{-N} (\delta^n V(x) + b). \quad (11)$$

Using the inequality

$$-\frac{\ln V(x)}{\ln \delta} - 1 \leq n(x) \leq -\frac{\ln V(x)}{\ln \delta}$$

and the fact that $\ln \delta \leq \ln |\lambda| \leq 0$, Inequality (11) with $n := n(x)$ gives:

$$\begin{aligned} |\lambda|^{-n(x)} |(P^{n(x)} f)(x)| &\leq \xi \delta^{-N} \left((\delta |\lambda|^{-1})^{n(x)} V(x) + b |\lambda|^{-n(x)} \right) \\ &= \xi \delta^{-N} \left(e^{n(x)(\ln \delta - \ln |\lambda|)} e^{\ln V(x)} + b e^{-n(x) \ln |\lambda|} \right) \\ &\leq \xi \delta^{-N} \left(e^{(\frac{\ln V(x)}{\ln \delta} + 1)(\ln |\lambda| - \ln \delta)} e^{\ln V(x)} + b e^{\frac{\ln V(x)}{\ln \delta} \ln |\lambda|} \right) \\ &= \xi \delta^{-N} \left(e^{\frac{\ln |\lambda|}{\ln \delta} \ln V(x)} e^{\ln |\lambda| - \ln \delta} + b V(x)^{\frac{\ln |\lambda|}{\ln \delta}} \right) \\ &= \xi \delta^{-N} (e^{\ln |\lambda| - \ln \delta} + b) V(x)^{\frac{\ln |\lambda|}{\ln \delta}}. \end{aligned}$$

This gives Inequality (8) with $c := \xi \delta^{-N} (e^{\ln |\lambda| - \ln \delta} + b)$. \square

Proof of Theorem 1. If $f \in \mathcal{B}_V \cap \text{Ker}(P - \lambda I)$, then $|\lambda|^{-n(x)} |(P^{n(x)} f)(x)| = |f(x)|$, so that (8) gives the expected conclusion when $p = 1$. Next, let us proceed by induction. Assume that the conclusion of Theorem 1 holds for some $p \geq 1$. Let $f \in \mathcal{B}_V \cap \text{Ker}(P - \lambda I)^{p+1}$. We can write

$$P^n f = (P - \lambda I + \lambda I)^n f = \lambda^n f + \sum_{k=1}^{\min(n,p)} \binom{n}{k} \lambda^{n-k} (P - \lambda I)^k f. \quad (12)$$

For $k \in \{1, \dots, p\}$, we have $f_k := (P - \lambda I)^k f \in \text{Ker}(P - \lambda I)^{p+1-k} \subset \text{Ker}(P - \lambda I)^p$, thus we have from the induction hypothesis :

$$\exists c' \in (0, +\infty), \forall k \in \{1, \dots, p\}, \forall x \in \mathbb{X}, \quad |f_k(x)| \leq c' V(x)^{\frac{\ln|\lambda|}{\ln \delta}} (\ln V(x))^{\frac{p(p-1)}{2}}. \quad (13)$$

Now, we obtain from (12) (with $n := n(x)$), (13) and Lemma 1 that for all $x \in \mathbb{X}$:

$$\begin{aligned} |f(x)| &\leq |\lambda|^{-n(x)} |(P^{n(x)} f)(x)| + c' V(x)^{\frac{\ln|\lambda|}{\ln \delta}} (\ln V(x))^{\frac{p(p-1)}{2}} |\lambda|^{-\min(n,p)} \sum_{k=1}^{\min(n,p)} \binom{n(x)}{k} \\ &\leq c V(x)^{\frac{\ln|\lambda|}{\ln \delta}} + c_1 V(x)^{\frac{\ln|\lambda|}{\ln \delta}} (\ln V(x))^{\frac{p(p-1)}{2}} n(x)^p \\ &\leq c_2 V(x)^{\frac{\ln|\lambda|}{\ln \delta}} (\ln V(x))^{\frac{p(p-1)}{2} + p} \end{aligned}$$

with some constants $c_1, c_2 \in (0, +\infty)$ independent of x . This gives the expected result. \square

