

## REGULAR PERTURBATION OF $V$ -GEOMETRICALLY ERGODIC MARKOV CHAINS

DÉBORAH FERRÉ,\* \*\*

LOÏC HERVÉ\* \*\*\* AND

JAMES LEDOUX,\* \*\*\*\* *Institut National des Sciences Appliquées*

### Abstract

In this paper, new conditions for the stability of  $V$ -geometrically ergodic Markov chains are introduced. The results are based on an extension of the standard perturbation theory formulated by Keller and Liverani. The continuity and higher regularity properties are investigated. As an illustration, an asymptotic expansion of the invariant probability measure for an autoregressive model with independent and identically distributed noises (with a nonstandard probability density function) is obtained.

*Keywords:* Stability; spectral method

2010 Mathematics Subject Classification: Primary 60J05; 47B07

### 1. Introduction and statements

Let  $\{P_\varepsilon\}_{|\varepsilon| < \varepsilon_0}$  be a family of transition kernels on a measurable space  $(\mathbb{X}, \mathcal{X})$ , where  $\varepsilon$  reads as a small perturbation parameter. Throughout the paper,  $V: \mathbb{X} \rightarrow [1, +\infty)$  is a fixed function. The (unperturbed) kernel  $P_0$  is assumed to satisfy the classical  $V$ -geometrical ergodicity property, that is,  $P_0$  admits a unique invariant probability measure  $\pi_0$ ,  $\pi_0(V) < \infty$ , and there exists some constants  $c \in (0, +\infty)$  and  $\kappa_1 \in (0, 1)$  such that

$$\sup_{|f| \leq V} |\mathbb{E}[f(X_n) \mid X_0 = x] - \pi_0(f)| \leq c\kappa_1^n V(x) \quad \text{for all } x \in \mathbb{X} \quad (\text{VG})$$

or, equivalently,

$$\sup_{|f| \leq V} |P_0^n f(x) - \pi_0(f)| \leq c\kappa_1^n V(x) \quad \text{for all } x \in \mathbb{X}.$$

This means that  $P_0$  has a spectral gap on the weighted supremum normed space  $\mathcal{B}_V$  of measurable functions  $f: \mathbb{X} \rightarrow \mathbb{C}$  such that  $\|f\|_V := \sup_{x \in \mathbb{X}} V(x)^{-1} |f(x)| < \infty$ . We are interested in the following two questions. For small enough  $|\varepsilon|$ ,

(Q1) does  $P_\varepsilon$  admit an invariant probability measure, say  $\pi_\varepsilon$ , and is  $P_\varepsilon$   $V$ -geometrically ergodic?

(Q2) do we have any control on  $\pi_0 - \pi_\varepsilon$ ?

Received 6 May 2011; revision received 16 July 2012.

\* Postal address: Mathematical Research Institute of Rennes, Institut National des Sciences Appliquées, 20 Avenue des Buttes de Coesmes, CS 70 839, 35708 Rennes cedex 7, France.

\*\* Email address: [deborah.ferre@insa-rennes.fr](mailto:deborah.ferre@insa-rennes.fr)

\*\*\* Email address: [loic.herve@insa-rennes.fr](mailto:loic.herve@insa-rennes.fr)

\*\*\*\* Email address: [james.ledoux@insa-rennes.fr](mailto:james.ledoux@insa-rennes.fr)

Under some classical aperiodicity and irreducibility conditions, property (VG) holds if and only if  $P_0$  satisfies the so-called drift condition based on the notion of a small set (see [15] for the definition of the drift condition which is not used here). Consequently, a natural and efficient way to study (Q1) is to prove that the perturbed Markov kernel  $P_\varepsilon$  also satisfies the drift condition (see, e.g. [3] and [18]). However, to the best of our knowledge, the theory of geometrical ergodic Markov chains does not provide any general answer to (Q2), except in terms of weak convergence in specific cases (see [18]). On the other hand, standard perturbation theory, which is a natural way to investigate (Q2), leads us to assume the following continuity condition:

$$\|P_\varepsilon - P_0\|_{\mathcal{B}_V} := \sup_{\|f\|_V \leq 1} \|P_\varepsilon f - P_0 f\|_V \rightarrow 0 \quad \text{when } \varepsilon \rightarrow 0; \quad (1)$$

that is, the operator norm of  $P_\varepsilon - P_0$  on  $\mathcal{B}_V$  goes to 0 as  $\varepsilon \rightarrow 0$ . In a series of papers, Kartashov introduced the concept of a ‘strongly stable Markov chain’ for a Markov chain with a transition kernel  $P_0$  such that, in some neighborhood of  $P_0$  with respect to  $\|\cdot\|_{\mathcal{B}_V}$ ,  $P_\varepsilon$  has a unique invariant probability measure  $\pi_\varepsilon$  with the property that  $\sup_{\|f\|_V \leq 1} |\pi_\varepsilon(f) - \pi_0(f)| \rightarrow 0$  as  $\|P_\varepsilon - P_0\|_{\mathcal{B}_V} \rightarrow 0$  uniformly in this neighborhood. Strong stability is shown to be equivalent to the convergence

$$\lim_{n \rightarrow \infty} \sup_{\|f\|_V \leq 1} \left\| \frac{1}{n} \sum_{k=1}^n P_0^k f - \pi_0(f) \mathbf{1}_{\mathbb{X}} \right\|_V = 0.$$

Moreover, if  $P_0$  is  $V$ -geometrically ergodic then, given  $\rho \in (0, 1)$ , one can consider  $N \in \mathbb{N}^*$  such that  $c\kappa_1^N \leq \rho$ , and if  $P_\varepsilon$  is such that  $\Delta_N := \|P_0^N - P_\varepsilon^N\|_{\mathcal{B}_V} < 1 - \rho$  then  $P_\varepsilon$  is  $V$ -geometrically ergodic and  $\sup_{\|f\|_V \leq 1} |\pi(f) - \pi_\varepsilon(f)| = O(\Delta_N/(1 - \rho - \Delta_N))$ . We refer the reader to [11] for an overview of results in this direction and to [17] for a related discussion for discrete state spaces. However, as discussed in [19, p. 1126] and in Example 1 below, the continuity condition (1) may be restrictive.

Similar questions arise in the context of dynamical systems. To overcome the previous difficulty, Keller introduced the more general assumption that

$$\lim_{\varepsilon \rightarrow 0} \sup_{\|f\|_0 \leq 1} \|P_\varepsilon f - P_0 f\|_1 = 0,$$

involving the two norms  $\|\cdot\|_0$  and  $\|\cdot\|_1$  (instead of a single norm) on the space on which  $P_0$  has a spectral gap [12]. This approach has been highly enhanced by the Keller–Liverani perturbation theorem [13], [14], which has proved to be a very powerful tool for studying the behavior of the Sinai–Ruelle–Bowen measures of certain perturbed dynamical systems (see, e.g. [2, Theorem 2.10] and [5, Theorem 2.8]).

The goal of this paper is to show that the Keller–Liverani theorem also provides an interesting way to investigate both (Q1) and (Q2) in the context of geometrical ergodic Markov chains. In this Markovian context, the closest work to ours is [19], where Keller’s approach is used. The results of [19] are improved here thanks to the Keller–Liverani perturbation theorem. Furthermore, in this paper, higher regularity properties than continuity are investigated when addressing (Q2). Note that the results of [13] and [14] have already been used in [4] to study some stability properties of parametric autoregressive models (for different purposes from those of Proposition 1 below).

