



Chernoff and Berry-Esséen inequalities for Markov Processes

Pascal Lezaud

► To cite this version:

Pascal Lezaud. Chernoff and Berry-Esséen inequalities for Markov Processes. ESAIM: Probability and Statistics, 2001, 5, pp 183-201. 10.1051/ps:2001108 . hal-00940906

HAL Id: hal-00940906

<https://enac.hal.science/hal-00940906>

Submitted on 1 Apr 2014

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.

CHERNOFF AND BERRY–ESSÉEN INEQUALITIES FOR MARKOV PROCESSES

PASCAL LEZAUD^{1,2}

Abstract. In this paper, we develop bounds on the distribution function of the empirical mean for general ergodic Markov processes having a spectral gap. Our approach is based on the perturbation theory for linear operators, following the technique introduced by Gillman.

Mathematics Subject Classification. 60F10.

Received February 19, 2001. Revised September 21 and October 1, 2001.

1. INTRODUCTION

Let (X_t) be an ergodic irreducible Markov process on a general state space E , whose stationary distribution is denoted by π . Then for any measurable bounded function f and any initial distribution q , the weak law of large numbers states that the empirical mean $t^{-1}S_t := t^{-1} \int_0^t f(X_s)ds$ converges in probability to $\pi(f) := \int f(y)\pi(dy)$. This result is the basis of the simulation method to evaluate the mean of the function f . In this paper, we will quantify the rate of convergence by studying the probability

$$P_q \left\{ |t^{-1}S_t - \pi(f)| \geq \gamma \right\},$$

where P_q denotes the probability measure of the process with the initial distribution q , and the size of deviation γ is a small number.

The aim of this work is to obtain some bounds usable to decide when exactly the simulation algorithm should be stopped. This bound will be useful if it is accurate and easy to compute. Even if the initial distribution is the stationary one, the problem of the convergence control subsists, since a confidence interval must be determined with a fixed level $0 < \alpha < 1$, such that $P_q \left\{ |t^{-1}S_t - \pi(f)| \geq u(\alpha) \right\} \leq \alpha$. The determination of a sufficient time for the simulation, with fixed α and γ , will be obtained by two different ways. The first one consists in bounding the quantity $P_q \left\{ |t^{-1}S_t - \pi(f)| > \gamma \right\}$, the second one uses the central limit theorem which claims that $t^{-1/2}(S_t - \pi(f))$ converges in law to a centered Gaussian with variance

$$\sigma_f^2 := \lim_{t \rightarrow \infty} t^{-1} \text{Var}_\pi(S_t) > 0.$$

The first bound is called Chernoff's bound, whereas the second method refers to a Berry–Esséen's bound.

Keywords and phrases: Markov process, Chernoff bound, Berry–Esséen, eigenvalues, perturbation theory.

¹ Centre d'Études de la Navigation Aérienne, 31055 Toulouse Cedex, France; e-mail: pascal.lezaud@cena.fr

² Université Paul Sabatier, 31055 Toulouse Cedex, France.

There are many works on deviation inequalities for Markov chains in the literature. An example is the large deviation principle theorem of Donsker and Varadhan [7] which involves the logarithmic spectral radius of the Feynman–Kac operator in the rate function. This result was simplified in [11]. However, asymptotic results are not satisfactory if one wants to achieve bounds that are useful for fixed time. In [16] the saddle point method is used to obtain the Edgeworth expansions of the deviation probability. Finally, an approach based on the concentration inequalities developed by [22] and [24] gives deviation inequalities for Markov chains.

Another approach based on the interpretation of the Feynman–Kac operator as a linear perturbation of the Markov kernel has been introduced by Nagaev [23] and Gillman [14]. Improvements of the result of Gillman was obtained in [9] for a finite state and reversible Markov chains and in [20] for finite state Markov chains and Markov processes. As the Feynman–Kac operator and its logarithmic spectral radius also is involved in the context of continuous-time processes [7], the perturbation interpretation is pertinent also in this context. So by using the perturbation theory for closed operator, we firstly improve the previous bound given in [20] and extend it to Markov processes with spectral gap. Our main result is the following Chernoff's bound, which yields a Gaussian behavior for the small values of γ and a Poissonian behavior for the large ones:

Theorem 1.1. *Let P_t an ergodic Markov semigroup with invariant probability measure π . Let assume that its infinitesimal generator L has a simple isolated eigenvalue $\lambda = 0$ and that the initial distribution q has a $L^2(\pi)$ -density relatively to the measure π . Then, for all $f \in D_2(L)$ such that $\pi(f) = 0$, $\|f\|_\infty \leq a$, for all $t > 0$ and all $0 < \gamma$,*

$$P_q \{t^{-1}S_t \geq \gamma\} \leq N_q \exp \left\{ -\frac{2t\gamma^2}{\sigma_f^2 \left(1 + \sqrt{1 + 4a\gamma/(\lambda_1 \sigma_f^2)}\right)^2} \right\}, \quad (1.1)$$

where λ_1 is the spectral gap of $(L + L^*)/2$ and N_q is the $L^2(\pi)$ -norm of the density of q related to the stationary distribution π .