3 Applications to discrete Markov chains

In this section, the state space \mathbb{X} is discrete. For the sake of simplicity, we assume that $\mathbb{X} := \mathbb{N}$ throughout the section. Let $P = (P(i, j))_{i, j \in \mathbb{N}^2}$ be a Markov kernel on \mathbb{N} . The function $V : \mathbb{N} \rightarrow [1, +\infty)$ is assumed to satisfy

$$\lim_n V(n) = +\infty \quad \text{and} \quad \sup_{n \in \mathbb{N}} \frac{(PV)(n)}{V(n)} < \infty.$$

The first focus is on the estimation of $r_{ess}(P)$ from Condition **(WD)**.

Proposition 2 *The two following conditions are equivalent:*

(a) Condition **(WD)** holds with V ;

(b) $L := \inf_{N \geq 1} (\ell_N)^{\frac{1}{N}} < 1$ where $\ell_N := \limsup_{n \rightarrow +\infty} (P^N V)(n)/V(n)$.

In this case, P is power-bounded and quasi-compact on \mathcal{B}_V with $r_{ess}(P) = \delta_V(P) = L$.

The proof of the the equivalence (a) \Leftrightarrow (b), as well as the equality $\delta_V(P) = L$, is straightforward (see [GHL11, Cor. 4]). That P is quasi-compact on \mathcal{B}_V under **(WD)** in the discrete case, with $r_{ess}(P) \leq \delta_V(P)$, can be derived from [Wu04] or [HL12, Th. 1] (see Subsection 2.1 and use the fact that the injection from \mathcal{B}_0 to \mathcal{B}_V is compact when $\mathbb{X} := \mathbb{N}$). Equality $r_{ess}(P) = \delta_V(P)$ can be proved by combining the results [Wu04, HL12] (see [GHL11, Cor. 1] for details).

In the next subsections, Proposition 2 is applied to random walks (RW) on \mathbb{N} with the following special sequence $V_\gamma := (\gamma^n)_{n \in \mathbb{N}}$ for some $\gamma \in (1, +\infty)$. The associated weighted-supremum space $\mathcal{B}_\gamma \equiv \mathcal{B}_{V_\gamma}$ is defined by:

$$\mathcal{B}_\gamma := \left\{ (f(n))_{n \in \mathbb{N}} \in \mathbb{C}^{\mathbb{N}} : \sup_{n \in \mathbb{N}} \gamma^{-n} |f(n)| < \infty \right\}. \quad (14)$$

3.1 Quasi-compactness of RW with bounded state-dependent increments

Let us fix $b \in \mathbb{N}^*$, and assume that the kernel P satisfies the following conditions:

$$\begin{aligned} \forall i \in \{0, \dots, b-1\}, \quad \sum_{j \geq 0} P(i, j) &= 1; \\ \forall i \geq b, \forall j \in \mathbb{N}, \quad P(i, j) &= \begin{cases} 0 & \text{if } |i-j| > b \\ a_{j-i}(i) & \text{if } |i-j| \leq b \end{cases} \end{aligned} \quad (15)$$

where $(a_{-b}(i), \dots, a_b(i)) \in [0, 1]^{2b+1}$ satisfies $\sum_{k=-b}^b a_k(i) = 1$ for all $i \geq b$. This kind of kernels arises, for instance, from time-discretization of Markovian queueing models.

Set $\phi_n(\gamma) := \sum_{k=-b}^b a_k(n) \gamma^k$. We have $(PV_\gamma)(n) = \phi_n(\gamma) V_\gamma(n)$ for each $n \geq b$. Then the next statement follows from Proposition 2 using (16)-(17).