*Notation.* For  $\beta \in [0, 1]$ , we denote by  $(\mathcal{B}_\beta, \|\cdot\|_\beta)$  the Banach space composed of the measurable functions  $f: \mathbb{X} \rightarrow \mathbb{C}$  such that  $\|f\|_\beta := \sup_{x \in \mathbb{X}} V(x)^{-\beta} |f(x)| < \infty$ . Note that  $\mathcal{B}_0$

corresponds to the space of bounded measurable functions on  $\mathbb{X}$ , with  $\|f\|_0 = \sup_{x \in \mathbb{X}} |f(x)|$ , and that  $\mathcal{B}_1 = \mathcal{B}_V$ . We denote by  $(\mathcal{L}(\mathcal{B}_\beta, \mathcal{B}_{\beta'}), \|\cdot\|_{\beta, \beta'})$  the space of all the bounded linear maps from  $\mathcal{B}_\beta$  to  $\mathcal{B}_{\beta'}$ , equipped with its usual norm:  $\|T\|_{\beta, \beta'} = \sup\{\|Tf\|_{\beta'}, f \in \mathcal{B}_\beta, \|f\|_\beta \leq 1\}$ . We write  $\mathcal{L}(\mathcal{B}_\beta)$  for  $\mathcal{L}(\mathcal{B}_\beta, \mathcal{B}_\beta)$  and  $\|T\|_\beta$  for  $\|T\|_{\beta, \beta}$ , which is a slight abuse of notation. Let  $(\mathcal{B}'_\beta, \|\cdot\|_{\mathcal{B}'_\beta})$  denote the dual space of  $\mathcal{B}_\beta$ . If  $T \in \mathcal{L}(\mathcal{B}_\beta)$  then  $T^*$  stands for the adjoint operator of  $T$ . By definition we have  $T^* \in \mathcal{L}(\mathcal{B}'_\beta)$  with the corresponding operator norm  $\|T^*\|_{\mathcal{B}'_\beta} = \|T\|_\beta$ . Note that  $T^*$  also defines an element of  $\mathcal{L}(\mathcal{B}'_\beta, \mathcal{B}'_0)$  with corresponding operator norm  $\|T^*\|_{\mathcal{B}'_\beta, \mathcal{B}'_0} \leq \|T^*\|_{\mathcal{B}'_\beta}$  from the continuous inclusion  $\mathcal{B}_0 \subset \mathcal{B}_\beta$ .

Each perturbed Markov kernel  $P_\varepsilon$  is assumed to continuously act on  $\mathcal{B}_1$ . The unperturbed kernel  $P_0$  is assumed to satisfy (VG), namely,  $P_0$  admits a unique invariant distribution  $\pi_0$  on  $(\mathbb{X}, \mathcal{X})$ ,  $\pi_0(V) < \infty$ , and

(V<sub>1</sub>) there exists  $\kappa_1 \in (0, 1)$  such that  $\|P_0^n - \pi_0(\cdot) \mathbf{1}_\mathbb{X}\|_1 = O(\kappa_1^n)$ .

We also assume that there exist  $N \in \mathbb{N}^*$ ,  $L \in (0, +\infty)$ , and  $\delta \in (0, 1)$  such that

(D) for all  $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$ ,  $P_\varepsilon^N V \leq \delta^N V + L \mathbf{1}_\mathbb{X}$ .

Under some classical aperiodicity and irreducibility assumptions, condition (V<sub>1</sub>) holds under the so-called drift condition (see [15, Chapter 4, Theorem 16.0.1]). Condition (D) on the family  $\{P_\varepsilon\}_{|\varepsilon| < \varepsilon_0}$  is weaker than the simultaneous geometrical ergodicity condition introduced in [18] since (D) involves no small set. In the  $\varepsilon = 0$  case (i.e. for a single transition kernel  $P_0$ ), the connection between the ‘best’ constants  $\kappa_1$  in (V<sub>1</sub>) and  $\delta$  in (D) is discussed in detail in [6]. See also Remark 1 below.

**Theorem 1.** *Conditions (V<sub>1</sub>), (D), and*

$$\lim_{\varepsilon \rightarrow 0} \|P_\varepsilon - P_0\|_{0,1} = 0 \quad (2)$$

*are assumed to hold. Then, setting  $\hat{\kappa} := \max(\kappa_1, \delta)$ , the following statements are satisfied.*

1. *For each  $\kappa \in (\hat{\kappa}, 1)$ , there exists  $\varepsilon_1 \in (0, \varepsilon_0]$  such that, for all  $\varepsilon \in (-\varepsilon_1, \varepsilon_1)$ ,  $P_\varepsilon$  has a unique invariant probability measure  $\pi_\varepsilon$ , with  $\pi_\varepsilon(V) < \infty$ , such that*

$$\sup_{|\varepsilon| < \varepsilon_1} \|P_\varepsilon^n - \pi_\varepsilon(\cdot) \mathbf{1}_\mathbb{X}\|_1 = O(\kappa^n). \quad (3)$$

2. *We have*

$$\lim_{\varepsilon \rightarrow 0} \sup_{\|f\|_0 \leq 1} |\pi_\varepsilon(f) - \pi_0(f)| = 0. \quad (4)$$

Note that, if each  $\pi_\varepsilon$  admits a density, say  $p_\varepsilon$ , with respect to a fixed positive measure  $\psi$  on  $(\mathbb{X}, \mathcal{X})$ , then  $\sup_{\|f\|_0 \leq 1} |\pi_\varepsilon(f) - \pi_0(f)| = \int_\mathbb{R} |p_\varepsilon(x) - p_0(x)| d\psi(x)$ .

Conclusion (3) means that  $P_\varepsilon$  is  $V$ -geometrically ergodic in a uniform way with respect to the perturbation parameter  $\varepsilon$ . Conclusion (4) means that the total variation norm of  $\pi_\varepsilon - \pi_0$  goes to 0 when  $\varepsilon \rightarrow 0$ . Consequently, Theorem 1 provides the same theoretical conclusions as those in [11, Chapter 3], but under the continuity condition (2), which is weaker than that in (1). Indeed, we have  $\|P_\varepsilon - P_0\|_{0,1} \leq \|P_\varepsilon - P_0\|_1$ . There are several examples (see, for instance, [19]) showing that the continuity condition (1) of [11] may fail, while (2) (or (8) below) holds. The case of an autoregressive process is investigated in Example 1 below. Note that the uniform bound (3) can be obtained from [16] when the kernels  $P_\varepsilon$  are assumed to satisfy the drift condition with constants and a small set which do not depend on  $\varepsilon$ . To derive (3) and

(4) from the weaker assumptions  $(V_1)$  and (D), some continuity assumption must be assumed on the map  $\varepsilon \mapsto P_\varepsilon$ .

Typically, condition  $(V_1)$  is Assumption AI of [19] (see the remark following their assumptions). Condition (2) is weaker than Assumption AII of [19] involving the sequence  $(P_\varepsilon^n)_{n \geq 1}$  (here Assumption AII is only required for  $n = 1$ ). Note that, given any  $r \in (\hat{\kappa}, 1)$  and setting  $\eta := 1 - \ln r / \ln \delta$  ( $\eta \in (0, 1)$ ), the proof of Theorem 1 and further results in [13] ensure that there exists  $D_r \in (0, +\infty)$  such that

$$\sup_{\|f\|_0 \leq 1} |\pi_\varepsilon(f) - \pi_0(f)| \leq D_r (\|P_\varepsilon - P_0\|_{0,1})^\eta. \quad (5)$$

This provides an alternative statement to [19, Theorem 3.1], which states an inequality of type (5) in assuming the existence of  $\pi_\varepsilon$ . The  $V$ -geometrical ergodicity of the perturbed kernel  $P_\varepsilon$  was an open question in [19]. Actually, under their assumptions, condition (D) is a quite natural hypothesis for  $P_\varepsilon$  to inherit the  $V$ -geometrical ergodicity of the unperturbed kernel  $P_0$  (note that condition (D) is not so far from the drift condition). In the proof of Theorem 1, condition (D) is viewed as a Doeblin–Fortet inequality on the dual space of  $\mathcal{B}_1$  in order to use the Keller–Liverani theorem.