Remark 1.2. If $\|f\|_2^2 \leq b^2$, the asymptotic variance of S_t is bounded by $2b^2/\lambda_1$, so in that case, we can replace in the inequality (1.1), the asymptotic variance by its upper bound. Therefore, if $\gamma \leq b^2/a$, we obtain a Gaussian bound since

$$P_q \{t^{-1}S_t \geq \gamma\} \leq N_q \exp \left\{ -\frac{\gamma^2 \lambda_1 t}{4b^2} \left(1 - \frac{a\gamma}{b^2}\right) \right\},$$

whereas for $\gamma > b^2/a$, we obtain a Poissonian bound

$$P_q \{t^{-1}S_t \geq \gamma\} \leq N_q \exp \left\{ -\frac{\gamma \lambda_1 t}{4} \left(1 - \frac{b^2}{a\gamma}\right) \right\}.$$

Remark 1.3. Obviously, we can take $b^2 = a^2$ in inequality (1.1), which gives

$$P_q \{t^{-1}S_t \geq \gamma\} \leq N_q \exp \left\{ -\frac{\lambda_1 t \gamma^2}{4a^2 \left(\sqrt{1 + 2\gamma/a} + 1\right)^2} \right\} \leq N_q \exp \left(-\frac{\lambda_1 t \gamma^2}{8a^2} \right). \quad (1.2)$$

Remark 1.4. In [1], the authors express their scepticism about the implications of the estimates of spectral gap for finite-time behavior of non-reversible Markov processes. For example, Dinwoodie, in [10], shows that for counterclockwise random walk on the circle $\mathbb{Z}/p\mathbb{Z}$, very much is lost by symmetric reversibilization. Some recent results to include non-reversible effects have been obtained in [10] and [8].

The central limit theorem was established by Doeblin for positive recurrent Markov chains by using a renewal approach [6] (I.16, Th. 1). In [18] Landers and Rogge derived a bound of the error for the central limit theorem for Markov chains. This result was improved by Bolthausen in [3]. Perturbation theory of linear operators is a fruitful method also to achieve explicit Berry–Esséen’s bound. In [21] Mann used this approach to obtain an explicit Berry–Esséen’s bound for discrete Markov chains. In this paper, we improve and extend the result of Mann to continuous-time Markov processes. Our result is the following Berry–Esséen’s bound:

Theorem 1.5. *Referring to the setting of Theorem 1.1, let assume futhermore that the Markov semigroup is reversible. Then, for all $f \in D_2(L)$ such that $\pi(f) = 0$, $\|f\|_\infty \leq a$, for all $t > 0$ and all $0 < \gamma$,*

$$\left| P_q \left\{ \frac{1}{\sigma_f \sqrt{t}} S_t \leq x \right\} - F(x) \right| \leq N_q \frac{26a}{\lambda_1 \sigma_f \sqrt{t}} + \frac{N_q}{\pi} e^{-\lambda_1 t}, \quad (1.3)$$

where $F(x)$ is the standard Gaussian distribution.

Example 1.6 (Ornstein–Uhlenbeck process). Let $\gamma_n(dx)$ be the standard Gauss measure on $\mathbb{R}^n$ and let consider the Ornstein–Uhlenbeck operator $L(f) = \Delta(f) - x \cdot \nabla(f)$, where Δ is the Laplace operator and ∇ the gradient operator. For all n -tuples $k = (k_1, \dots, k_n)$ with $k_i \in \mathbb{N}$, the Hermite polynomials

$$H_k(x) = \Pi_{i=1}^n h_{k_i}(x) / \sqrt{k_i!}, \quad \text{with } h_k(x) = (-1)^k e^{x^2/2} \frac{d^k}{dx^k} e^{-x^2/2},$$

form a complete system of orthonormal eigenfunctions in $L^2(\gamma_n)$ for the operator L with eigenvalue $k_1 k_2 \cdots k_n$. So the Theorems 1.1 and 1.5 apply with $\lambda_1 = 1$ only for bounded functions f .