Proposition 3 *Assume that, for every $k \in \mathbb{Z}$ such that $|k| \leq b$, $\lim_n a_k(n) = a_k \in [0, 1]$, and that $\gamma \in (1, +\infty)$ is such that*

$$\phi(\gamma) := \sum_{k=-b}^b a_k \gamma^k < 1 \quad (16)$$

$$\forall i \in \{0, \dots, b-1\}, \quad \sum_{j \geq 0} P(i, j) \gamma^j < \infty. \quad (17)$$

Then P is power-bounded and quasi-compact on \mathcal{B}_γ with $r_{\text{ess}}(P) \leq \phi(\gamma)$.

Example 1 (State-dependent birth-and-death Markov chains) *When $b := 1$ in (15), we obtain the standard class of state-dependent birth-and-death Markov chains:*

$$\forall n \geq 1, \quad P(n, n-1) := p_n, \quad P(n, n) := r_n, \quad P(n, n+1) := q_n,$$

where $(p_n, r_n, q_n) \in [0, 1]^2$ and $p_n + r_n + q_n = 1$. Assume that the following limits exist:

$$\lim_n p_n := p \in (0, 1], \quad \lim_n r_n := r \in [0, 1), \quad \lim_n q_n := q.$$

If $\gamma \in (1, +\infty)$ is such that $\phi(\gamma) := p/\gamma + r + q\gamma < 1$ and $\sum_{n \geq 0} P(0, n) \gamma^n < \infty$, then it follows from Proposition 3 that $r_{\text{ess}}(P) \leq p/\gamma + r + q\gamma$. The conditions $\gamma > 1$ and $p/\gamma + r + q\gamma < 1$ are equivalent to the following ones (use $r = 1 - p - q$ for (i)):

(i) either $p > q > 0$ and $1 < \gamma < p/q$; (ii) or $q = 0$ and $\gamma > 1$.

(i) When $p > q > 0$ and $1 < \gamma < p/q$: if $\sum_{n \geq 0} P(0, n) \gamma^n < \infty$, then P is power-bounded and quasi-compact on \mathcal{B}_γ with $r_{\text{ess}}(P) \leq \phi(\gamma)$. Set $\hat{\gamma} := \sqrt{p/q}$. Then

$$\min_{\gamma > 1} \phi(\gamma) = \phi(\hat{\gamma}) = r + 2\sqrt{pq} \in (r^2, 1). \quad (18)$$

Consequently, if $\sum_{n \geq 0} P(0, n) (\hat{\gamma})^n < \infty$, then the previous conclusions holds for $\gamma := \hat{\gamma}$, with essential spectral radius on $\mathcal{B}_{\hat{\gamma}}$ satisfying $r_{\text{ess}}(P) \leq r + 2\sqrt{pq}$.

(ii) When $q := 0$ and $\gamma > 1$: if $\sum_{n \geq 0} P(0, n) \gamma^n < \infty$, then $r_{\text{ess}}(P) \leq \phi(\gamma) = p/\gamma + r$.

Remark 1 (Random walks with i.d. bounded increments) Consider the case when the increments $a_k(n)$ do not depend on the state n , that is when the kernel P is

$$\forall i \in \{0, \dots, b-1\}, \quad \sum_{j \geq 0} P(i, j) = 1; \quad \forall i \geq b, \forall j \in \mathbb{N}, \quad P(i, j) = \begin{cases} a_{j-i} & \text{if } |i-j| \leq b \\ 0 & \text{if } |i-j| > b \end{cases}$$

where $(a_{-b}, \dots, a_b) \in [0, 1]^{2b+1}$ and $\sum_{k=-b}^b a_k = 1$. Obviously the statements of Example 1 apply but some additional facts can be deduced for such Markov chains. First note that

$$\forall \gamma \in (1, +\infty), \quad \forall N \geq 1, \quad \forall n \geq Nb, \quad (P^N V_\gamma)(n) = \phi(\gamma)^N V_\gamma(n). \quad (19)$$

Consequently, under the assumptions (17) and $\phi(\gamma) < 1$ where $\phi(\cdot)$ is given by (16), we obtain from Proposition 2 that Condition **(WD)** is fulfilled with V_γ and

$$r_{\text{ess}}(P) = \delta_{V_\gamma}(P) = \phi(\gamma). \quad (20)$$

For the birth-and-death Markov chains, that is when $b := 1$, the convergence rate is computed in the next subsection using (20) and Theorem 1.