**Example 1.** (*Autoregressive process.*) Assume that  $\mathbb{X} := \mathbb{R}$  and that  $(X_n)_{n \in \mathbb{N}}$  is the autoregressive model defined by

$$X_n = \alpha X_{n-1} + \vartheta_n, \quad n \in \mathbb{N}^*, \quad (6)$$

where  $X_0$  is a real-valued random variable,  $\alpha \in (-1, 1)$ , and  $(\vartheta_n)_{n \geq 1}$  is a sequence of independent and identically distributed real-valued random variables, independent of  $X_0$ . Assume that  $\vartheta_1$  has a Lebesgue probability density function on  $\mathbb{X}$ , say  $\nu(\cdot)$ , and admits a first moment,  $\int |x| \nu(x) dx < \infty$ . We know that  $(X_n)_{n \in \mathbb{N}}$  is a Markov chain with transition kernel

$$P_\alpha(x, A) = \int_{\mathbb{R}} \mathbf{1}_A(\alpha x + y) \nu(y) dy = \int_{\mathbb{R}} \mathbf{1}_A(y) \nu(y - \alpha x) dy. \quad (7)$$

Set  $V(x) := 1 + |x|$ ,  $x \in \mathbb{R}$ . It is known (see, e.g. [20, Section 8] or [6, Section 5.5]) that, for each  $\alpha \in (-1, 1)$ ,  $(X_n)_{n \in \mathbb{N}}$  is  $V$ -geometrically ergodic with an invariant distribution  $\pi_\alpha$ . Next, given any  $a_0 \in (0, 1)$ , it can be easily checked that the family  $\{P_\alpha, \alpha \in (-a_0, a_0)\}$  satisfies condition (D) with  $N = 1$  and any  $\delta \in (a_0, 1)$ . Finally, we prove below that, for every  $\alpha_0 \in (-1, 1)$ , the kernels  $P_{\alpha_0 + \varepsilon}$  satisfy the weak continuity condition (2). The previous facts ensure that Theorem 1 applies to the family  $(P_{\alpha_0 + \varepsilon})_\varepsilon$ , so that the total variation norm of  $\pi_\alpha - \pi_{\alpha_0}$  goes to 0 when  $\alpha \rightarrow \alpha_0$ . Note that no specific assumptions are required on the density  $\nu$ , except that it admits a first moment. To the best of our knowledge, this last result is new. It cannot be deduced from a known method, as in [11], since condition (1) does not hold, as we prove below.

First we show that, for any  $\alpha_0 \in (-1, 1)$ , we have  $\lim_{\alpha \rightarrow \alpha_0} \|P_\alpha - P_{\alpha_0}\|_{0,1} = 0$ , that is, for all  $\varepsilon > 0$ , there exists  $\eta > 0$  such that

$$|\alpha - \alpha_0| < \eta \implies \|P_\alpha - P_{\alpha_0}\|_{0,1} = \sup_{\|f\|_0 \leq 1} \sup_{x \in \mathbb{R}} \frac{|P_\alpha f(x) - P_{\alpha_0} f(x)|}{V(x)} < \varepsilon.$$

From (7), we have  $P_\alpha f(x) = \int_{\mathbb{R}} f(y) \nu(y - \alpha x) dy$ , so

$$\|P_\alpha - P_{\alpha_0}\|_{0,1} \leq \sup_{x \in \mathbb{R}} \frac{\|\nu(\cdot - \alpha x) - \nu(\cdot - \alpha_0 x)\|_{\mathbb{L}^1(\mathbb{R})}}{V(x)}.$$

Let  $\varepsilon > 0$ . Since  $V(x) \rightarrow +\infty$  when  $x \rightarrow \infty$ , we can choose a constant  $A > 0$  such that  $4/V(x) < \varepsilon$  for any  $|x| > A$ . Therefore, using the fact that  $v$  is a probability density function, we obtain

$$\sup_{|x| > A} \frac{\|v(\cdot - \alpha x) - v(\cdot - \alpha_0 x)\|_{\mathbb{L}^1(\mathbb{R})}}{V(x)} \leq \sup_{|x| > A} \frac{2}{V(x)} < \varepsilon.$$

Next, the map from  $\mathbb{R}$  into  $\mathbb{L}^1(\mathbb{R})$  defined by  $t \mapsto v(\cdot - t)$  is continuous, so we can pick  $\eta' > 0$  such that, for all  $|x| \leq A$ ,

$$|\alpha x - \alpha_0 x| \leq |\alpha - \alpha_0|A < \eta' \implies \|v(\cdot - \alpha x) - v(\cdot - \alpha_0 x)\|_{\mathbb{L}^1(\mathbb{R})} < \varepsilon.$$

Then, using  $V \geq 1$ , there exists  $\eta := \eta'/A$  such that, for any  $\alpha$  satisfying  $|\alpha - \alpha_0| < \eta$ ,

$$\sup_{|x| \leq A} \frac{\|v(\cdot - \alpha x) - v(\cdot - \alpha_0 x)\|_{\mathbb{L}^1(\mathbb{R})}}{V(x)} \leq \sup_{|x| \leq A} \|v(\cdot - \alpha x) - v(\cdot - \alpha_0 x)\|_{\mathbb{L}^1(\mathbb{R})} < \varepsilon.$$

Second, let us check that, whatever the density  $v(\cdot)$ , the strong continuity condition (1) is never fulfilled. Let us consider a positive real number  $a$  such that

$$\int_{-a}^a v(y) dy \neq \int_{-2a}^{-a} v(y) dy + \int_a^{2a} v(y) dy.$$

Such a real number exists for any density  $v(\cdot)$  (if the previous terms coincide for every  $a > 0$  then we obtain a contradiction when  $a \rightarrow +\infty$ ). Let  $\alpha_0 \in (0, 1)$  be fixed. For each  $\alpha \in (\alpha_0, 1)$ , define

$$x_\alpha := \frac{a}{\alpha - \alpha_0} > 0.$$

Note that  $-a + \alpha_0 x_\alpha < -a + \alpha x_\alpha < a + \alpha_0 x_\alpha < a + \alpha x_\alpha$ . Next let us introduce the following function:

$$f_\alpha(y) := y \mathbf{1}_{[a+\alpha_0 x_\alpha, a+\alpha x_\alpha]}(y) - y \mathbf{1}_{[-a+\alpha_0 x_\alpha, -a+\alpha x_\alpha]}(y) \quad \text{for all } y \in \mathbb{R}.$$

From (7) we have

$$\begin{aligned} (P_\alpha f_\alpha)(x_\alpha) &= \int_0^a (y + \alpha x_\alpha) v(y) dy - \int_{-2a}^{-a} (y + \alpha x_\alpha) v(y) dy, \\ (P_{\alpha_0} f_\alpha)(x_\alpha) &= \int_a^{2a} (y + \alpha_0 x_\alpha) v(y) dy - \int_{-a}^0 (y + \alpha_0 x_\alpha) v(y) dy. \end{aligned}$$