3.2 Study of the convergence rate for the birth-and-death Markov chains

Let $p, q, r \in [0, 1]$ be such that $p + r + q = 1$, $p > q > 0$, and let P be defined by

$$\begin{aligned} \forall n \geq 1, \quad P(n, n-1) &:= p, \quad P(n, n) := r \in [0, 1), \quad P(n, n+1) := q, \\ P(0, 0) &\in (0, 1), \quad \sum_{n \geq 0} P(0, n) (\hat{\gamma})^n < \infty \text{ where } \hat{\gamma} := \sqrt{\frac{p}{q}} \in (1, +\infty). \end{aligned} \quad (21)$$

Let $V_{\hat{\gamma}} := (\hat{\gamma}^n)_{n \in \mathbb{N}}$ and its associated weighted-supremum space $\mathcal{B}_{\hat{\gamma}} := \mathcal{B}_{V_{\hat{\gamma}}}$. It is well known that P is $V_{\hat{\gamma}}$ -geometrically ergodic. Here we compute the convergence rate $\rho_{V_{\hat{\gamma}}}(P)$ (see Proposition 4). First note that (18) (20) give

$$r_{\text{ess}}(P) = \delta_{V_{\hat{\gamma}}}(P) = r + 2\sqrt{pq}.$$

Second Theorem 1 allows us to prove the following.

Lemma 2 Assume that Conditions (21) hold true. If f is a nontrivial eigenvector in $\mathcal{B}_{\hat{\gamma}}$ associated with a complex eigenvalue λ of P such that $r + 2\sqrt{pq} < |\lambda| \leq 1$ then

$$\exists \alpha_1 \in \mathbb{C} \setminus \{0\}, \quad \forall n \geq 0, \quad f(n) = \alpha_1 z_\lambda^n, \quad (22)$$

with z_λ satisfying the following conditions:

$$|z_\lambda| < \hat{\gamma}, \quad (23a)$$

$$qz_\lambda^2 + (r - \lambda)z_\lambda + p = 0, \quad (23b)$$

$$\sum_{n \geq 0} P(0, n) z_\lambda^n = \lambda. \quad (23c)$$

Proof. Let $\lambda \in \mathbb{C}$, be such $r + 2\sqrt{pq} < |\lambda| \leq 1$ and $f \in \mathcal{B}_{\widehat{\gamma}}$, $f \neq 0$ satisfying $Pf = \lambda f$, that is

$$\forall n \geq 1, \quad \lambda f(n) = pf(n-1) + rf(n) + qf(n+1). \quad (24)$$

Let us denote by z_λ, z'_λ the two complex solutions of the characteristic equation

$$qz^2 + (r - \lambda)z + p = 0.$$

Observe that $z_\lambda z'_\lambda = p/q = \widehat{\gamma}^2$. Recall that the solutions of (24) are of the form, either $f(n) = \alpha_1 z_\lambda^n + \alpha_2 z'_\lambda^n$ if $z_\lambda \neq z'_\lambda$, or $f(n) = \alpha_1 z_\lambda^n + \alpha_2 n z_\lambda^n$ if $z_\lambda = z'_\lambda$, with $\alpha_1, \alpha_2 \in \mathbb{C}$. We have $|z_\lambda| \neq |z'_\lambda|$. Indeed, Theorem 1 applied with $p := 1$ and $\delta := r + 2\sqrt{pq}$ implies that $|f| \leq cV_{\widehat{\gamma}}^\tau$ with $\tau := \ln|\lambda|/\ln\delta \in (0, 1)$ and some constant c . Consequently we have $|\alpha_1 z_\lambda^n + \alpha_2 z'_\lambda^n| \leq c\widehat{\gamma}^{\tau n}$ in case $z_\lambda \neq z'_\lambda$, and $|\alpha_1 z_\lambda^n + \alpha_2 n z_\lambda^n| \leq c\widehat{\gamma}^{\tau n}$ in case $z_\lambda = z'_\lambda$. If $|z_\lambda| = |z'_\lambda|$, then we would have $|z_\lambda| = |z'_\lambda| = \widehat{\gamma}$, but the two previous inequalities then easily imply that $\alpha_1 = \alpha_2 = 0$, that is $f = 0$.