Set  $J := \int_{-a}^a y v(y) dy - \int_{-2a}^{-a} y v(y) dy - \int_a^{2a} y v(y) dy$ . We obtain

$$\begin{aligned} (P_\alpha f_\alpha)(x_\alpha) - (P_{\alpha_0} f_\alpha)(x_\alpha) \\ = J + \alpha x_\alpha \left( \int_0^a v(y) dy - \int_{-2a}^{-a} v(y) dy \right) + \alpha_0 x_\alpha \left( \int_{-a}^0 v(y) dy - \int_a^{2a} v(y) dy \right). \end{aligned}$$

Since  $V(x_\alpha) = 1 + x_\alpha$  and  $x_\alpha \rightarrow +\infty$  as  $\alpha \downarrow \alpha_0$ , it follows that

$$\lim_{\alpha \downarrow \alpha_0} \frac{(P_\alpha f_\alpha)(x_\alpha) - (P_{\alpha_0} f_\alpha)(x_\alpha)}{V(x_\alpha)} = \alpha_0 \left( \int_{-a}^a v(y) dy - \int_{-2a}^{-a} v(y) dy - \int_a^{2a} v(y) dy \right) \neq 0.$$

Finally, from

$$\begin{aligned} \|P_\alpha - P_{\alpha_0}\|_1 &= \|P_\alpha - P_{\alpha_0}\|_{\mathcal{B}_V} \\ &= \sup_{|f| \leq V} \sup_{x \in \mathbb{R}} \frac{|(P_\alpha f)(x) - (P_{\alpha_0} f)(x)|}{V(x)} \\ &\geq \frac{|(P_\alpha f_\alpha)(x_\alpha) - (P_{\alpha_0} f_\alpha)(x_\alpha)|}{V(x_\alpha)}, \end{aligned}$$

we deduce that  $\|P_\alpha - P_{\alpha_0}\|_{\mathcal{B}_V}$  does not go to 0 when  $\alpha \downarrow \alpha_0$ .

Reinforcing the continuity condition (2) of Theorem 1 allows us to obtain the following refinement.

**Corollary 1.** *Assume that conditions  $(V_1)$  and  $(D)$  hold, and that, for  $\beta \in (0, 1]$ ,*

*$(V_\beta)$  there exists  $\kappa_\beta \in (0, 1)$  such that  $\|P_0^n - \pi_0(\cdot) \mathbf{1}_{\mathbb{X}}\|_\beta = O(\kappa_\beta^n)$*

*and*

$$\lim_{\varepsilon \rightarrow 0} \|P_\varepsilon - P_0\|_{\beta,1} = 0. \quad (8)$$

*Then there exist  $\varepsilon_1 \in (0, \varepsilon_0]$  (from Theorem 1) and  $C \in (0, +\infty)$  such that, for all  $\varepsilon \in (-\varepsilon_1, \varepsilon_1)$ ,*

$$\sup_{\|f\|_\beta \leq 1} |\pi_\varepsilon(f) - \pi_0(f)| \leq C \|P_\varepsilon - P_0\|_{\beta,1}.$$

Note that, under the drift condition,  $(V_1)$  implies  $(V_\beta)$  by Jensen's inequality for  $\beta \in (0, 1]$ .

**Remark 1.** Let  $P$  be any Markov kernel on  $\mathbb{X}$ . If we have  $P^N V \leq \delta^N V + L \mathbf{1}_{\mathbb{X}}$  for some  $N \in \mathbb{N}^*$ ,  $L \in (0, +\infty)$ , and  $\delta \in (0, 1)$  and if  $P^\ell$  is compact from  $\mathcal{B}_0$  to  $\mathcal{B}_V$  (for some  $\ell \geq 1$ ), then  $P$  is a power-bounded quasi-compact operator on  $\mathcal{B}_V$ . Moreover, its essential spectral radius is such that  $r_{\text{ess}}(P) \leq \delta$ . Actually, denoting by  $\delta_V(P)$  the infimum of the real numbers  $\delta$  for which the above inequality holds for some  $N \in \mathbb{N}^*$  and  $L \in (0, +\infty)$ , we have  $r_{\text{ess}}(P) \leq \delta_V(P)$ , with equality in many cases. This follows from the Doeblin–Fortet inequalities (10) below and [8], since  $(P^*)^\ell$  is compact from  $\mathcal{B}'_V$  to  $\mathcal{B}'_0$ . If, in addition,  $P$  satisfies aperiodicity and irreducibility conditions, then  $P$  is  $V$ -geometrically ergodic. Such results are fully detailed and applied in [6] (without any perturbation issues). Now, if  $\rho_V(P)$  denotes the infimum bound of the real numbers  $\kappa_1$  such that  $(V_1)$  holds then we have  $\rho_V(P) \geq r_{\text{ess}}(P)$ . If there are eigenvalues  $\lambda$  of  $P$  on  $\mathcal{B}_V$  such that  $r_{\text{ess}}(P) < |\lambda| < 1$  then  $\rho_V(P)$  is the maximal modulus of such eigenvalues (which are in finite number from the definition of  $r_{\text{ess}}(P)$ ); if not, we have  $\rho_V(P) = r_{\text{ess}}(P)$ . For instance, if  $(X_n)_{n \in \mathbb{N}}$  is defined by (6) and if the probability density function  $v(\cdot)$  of the noise has a moment of order  $r \in [1, +\infty)$ , then  $(X_n)_{n \in \mathbb{N}}$  is  $(1 + |\cdot|)^r$ -geometrically ergodic with  $r_{\text{ess}}(P) \leq |\alpha|^r$  and  $\rho_V(P) = |\alpha|$ . See [20, Section 8] or [6, Section 5] for details.

Now we are interested in asymptotic expansions of the perturbed invariant distributions. Under condition (VG) and the continuity condition (1), expansions related to the generalized potential  $R = (I - P_0 + \Pi_0)^{-1}$ , with  $\Pi_0(\cdot) = \pi_0(\cdot) \mathbf{1}_{\mathbb{X}}$ , are given in [7], [10], and [11], namely,

$$\pi_\varepsilon = \pi_0(I - D_\varepsilon R)^{-1} = \sum_{k \geq 0} \pi_0(D_\varepsilon R)^k = \pi_0 + \sum_{k=1}^n \pi_0(D_\varepsilon R)^k + O((D_\varepsilon R)^{n+1}),$$

when  $D_\varepsilon = P_\varepsilon - P_0$  is such that  $\|D_\varepsilon R\|_1 < 1$ . In general, this is not a Taylor expansion, except, for instance, when  $D_\varepsilon$  has the form  $D_\varepsilon = \varepsilon D$  with  $D \in \mathcal{L}(\mathcal{B}_1)$ . This special case

is discussed in [1]. In fact, for general perturbations, even in the case when (1) is fulfilled, obtaining Taylor expansions for  $\pi_\varepsilon$  causes difficulties when the derivatives of the perturbed kernels  $P_\varepsilon(x, \cdot)$  with respect to  $\varepsilon$  yield some weights (see the term  $x^k$  in (14) below for the order of regularity  $k \geq 1$ ). For this difficulty, the derivation procedure (see [5] and [9]) based on the Keller–Liverani theorem is of great interest.