From $|z_\lambda| \neq |z'_\lambda|$, we can suppose that (for instance) $|z_\lambda| < \widehat{\gamma}$ and $|z'_\lambda| > \widehat{\gamma}$. Since $f, (z_\lambda^n)_{n \in \mathbb{N}}$ are in $\mathcal{B}_{\widehat{\gamma}}$ and $(z'_\lambda^n)_n$ is not in $\mathcal{B}_{\widehat{\gamma}}$, we obtain: $\forall n \geq 0, f(n) = \alpha_1 z_\lambda^n$. Since $f \neq 0$ (i.e. $\alpha_1 \neq 0$), the equation $(Pf)(0) = \lambda f(0)$ implies that z_λ must satisfy (23c). \square

Proposition 4 *In addition to Conditions (21), the boundary transition probabilities are assumed to satisfy, for some $a \in (0, 1)$:*

$$P(0, 0) := a, \quad P(0, 1) := 1 - a.$$

Then P is $V_{\widehat{\gamma}}$ -geometrically ergodic. Furthermore, defining $a_0 := 1 - q - \sqrt{pq}$, the convergence rate $\rho_{V_{\widehat{\gamma}}}(P)$ of P is given by:

- when $a \in [a_0, 1)$:

$$\rho_{V_{\widehat{\gamma}}}(P) = r + 2\sqrt{pq}; \quad (25)$$

- when $a \in (0, a_0]$:

(a) in case $2p \leq (1 - q + \sqrt{pq})^2$:

$$\rho_{V_{\widehat{\gamma}}}(P) = r + 2\sqrt{pq}; \quad (26)$$

(b) in case $2p > (1 - q + \sqrt{pq})^2$, setting $a_1 := p - \sqrt{pq} - \sqrt{r(r + 2\sqrt{pq})}$:

$$\rho_{V_{\widehat{\gamma}}}(P) = \left| a + \frac{p(1-a)}{a-1+q} \right| \quad \text{when } a \in (0, a_1] \quad (27a)$$

$$\rho_{V_{\widehat{\gamma}}}(P) = r + 2\sqrt{pq} \quad \text{when } a \in [a_1, a_0]. \quad (27b)$$

When $r := 0$, such results have been obtained in [RT99, Bax05, LT96] by using various methods involving conditions on a (see the end of Introduction). Let us specify the above formulas in case $r := 0$. We have $a_0 = a_1 = p - \sqrt{pq} = (p - q)/(1 + \sqrt{q/p})$, and it can be

easily checked that $2p > (1 - q + \sqrt{pq})^2$. The properties (25) (27a) (27b) then rewrite as: $\rho_{V_{\hat{\gamma}}}(P) = (pq + (a - p)^2)/|a - p|$ when $a \in (0, a_0]$, and $\rho_{V_{\hat{\gamma}}}(P) = 2\sqrt{pq}$ when $a \in (a_0, 1)$.