To illustrate this approach, we only consider the special instance of the autoregressive model introduced in Example 1. More specifically, we assume that  $(X_n)_{n \in \mathbb{N}}$  is given by (6) and that the probability density function  $v(\cdot)$  of the noise has a moment of order  $r$ ,  $\int |x|^r v(x) dx < \infty$  for some  $r \in [1, +\infty)$ . Set  $V(x) := (1 + |x|)^r$ ,  $x \in \mathbb{R}$ . Then, for each  $\alpha \in (-1, 1)$ ,  $(X_n)_{n \in \mathbb{N}}$  is  $V$ -geometrically ergodic with an invariant distribution  $\pi_\alpha$ . Moreover, we assume that  $r$  is not an integer (i.e.  $r > \lfloor r \rfloor$ , where  $\lfloor \cdot \rfloor$  stands for the integer part function on  $\mathbb{R}$ ), and that  $v(\cdot)$  is positive and  $(\lfloor r \rfloor + 1)$ -times continuously differentiable on  $\mathbb{R}$ , with

$$\sup_{t \in \mathbb{R}} \frac{|v^{(j)}(t)|}{v(t)} < \infty, \quad j = 1, \dots, \lfloor r \rfloor + 1. \quad (9)$$

Finally, suppose that, for all  $x_0 \in \mathbb{R}$ , there exist a neighborhood  $V_{x_0}$  of  $x_0$  and a nonnegative measurable function  $q_{x_0}(\cdot)$  such that  $\int_{\mathbb{R}} (1 + |y|)^r q_{x_0}(y) dy < \infty$  and, for all  $y \in \mathbb{R}$  and  $v \in V_{x_0}$ ,  $v(y + v) \leq q_{x_0}(y)$ .

**Proposition 1.** *Under the above assumptions, there exists  $\beta \equiv \beta(r) \in (0, 1)$  such that the map  $\alpha \mapsto \pi_\alpha$  is  $\lfloor r \rfloor$ -times continuously differentiable from  $(-1, 1)$  to the dual space  $\mathcal{B}'_\beta$  of  $\mathcal{B}_\beta$ . In particular, for all  $\alpha \in (-1, 1)$ , there exist  $\lfloor r \rfloor$  signed measures  $\mu_{\alpha,1}, \dots, \mu_{\alpha,\lfloor r \rfloor}$  on  $\mathbb{R}$  such that, for all  $A \in \mathcal{X}$ ,*

$$\pi_{\alpha+\varepsilon}(A) = \pi_\alpha(A) + \sum_{j=1}^{\lfloor r \rfloor} \frac{\varepsilon^j}{j!} \mu_{\alpha,j}(A) + \varepsilon^{\lfloor r \rfloor} R_\varepsilon(A) \quad \text{with} \quad \lim_{\varepsilon \rightarrow 0} \sup_{A \in \mathcal{X}} |R_\varepsilon(A)| = 0.$$

Condition (9) is well suited to densities of the form  $v(x) = c(1 + |x|)^{-\gamma}$ . For other forms of the density, (9) can be easily adapted so that Proposition 1 is still valid.

## 2. Proofs

*Proof of Theorem 1.* We will repeatedly use the fact that a bounded linear operator (between two normed vector spaces) and its adjoint have the same norm. The adjoint of the rank 1 projection  $\Pi_0(f) = \pi_0(f) \mathbf{1}_{\mathbb{X}}$  is defined by  $\Pi_0^*(f') := f'(\mathbf{1}_{\mathbb{X}}) \pi_0$  for all  $f' \in \mathcal{B}'_1$ .

First, assumption (V<sub>1</sub>) gives

$$\|(P_0^*)^n - \Pi_0^*\|_{\mathcal{B}'_1} = O(\kappa_1^n).$$

Second, it follows from the continuity assumption (2) of Theorem 1 that

$$\lim_{\varepsilon \rightarrow 0} \|P_\varepsilon^* - P_0^*\|_{\mathcal{B}'_1, \mathcal{B}'_0} = 0.$$

Third, assumption (D) gives

$$\|P_\varepsilon^{*N} f'\|_{\mathcal{B}'_1} \leq \delta^N \|f'\|_{\mathcal{B}'_1} + L \|f'\|_{\mathcal{B}'_0} \quad \text{for all } \varepsilon \in (-\varepsilon_0, \varepsilon_0) \text{ and } f' \in \mathcal{B}'_1. \quad (10)$$

Indeed, recall that  $\mathcal{B}'_1$  and  $\mathcal{B}'_0$  are Banach lattices and that, for each  $g' \in \mathcal{B}'_1$ , we have  $\|g'\|_{\mathcal{B}'_0} = \| |g'| \|_{\mathcal{B}'_0} = \langle |g'|, \mathbf{1}_{\mathbb{X}} \rangle$  and  $\|g'\|_{\mathcal{B}'_1} = \| |g'| \|_{\mathcal{B}'_1} = \langle |g'|, V \rangle$ , where  $\langle \cdot, \cdot \rangle$  stands for the duality

brackets in both  $\mathcal{B}'_0 \times \mathcal{B}_0$  and  $\mathcal{B}'_1 \times \mathcal{B}_1$ . Next, observe that, for every  $f' \in \mathcal{B}'_1$  and every  $f \in \mathcal{B}_1$  such that  $|f| \leq V$ , we have  $|\langle (P^*)^N f', f \rangle| \leq \langle |f'|, P^N V \rangle$ . Hence, (D) gives (10) as claimed.

The three above statements show that  $\{P_\varepsilon^*\}_{|\varepsilon| < \varepsilon_0}$  satisfies the assumptions of [14] on  $\mathcal{B}'_1$  (see also [9] for the use of the perturbation Keller–Liverani theorem in a Markov context). Therefore, for all  $\kappa \in (\hat{\kappa}, 1)$ , there exists  $\varepsilon_1 \in (0, \varepsilon_0]$  such that, for all  $\varepsilon \in (-\varepsilon_1, \varepsilon_1)$ , the following properties hold: there exist  $\lambda_\varepsilon \in \mathbb{C}$  satisfying  $\lim_{\varepsilon \rightarrow 0} \lambda_\varepsilon = 1$  and a rank 1 projection  $\Pi'_\varepsilon$  on  $\mathcal{B}'_1$  such that

$$\sup_{|\varepsilon| < \varepsilon_1} \|(P_\varepsilon^*)^n - \lambda_\varepsilon^n \Pi'_\varepsilon\|_{\mathcal{B}'_1} = O(\kappa^n) \quad \text{and} \quad \lim_{\varepsilon \rightarrow 0} \|\Pi'_\varepsilon - \Pi_0^*\|_{\mathcal{B}'_1, \mathcal{B}_0} = 0. \quad (11)$$