Proof of Proposition 4. By elimination, given some $\lambda \in \mathbb{C}$, a necessary and sufficient condition for the two following equations

$$qz^2 + (r - \lambda)z + p = 0, \quad (28a)$$

$$a + (1 - a)z = \lambda. \quad (28b)$$

to have a common solution $z \in \mathbb{C}$ is that

$$(1 - \lambda)[(\lambda - a)(1 - a - q) + p(1 - a)]. \quad (29)$$

Assume that $a \neq 1 - q$. Then $\lambda = 1$ is a solution of (29) and the other solution of (29), say $\lambda(a)$, and the associated complex number in (28b), say $z(a)$, are given by:

$$\lambda(a) := a + \frac{p(1 - a)}{a - 1 + q} \in \mathbb{R} \quad \text{and} \quad z(a) := \frac{p}{a + q - 1} \in \mathbb{R}. \quad (30)$$

Now, let $\lambda \in \mathbb{C}$ be such that $r + 2\sqrt{pq} < |\lambda| < 1$, and assume that there exists $f \in \mathcal{B}_{\hat{\gamma}}$, $f \neq 0$, such that $Pf = \lambda f$. Then Lemma 2 gives $f := (z\lambda^n)_{n \geq 0}$ (up to a multiplicative constant), with $z_\lambda \in \mathbb{C}$ satisfying $|z_\lambda| < \hat{\gamma}$ and Equations (28a)-(28b). Thus we have $\lambda = \lambda(a)$ and $z_\lambda = z(a)$, with $\lambda(a)$ and $z(a)$ given by (30). Conversely, we have $Pf_a = \lambda(a)f_a$ with $f_a = (z(a)^n)_{n \geq 0}$ since, by definition, $z(a)$ satisfies the equations (28a)-(28b) associated with $\lambda = \lambda(a)$. Now we must find the values $a \in (0, 1)$ for which we have $r + 2\sqrt{pq} < |\lambda(a)| < 1$ and $|z(a)| \leq \hat{\gamma}$. This is the relevant question since Proposition 1 gives the following properties:

- (i) if $r + 2\sqrt{pq} < |\lambda(a)| < 1$ and $|z(a)| < \hat{\gamma}$, then we have $\rho_{V_{\hat{\gamma}}}(P) = |\lambda(a)|$ since $\lambda(a)$ is the only eigenvalue λ of P on $\mathcal{B}_{\hat{\gamma}}$ such that $r + 2\sqrt{pq} < |\lambda| < 1$ (apply Proposition 1 with any r_0 such that $r + 2\sqrt{pq} < r_0 < |\lambda(a)|$),
- (ii) if $\lambda(a)$ or $z(a)$ do not satisfy the previous conditions, then we have $\rho_{V_{\hat{\gamma}}}(P) = r + 2\sqrt{pq}$ since there is no eigenvalue λ of P on $\mathcal{B}_{\hat{\gamma}}$ such that $r + 2\sqrt{pq} < |\lambda| < 1$ (apply Proposition 1 with any r_0 such that $r + 2\sqrt{pq} < r_0 < 1$).

Observe that

$$|z(a)| \leq \hat{\gamma} \Leftrightarrow |a - 1 + q| \geq \sqrt{pq}. \quad (31)$$

Hence, if $a \in (a_0, 1)$ (recall that $a_0 := 1 - q - \sqrt{pq}$), then $|z(a)| > \hat{\gamma}$. Then (ii) gives (25).

Now let $a \in (0, a_0]$. Then $|z(a)| \leq \hat{\gamma}$. Let us study $\lambda(a)$. We have $\lambda'(a) = 1 - pq/(a - 1 + q)^2$, so that $a \mapsto \lambda(a)$ is increasing on $(-\infty, a_0]$ from $-\infty$ to $\lambda(a_0) = r - 2\sqrt{pq}$. Thus

$$\forall a \in (0, a_0], \quad \lambda(a) \leq r - 2\sqrt{pq} < r + 2\sqrt{pq}.$$

and the equation $\lambda(a) = -(r + 2\sqrt{pq})$ has a unique solution $a_1 \in (-\infty, a_0)$. Note that $a_1 < a_0$ and $\lambda(a_1) = -(r + 2\sqrt{pq})$, that $\lambda(0) = p/(q - 1) \in [-1, 0)$ and finally that

$$\lambda(0) - \lambda(a_1) = p/(q - 1) + r + 2\sqrt{pq} = \frac{(q - \sqrt{pq} - 1)^2 - 2p}{1 - q}.$$