For small enough  $\varepsilon_1$ , we have  $\lambda_\varepsilon = 1$ . Indeed, one may assume that  $|\lambda_\varepsilon| > \kappa$ . Then we deduce from  $\mathbf{1}_\mathbb{X} = P_\varepsilon^n \mathbf{1}_\mathbb{X}$  and (11) that, for any  $f' \in \mathcal{B}'_1$ , we have  $\lim_n \lambda_\varepsilon^{-n} \langle f', \mathbf{1}_\mathbb{X} \rangle = \lim_n \lambda_\varepsilon^{-n} \langle (P_\varepsilon^*)^n f', \mathbf{1}_\mathbb{X} \rangle = \langle \Pi'_\varepsilon f', \mathbf{1}_\mathbb{X} \rangle$ . Thus,  $(\lambda_\varepsilon^{-n})_{n \in \mathbb{N}}$  converges in  $\mathbb{C}$ , so we have either  $\lambda_\varepsilon = 1$  or  $|\lambda_\varepsilon| > 1$ . Moreover, the sequence  $(P_\varepsilon^n)_{n \in \mathbb{N}}$  is bounded in  $\mathcal{L}(\mathcal{B}_V)$  from (D). Thus,  $((P_\varepsilon^*)^n)_{n \in \mathbb{N}}$  is bounded in  $\mathcal{L}(\mathcal{B}'_V)$ , and (11) then implies that  $(\lambda_\varepsilon^n)_{n \in \mathbb{N}}$  is bounded in  $\mathbb{C}$ . Thus,  $\lambda_\varepsilon = 1$ . Hence,  $((P_\varepsilon^*)^n)_{n \in \mathbb{N}}$  is Cauchy, and so is  $(P_\varepsilon^n)_{n \in \mathbb{N}}$  in  $\mathcal{L}(\mathcal{B}_1)$ . Therefore, there exists  $\Pi_\varepsilon \in \mathcal{L}(\mathcal{B}_1)$  such that

$$\lim_n P_\varepsilon^n = \Pi_\varepsilon \quad \text{in } \mathcal{L}(\mathcal{B}_1). \quad (12)$$

Thus,  $(P_\varepsilon^n)^* \rightarrow \Pi_\varepsilon^*$  in  $\mathcal{L}(\mathcal{B}'_1)$ , and so  $\Pi'_\varepsilon = \Pi_\varepsilon^*$ . We have obtained

$$\sup_{|\varepsilon| < \varepsilon_1} \|P_\varepsilon^n - \Pi_\varepsilon\|_1 = O(\kappa^n) \quad \text{and} \quad \lim_{\varepsilon \rightarrow 0} \|\Pi_\varepsilon - \Pi_0\|_{0,1} = 0.$$

From (12), it follows that  $\Pi_\varepsilon$  is a positive projection on  $\mathcal{B}_1$  satisfying  $\Pi_\varepsilon P_\varepsilon = P_\varepsilon \Pi_\varepsilon = \Pi_\varepsilon$ . Let us prove that  $\Pi_\varepsilon$  is rank 1. We know that  $\Pi'_\varepsilon$  is rank 1, namely,  $\Pi'_\varepsilon(\cdot) := \langle \phi_\varepsilon, \cdot \rangle a'_\varepsilon$  for some  $\phi_\varepsilon \neq 0$  in the dual space of  $\mathcal{B}'_1$  and for some  $a'_\varepsilon \neq 0$  in  $\mathcal{B}'_1$ . From  $\Pi_\varepsilon^* = \Pi'_\varepsilon$ , we obtain  $\ker \Pi_\varepsilon = \ker a'_\varepsilon$ . Thus,  $\ker \Pi_\varepsilon$  is of codimension 1, so we have  $\dim \operatorname{Im} \Pi_\varepsilon = 1$  (use  $\mathcal{B}_1 = \ker \Pi_\varepsilon \oplus \operatorname{Im} \Pi_\varepsilon$ ).

The last fact shows that  $\Pi_\varepsilon(\cdot) := e'_\varepsilon(\cdot) \mathbf{1}_\mathbb{X}$  for some  $e'_\varepsilon \geq 0$  in  $\mathcal{B}'_1$ . Since, for all  $A \in \mathcal{X}$ ,  $\lim_n P_\varepsilon^n(x_0, A) = e'_\varepsilon(\mathbf{1}_A)$  (for some fixed  $x_0 \in \mathbb{X}$ ) and  $P_\varepsilon^n(x_0, \cdot)$  is a probability measure on  $(\mathbb{X}, \mathcal{X})$  for each  $n \geq 1$ , we deduce from the Vitali–Hahn–Saks theorem that  $A \mapsto \pi_\varepsilon(A) := e'_\varepsilon(\mathbf{1}_A)$  is a probability measure on  $(\mathbb{X}, \mathcal{X})$ . Clearly,  $\pi_\varepsilon$  is  $P_\varepsilon$ -invariant. Next  $\pi_\varepsilon$  and  $e'_\varepsilon$  coincide on  $\mathcal{B}_0$ : indeed, every  $f \in \mathcal{B}_0$  can be approached uniformly on  $\mathbb{X}$  by a sequence  $(f_n)_n$  of simple functions, so  $\pi_\varepsilon(f) = \lim_n \pi_\varepsilon(f_n) = \lim_n e'_\varepsilon(f_n) = e'_\varepsilon(f)$  (the last convergence holds since  $e'_\varepsilon \in \mathcal{B}'_1$ ). We obtain  $\pi_\varepsilon(V \mathbf{1}_{[V \leq n]}) = e'_\varepsilon(V \mathbf{1}_{[V \leq n]}) \leq e'_\varepsilon(V) < \infty$  for every  $n \geq 1$  since  $e'_\varepsilon \geq 0$ . The monotone convergence theorem gives  $\pi_\varepsilon(V) < \infty$ . Thus,  $f \mapsto \pi_\varepsilon(f)$  is in  $\mathcal{B}'_1$ . Since, for all  $f \in \mathcal{B}_1$ ,  $\lim_n P_\varepsilon^n f = e'_\varepsilon(f) \mathbf{1}_\mathbb{X}$  in  $\mathcal{B}_1$ , we deduce that  $e'_\varepsilon = \pi_\varepsilon$  on  $\mathcal{B}_1$  from the invariance of  $\pi_\varepsilon$ . Finally, the last assertion of Theorem 1 follows from  $\lim_{\varepsilon \rightarrow 0} \|\Pi_\varepsilon - \Pi_0\|_{0,1} = 0$ .

*Proof of Corollary 1.* Let  $\kappa \in (\max(\kappa_\beta, \kappa_1), 1)$ , and let  $\Gamma$  be the oriented circle in  $\mathbb{C}$ , centered at  $z = 1$  and with radius less than  $(1 - \kappa)/2$ . Note that the assumptions of Corollary 1 imply those of Theorem 1. We know from [14] and the proof of Theorem 1 that  $M := \sup_{z \in \Gamma, |\varepsilon| < \varepsilon_1} \|(zI - P_\varepsilon)^{-1}\|_1 < \infty$ . Moreover, the rank 1 eigenprojection  $\Pi_\varepsilon$  is from standard spectral theory:

$$\Pi_\varepsilon = \frac{1}{2i\pi} \oint_\Gamma (zI - P_\varepsilon)^{-1} dz. \quad (13)$$



Next condition  $(V_\beta)$  gives  $M_0 := \sup_{z \in \Gamma} \|(zI - P_0)^{-1}\|_\beta < \infty$ . Assume that  $\|\mathbf{1}_\mathbb{X}\|_1 = 1$  (to simplify). Then, from

$$(zI - P_\varepsilon)^{-1} - (zI - P_0)^{-1} = (zI - P_\varepsilon)^{-1}(P_\varepsilon - P_0)(zI - P_0)^{-1},$$

we obtain, for every  $f \in \mathcal{B}_\beta$ ,

$$|\pi_\varepsilon(f) - \pi_0(f)| = \|\Pi_\varepsilon(f) - \Pi_0(f)\|_1 \leq \frac{1-\kappa}{2} M \|P_\varepsilon - P_0\|_{\beta,1} M_0 \|f\|_\beta.$$