When $2p \leq (1 - q + \sqrt{pq})^2$, (26) follows from (ii). Indeed $|\lambda(a)| < r + 2\sqrt{pq}$ since

$$\forall a \in (0, a_0], \quad -(r + 2\sqrt{pq}) = \lambda(a_1) \leq \lambda(0) < \lambda(a) < r + 2\sqrt{pq}.$$

When $2p > (1 - q + \sqrt{pq})^2$, we have $a_1 \in (0, a_0]$ and:

- if $a \in (0, a_1)$, then (27a) follows from (i). Indeed $r + 2\sqrt{pq} < |\lambda(a)| < 1$ since

$$\forall a \in (0, a_1], \quad -1 \leq \lambda(0) < \lambda(a) < \lambda(a_1) = -(r + 2\sqrt{pq});$$

- if $a \in [a_1, a_0]$, then (27b) follows (ii). Indeed $|\lambda(a)| < r + 2\sqrt{pq}$ since

$$-(r + 2\sqrt{pq}) = \lambda(a_1) \leq \lambda(a) < r + 2\sqrt{pq}.$$

It remains to study the special case $a = 1 - q$. Then $\lambda = 1$ is the only solution of (29). Again let $\lambda \in \mathbb{C}$ be such that $r + 2\sqrt{pq} < |\lambda| < 1$, and let $f \in \mathcal{B}_{\tilde{\gamma}}$, $f \neq 0$, such that $Pf = \lambda f$. Then Lemma 2 gives $f := (z_\lambda^n)_{n \geq 0}$, with $z_\lambda \in \mathbb{C}$ satisfying Equations (28a)-(28b), thus Equation (29). Consequently there is no eigenvalue of P such that $r + 2\sqrt{pq} < |\lambda| < 1$. Proposition 1 applied with any $r_0 \in (r + 2\sqrt{pq}, 1)$ then gives $\rho_{V_{\tilde{\gamma}}}(P) = r + 2\sqrt{pq}$. \square

3.3 Convergence rate for a RW with unbounded increments

Proposition 2 and Theorem 1 may be also useful to estimate the convergence rate for RW on $\mathbb{X} := \mathbb{N}$ with unbounded increments. For instance, let P be defined by [MS95]

$$\forall n \geq 1, \quad P(0, n) := q_n, \quad \forall n \geq 1, \quad P(n, 0) := p, \quad P(n, n+1) := q = 1 - p,$$

with $p \in (0, 1)$ and $q_n \in [0, 1]$ such that $\sum_{n \geq 1} q_n = 1$.

Proposition 5 *Assume that $\gamma \in (1, 1/q)$ is such that $\sum_{n \geq 1} q_n \gamma^n < \infty$. Then $r_{ess}(P) \leq q\gamma$. Moreover P is V_γ -geometrically ergodic with convergence rate $\rho_{V_\gamma}(P) \leq \max(q\gamma, p)$.*

Proof. We have: $\forall n \geq 1, (PV_\gamma)(n) = q\gamma^{n+1} + p$. Thus, if $\gamma \in (1, 1/q)$ and $\sum_{n \geq 1} q_n \gamma^n < \infty$, then Condition **(WD)** holds with V_γ , and we have $\delta_{V_\gamma}(P) \leq q\gamma$. Therefore it follows from Proposition 2 that $r_{ess}(P) \leq q\gamma$. Now Proposition 1 is applied with any $r_0 > \max(q\gamma, p)$. Let $\lambda \in \mathbb{C}$ be such that $\max(q\gamma, p) < |\lambda| \leq 1$, and let $f \in \mathcal{B}_\gamma$, $f \neq 0$, be such that $Pf = \lambda f$. We obtain $f(n) = (\lambda/q)f(n-1) - pf(0)/q$ for any $n \geq 2$, so that

$$\forall n \geq 2, \quad f(n) = \left(\frac{\lambda}{q}\right)^{n-1} \left(f(1) - \frac{pf(0)}{\lambda - q}\right) + \frac{pf(0)}{\lambda - q}.$$