*Proof of Proposition 1.* Recall that  $V(x) := (1 + |x|)^r$ ,  $x \in \mathbb{R}$ . Let  $a_0 \in (0, 1)$ . We apply the derivation procedure of [9, Appendix A] to the family  $\{P_\alpha, \alpha \in (-a_0, a_0)\}$  with respect to the Banach spaces  $\mathcal{B}_\beta$ . Let  $\beta \in (0, 1]$ . For any  $\alpha \in (-a_0, a_0)$ , the Markov kernel  $P_\alpha$  of the autoregressive model  $(X_n)_{n \in \mathbb{N}}$  satisfies the drift condition and the aperiodicity/irreducibility assumptions of [15]. Hence, each  $P_\alpha$  satisfies condition  $(V_\beta)$ . In the same way,  $\{P_\alpha\}_{|\alpha| < a_0}$  satisfies condition (D), so it also satisfies (D) with respect to the function  $V^\beta$  thanks to Jensen's inequality. Let  $\alpha \in (-a_0, a_0)$ . Then  $\lim_{\varepsilon \rightarrow 0} \|P_{\alpha+\varepsilon} - P_\alpha\|_{0,\beta} = 0$  from Lemma 1 below. Hence, the perturbation Keller–Liverani theorem can be applied in  $\mathcal{L}(\mathcal{B}_\beta)$  to the family  $(P_{\alpha+\varepsilon})_\varepsilon$  (as seen with  $\beta = 1$  in the proof of Theorem 1). This gives the following spectral properties: there exists  $\kappa_{\alpha,\beta} \in (0, 1)$  such that, for all  $\kappa \in (\kappa_{\alpha,\beta}, 1)$ , there exists  $\hat{\varepsilon} > 0$  such that the resolvents  $(zI - P_{\alpha+\varepsilon})^{-1}$  are well defined and uniformly bounded in  $\mathcal{L}(\mathcal{B}_\beta)$  provided that  $|\varepsilon| \leq \hat{\varepsilon}$ , and  $z \in \mathbb{C}$  satisfies  $|z| \geq \kappa$  and  $|z - 1| \geq (1 - \kappa)/2$ .

Now let us introduce the formal derivative operators of  $\alpha \mapsto P_\alpha$ :

$$(P_{k,\alpha} f)(x) = (-1)^k x^k \int_{\mathbb{R}} f(y) v^{(k)}(y - \alpha x) dy \quad \text{for all } x \in \mathbb{R}. \quad (14)$$

**Lemma 1.** *Let  $(\beta, \beta') \in [0, 1]^2$ .*

1. *If  $\beta + k/r < \beta' \leq 1$  for  $k \in \{0, \dots, \lfloor r \rfloor\}$  then  $\alpha \mapsto P_{k,\alpha}$  is continuous from  $(-a_0, a_0)$  to  $\mathcal{L}(\mathcal{B}_\beta, \mathcal{B}_{\beta'})$ .*
2. *If  $\beta + (k+1)/r < \beta' \leq 1$  for  $k \in \{0, \dots, \lfloor r \rfloor - 1\}$  then  $\alpha \mapsto P_{k,\alpha}$  is continuously differentiable from  $(-a_0, a_0)$  to  $\mathcal{L}(\mathcal{B}_\beta, \mathcal{B}_{\beta'})$ , with  $dP_{k,\alpha}(\alpha)/d\alpha = P_{k+1,\alpha}$ .*

Let Lemma 1 hold and apply [9, Appendix A]. For  $\beta \in (0, 1)$  and  $\sigma > 0$ , we set  $T_0(\beta) = \beta + \sigma/r$  and  $T_1(\beta) = \beta + (1 + \sigma)/r$ . It follows from Lemma 1 that, for  $j = 1, \dots, \lfloor r \rfloor$ ,

- if  $T_0(\beta) \in (0, 1]$  then  $\alpha \mapsto P_\alpha$  is continuous from  $(-a_0, a_0)$  to  $\mathcal{L}(\mathcal{B}_\beta, \mathcal{B}_{T_0(\beta)})$ ,
- if  $T_1(T_0 T_1)^{j-1}(\beta) \in (0, 1]$  then  $\alpha \mapsto P_\alpha$  is  $j$ -times continuously differentiable from  $(-a_0, a_0)$  to  $\mathcal{L}(\mathcal{B}_\beta, \mathcal{B}_{T_1(T_0 T_1)^{j-1}(\beta)})$ .

Let  $\beta_r \in (0, 1 - \lfloor r \rfloor/r)$ , and let  $\sigma > 0$  be such that  $\beta_r + [(2\lfloor r \rfloor + 1)\sigma + \lfloor r \rfloor]/r = 1$ . In other words, we have  $(T_0 T_1)^{\lfloor r \rfloor} T_0(\beta_r) = 1$ . Let  $\alpha_0 \in (-a_0, a_0)$ . Then it follows from [9, Proposition A.1] that there exists  $\tilde{\kappa} \in (0, 1)$  such that, for all  $z \in \mathcal{D} := \{z \in \mathbb{C} : |z| \geq \tilde{\kappa}, |z - 1| \geq (1 - \tilde{\kappa})/2\}$ , the map  $\alpha \mapsto (zI - P_\alpha)^{-1}$  is  $\lfloor r \rfloor$ -times continuously differentiable from some open interval  $I_{\alpha_0}$  centered at  $\alpha_0$  in  $\mathcal{L}(\mathcal{B}_{\beta_r}, \mathcal{B}_1)$ . Moreover, the derivatives (up to the order  $\lfloor r \rfloor$ ) of the last map are uniformly bounded in  $(\alpha, z) \in I_{\alpha_0} \times \mathcal{D}$ . Next, let us define  $\Pi_\alpha(f) = \pi_\alpha(\cdot) \mathbf{1}_\mathbb{R}$  for any  $f \in \mathcal{B}_{\beta_r}$ . From the above facts and standard spectral calculus (see (13)), the map  $\alpha \mapsto \Pi_\alpha(f)$  is  $\lfloor r \rfloor$ -times continuously differentiable from  $I_{\alpha_0}$  to

$\mathcal{L}(\mathcal{B}_{\beta_r}, \mathcal{B}_1)$ , so is  $\alpha \mapsto \pi_\alpha(\cdot)$  from  $I_{\alpha_0}$  to  $\mathcal{B}'_{\beta_r}$ . Since  $\alpha_0$  is any element in  $(-a_0, a_0)$ , with arbitrary  $a_0 \in (-1, 1)$ , this proves the first assertion of Proposition 1.

Let  $\alpha \in (-1, 1)$ . For all Borel sets  $A$  of  $\mathbb{R}$ , we have  $d\pi_\alpha(A)/d\alpha = \lim_{h \rightarrow 0} (\pi_{\alpha+h}(A) - \pi_\alpha(A))/h$ . Since  $\pi_{\alpha+h}(\cdot) - \pi_\alpha(\cdot)$  is a (signed) measure on  $\mathbb{R}$ , it follows from the Vitali–Hahn–Saks theorem that there exists a (signed) measure  $\mu_{\alpha,1}$  on  $\mathbb{R}$  such that we have  $d\pi_\alpha(A)/d\alpha = \mu_{\alpha,1}(A)$ . An obvious induction gives the same conclusion for the derivatives of order  $j = 2, \dots, \lfloor r \rfloor$ .

*Proof of Lemma 1.* Set  $A_k := \sup_{t \in \mathbb{R}} |v^{(k)}(t)|/v(t)$  and  $A := \max_{0 \leq k \leq \lfloor r \rfloor + 1} A_k$ . Note that there exists  $B \equiv B(\beta)$  such that, for all  $\alpha \in [-a_0, a_0]$ , we have  $P_\alpha V^\beta \leq B V^\beta$ .