Since $f \in \mathcal{B}_\gamma$ and $|\lambda|/q > \gamma$, we obtain $f(1) = pf(0)/(\lambda - q)$, and consequently: $\forall n \geq 1, f(n) = pf(0)/(\lambda - q)$. Next the equality $\lambda f(0) = (Pf)(0) = \sum_{n \geq 1} q_n f(n)$ gives: $\lambda f(0) = pf(0)/(\lambda - q)$ since $\sum_{n \geq 1} q_n = 1$. We have $f(0) \neq 0$ since we look for $f \neq 0$. Thus λ satisfies $\lambda^2 - q\lambda - p = 0$, namely: $\lambda = 1$ or $\lambda = -p$. The case $\lambda = -p$ has not to be considered since $|\lambda| > \max(q\gamma, p)$. If $\lambda = 1$, then $f(n) = f(0)$ for each $n \in \mathbb{N}$, so that 1 is a simple eigenvalue. Thus $\lambda = 1$ is a simple eigenvalue of P on \mathcal{B}_γ and the only eigenvalue such that $\max(q\gamma, p) < |\lambda| \leq 1$. Then Proposition 1 gives the second conclusion of Proposition 5. \square

Note that p cannot be dropped in the inequality $\rho_{V_\gamma}(P) \leq \max(q\gamma, p)$ since $\lambda = -p$ is an eigenvalue of P on \mathcal{B}_γ with corresponding eigenvector $f_p := (1, -p, -p, \dots)$.

References

- [Bax05] P. H. Baxendale. Renewal theory and computable convergence rates for geometrically ergodic Markov chains. *Ann. Appl. Probab.*, 15(1B):700–738, 2005.
- [GHL11] D. Guibourg, L. Hervé, and J. Ledoux. Quasi-compactness of Markov kernels on weighted-supremum spaces and geometrical ergodicity. arXiv : 1110.3240, 2011.
- [Hen93] H. Hennion. Sur un théorème spectral et son application aux noyaux lipchitziens. *Proc. Amer. Math. Soc.*, 118:627–634, 1993.
- [HL12] L. Hervé and J. Ledoux. Truncated Markov chains and V -geometric ergodicity via weak perturbation theory. Submitted, 2012.
- [HS92] A. Hordijk and F. Spieksma. On ergodicity and recurrence properties of a Markov chain with an application to an open Jackson network. *Adv. in Appl. Probab.*, 24(2):343–376, 1992.
- [KM03] I. Kontoyiannis and S. P. Meyn. Spectral theory and limit theorems for geometrically ergodic Markov processes. *Ann. Appl. Probab.*, 13(1):304–362, 2003.
- [KM11] I. Kontoyiannis and S. Meyn. Geometric ergodicity and the spectral gap of non-reversible Markov chains. *Probab. Theory Related Fields*, 2011. To appear.
- [LT96] R. B. Lund and R. L. Tweedie. Geometric convergence rates for stochastically ordered Markov chains. *Math. Oper. Res.*, 21(1):182–194, 1996.
- [MS95] V. A. Malyshev and F. M. Spieksma. Intrinsic convergence rate of countable Markov chains. *Markov Process. Related Fields*, 1(2):203–266, 1995.
- [MT93] S. P. Meyn and R. L. Tweedie. *Markov chains and stochastic stability*. Springer-Verlag London Ltd., London, 1993.
- [MT94] S. P. Meyn and R. L. Tweedie. Computable bounds for geometric convergence rates of Markov chains. *Ann. Probab.*, 4:981–1011, 1994.
- [RT99] G. O. Roberts and R. L. Tweedie. Bounds on regeneration times and convergence rates for Markov chains. *Stochastic Process. Appl.*, 80(2):211–229, 1999.
- [Wu04] L. Wu. Essential spectral radius for Markov semigroups. I. Discrete time case. *Probab. Theory Related Fields*, 128(2):255–321, 2004.