For  $\alpha \in (-a_0, a_0)$ ,  $\beta \in (0, 1]$ ,  $f \in \mathcal{B}_\beta$ , and  $x \in \mathbb{R}$ , we define  $F_k(\alpha) := (P_{k,\alpha} f)(x)$ . From the assumptions on  $v$  and Lebesgue’s theorem, for every  $k = 0, \dots, \lfloor r \rfloor$ ,  $F_k$  is differentiable on  $[-a_0, a_0]$ , with  $dF_k(\alpha)/d\alpha = F_{k+1}(\alpha)$ .

Let  $k \in \{0, \dots, \lfloor r \rfloor\}$ , let  $(\beta, \beta') \in [0, 1]^2$  be such that  $\beta + k/r < \beta' \leq 1$ , and let  $0 < \sigma \leq 1$  be such that  $\beta + (k + \sigma)/r = \beta'$ . For any  $(\alpha, \alpha') \in (-a_0, a_0)^2$ , we obtain

$$|F_k(\alpha) - F_k(\alpha')| \leq 2A_k B \|f\|_\beta V(x)^{\beta+k/r}, \quad (15a)$$

$$|F_k(\alpha) - F_k(\alpha')| \leq |\alpha - \alpha'| A_{k+1} B \|f\|_\beta V(x)^{\beta+(k+1)/r}, \quad (15b)$$

where we have used a Taylor expansion of  $F_k$  in (15b). Multiplying (15a) (to the power  $1 - \sigma$ ) by (15b) (to the power  $\sigma$ ) gives

$$|F_k(\alpha) - F_k(\alpha')| \leq 2|\alpha - \alpha'|^\sigma A B \|f\|_\beta V(x)^{\beta+(k+\sigma)/r},$$

which can be rewritten as  $\|P_{k,\alpha} f - P_{k,\alpha'} f\|_{\beta'} \leq 2AB|\alpha - \alpha'|^\sigma \|f\|_\beta$ . The first assertion of Lemma 1 is proved.

Next let  $k \in \{0, \dots, \lfloor r \rfloor - 1\}$ , let  $(\beta, \beta') \in [0, 1]^2$  be such that  $\beta + (k + 1)/r < \beta' \leq 1$ , and let  $0 < \sigma \leq 1$  be such that  $\beta + (k + 1 + \sigma)/r = \beta'$ . For any  $(\alpha, \alpha') \in (-a_0, a_0)^2$ , we obtain

$$|F_k(\alpha) - F_k(\alpha') - (\alpha - \alpha') F_{k+1}(\alpha')| \leq 2|\alpha - \alpha'| A_{k+1} B \|f\|_\beta V(x)^{\beta+(k+1)/r}, \quad (16a)$$

$$|F_k(\alpha) - F_k(\alpha') - (\alpha - \alpha') F_{k+1}(\alpha')| \leq \frac{|\alpha - \alpha'|^2}{2} A_{k+2} B \|f\|_\beta V(x)^{\beta+(k+2)/r}. \quad (16b)$$

Then, by multiplying (16a) (with the power  $1 - \sigma$ ) and (16b) (with the power  $\sigma$ ), we obtain

$$|F_k(\alpha) - F_k(\alpha') - (\alpha - \alpha') F_{k+1}(\alpha')| \leq 2|\alpha - \alpha'|^{1+\sigma} A B \|f\|_\beta V(x)^{\beta+(k+1+\sigma)/r},$$

which can be rewritten as  $\|P_{k,\alpha} f - P_{k,\alpha'} f - (\alpha - \alpha') P_{k+1,\alpha'} f\|_{\beta'} \leq 2AB|\alpha - \alpha'|^{1+\sigma} \|f\|_\beta$ .

## References

- [1] ALTMAN, E., AVRACHENKOV, K. E. AND NÚÑEZ-QUEJIA, R. (2004). Perturbation analysis for denumerable Markov chains with application to queueing models. *Adv. Appl. Prob.* **36**, 839–853.
- [2] BALADI, V. (2000). *Positive Transfer Operators and Decay of Correlations* (Adv. Ser. Nonlinear Dynamics **16**). World Scientific, River Edge, NJ.
- [3] BREYER, L., ROBERTS, G. O. AND ROSENTHAL, J. S. (2001). A note on geometric ergodicity and floating-point roundoff error. *Statist. Prob. Lett.* **53**, 123–127.
- [4] FERRÉ, D. (2011). A parametric first-order edgeworth expansion for markov additive functionals. Application to M-estimations. Preprint. Available at <http://hal.archives-ouvertes.fr/hal-00668894>.
- [5] GOUÉZEL, S. AND LIVERANI, C. (2006). Banach spaces adapted to Anosov systems. *Ergodic Theory Dynam. Systems* **26**, 189–217.

- [6] GUIBOURG, D., HERVÉ, L. AND LEDOUX, J. (2012). Quasi-compactness of Markov kernels on weighted-supremum spaces and geometrical ergodicity. Preprint. Available at <http://arxiv.org/abs/1110.3240v5>.
- [7] HEIDERGOTT, B. AND HORDIJK, A. (2003). Taylor series expansions for stationary Markov chains. *Adv. Appl. Prob.* **35**, 1046–1070.
- [8] HENNION, H. (1993). Sur un théorème spectral et son application aux noyaux lipchitziens. *Proc. Amer. Math. Soc.* **118**, 627–634.
- [9] HERVÉ, L. AND PÈNE, F. (2010). The Nagaev–Guivarc’h method via the Keller–Liverani theorem. *Bull. Soc. Math. France* **138**, 415–489.
- [10] KARTASHOV, N. V. (1981). Strongly stable Markov chains. In *Problems of Stability of Stochastic Models* (Panevezhis, 1980), Vsesoyuz. Nauch.-Issled. Inst. Sistem. Issled., Moscow, pp. 54–59.
- [11] KARTASHOV, N. V. (1996). *Strong Stable Markov Chains*. VSP, Utrecht.
- [12] KELLER, G. (1982). Stochastic stability in some chaotic dynamical systems. *Monatsh. Math.* **94**, 313–333.
- [13] KELLER, G. AND LIVERANI, C. (1999). Stability of the spectrum for transfer operators. *Ann. Scuola Norm. Sup. Pisa Classe Sci.* **28**, 141–152.
- [14] LIVERANI, C. (2004). Invariant measure and their properties. a functional analytic point of view. In *Dynamical Systems, Part II*, Nuove Pubblicazioni della Classe di Scienze, Scuola Normale Superiore.
- [15] MEYN, S. P. AND TWEEDIE, R. L. (1993). *Markov Chains and Stochastic Stability*. Springer, London.
- [16] MEYN, S. P. AND TWEEDIE, R. L. (1994). Computable bounds for geometric convergence rates of Markov chains. *Ann. Appl. Prob.* **4**, 981–1011.
- [17] MOUHOUBI, Z. AND AÏSSANI, D. (2010). New perturbation bounds for denumerable Markov chains. *Linear Algebra Appl.* **432**, 1627–1649.
- [18] ROBERTS, G. O., ROSENTHAL, J. S. AND SCHWARTZ, P. O. (1998). Convergence properties of perturbed Markov chains. *J. Appl. Prob.* **35**, 1–11.
- [19] SHARDLOW, T. AND STUART, A. M. (2000). A perturbation theory for ergodic Markov chains and application to numerical approximations. *SIAM J. Numer. Analysis* **37**, 1120–1137.
- [20] WU, L. (2004). Essential spectral radius for Markov semigroups. I. Discrete time case. *Prob. Theory Relat. Fields* **128**, 255–321.