Example 1.7 (Continuous-time Markovian jump processes). Let consider a positive-recurrent continuous-time Markovian jump process with measurable state space $(E, \mathcal{E})$, transition kernel $J(x, dy)$ and reversible measure π . The transition rates are supposed essentially bounded, *i.e.*,

$$\pi - \text{ess sup}_x J(x, \{x\}^c) \leq M < \infty.$$

Then the infinitesimal generator L of this jump process:

$$Lf(x) = \int (f(x) - f(y)) J(x, dy),$$

is selfadjoint. Bounds on the spectral gap λ_1 are derived in [19] by a generalization of Cheeger’s inequality. More precisely, the Cheeger’s isoperimetric constant is

$$k \equiv \inf_{\substack{A \in \mathcal{E} \\ 0 < \pi(A) < 1}} k(A),$$

where

$$k(A) \equiv \frac{\int I_A(x) J(x, A^c) \pi(dx)}{\pi(A) \pi(A^c)} = \frac{\langle f, g \rangle_{L^2(\pi)}}{\pi(A) \pi(A^c)}.$$

Then [19] (Th. 2.1)

$$\kappa k^2 / 8M \leq \lambda_1 \leq k,$$

where κ is an explicit positive constant such that $\kappa \geq 1$. So the Theorems 1.1 and 1.5 apply with $\lambda_1 \geq k^2 / 8M$ only for bounded functions f .

Example 1.8 (Diffusion processes). Let M be a d -dimensional, connected, compact Riemannian manifold. Let Δ denote the Laplace–Beltrami operator, and consider the associated heat flow semigroup P_t , which is reversible with respect to the normalized Riemann measure μ . The spectrum of Δ is discrete and each eigenprojection is a finite-rank operator [13], so Theorems 1.1 and 1.5 can be applied.

The previous theorem applications can be expanded to the framework of general diffusion processes with spectral gap. For instance, let M be a connected, noncompact, complete Riemannian manifold and let consider the elliptic operator $L = \Delta + \nabla V$ for some $V \in C^2(M)$ with $\exp[V]$ integrable with respect to the Riemannian volume measure. In [26], Wang has given conditions for the existence of the spectral gap of L in $L^2(\mu)$, where $d\mu = Z^{-1} \exp[V]dx$, with $Z := \int_M \exp[V]dx$. Estimation of spectral gap for elliptic operators was studied by [5].

A short description of the paper is as follows. Section 2 of this paper proves Theorem 1.1 for finite-state Markov process. We will prove all results without using the large deviation theory, since we wish a self-contained article. In Section 3, by following [21], we obtain a Berry–Esséen’s bound for finite state Markov process. Nevertheless, by using estimates on eigenvalues and eigenvectors given in [17], we simplify the proof of Mann and improve the numerical constant by a factor of 400. Section 4 shows how a Chernoff’s bound and a Berry–Esséen’s bound can be achieved for a general-state Markov processes having a spectral gap. Finally, Annexe A sets out preliminaries on the perturbation theory of linear operator. In particular, we explain why the estimates on eigenvalues which have been obtained in finite-dimensional vector space, may be extended to closed operator with isolated eigenvalue. Always following the Kato’s approach, we show that these estimates can be used for semigroups with spectral gap which are perturbed by addition of a bounded operator.

2. CHERNOFF-TYPE BOUND FOR A FINITE-STATE MARKOV PROCESSES

This section proves Theorem 2.4, stated in the introduction, for a finite-state Markov process. So that, we consider an irreducible time homogeneous Markov process (X_t) on a finite-state space E . Its infinitesimal generator, denoted by L , defines an operator acting on $\ell^2(\pi)$ endowed with the inner product

$$\langle f, g \rangle = \sum_{x \in E} f(x)g(x)\pi(x),$$

by

$$Lf(x) = \sum_{y \in E} L(x, y)f(y).$$

Throughout this paper, $\|\cdot\|$ will always denote the $\ell^2(\pi)$ – norm. We first proves the following lemma:

Lemma 2.1. *Referring to the setting of Theorem 2.4, let $r > 0$. Then for any $t > 0$,*

$$P_q \{t^{-1}S_t > \gamma\} \leq e^{-rt\gamma} \langle q/\pi, \exp(-\Lambda(r)t).1 \rangle, \quad (2.1)$$

where $\Lambda(r) = -L - rM_f$, M_f being the multiplicative operator by the function f , i.e. $M_f g(x) := f(x)g(x)$.

Proof. The Markov process being on finite-state space, we may assume that for each ω the sample function $t \rightarrow X_t(\omega)$ is right continuous and has left hand limits everywhere in $[0, \infty)$. So, $t^{-1}S_t$ is the limit of Riemann sums:

$$t^{-1}S_t = \lim_{k \rightarrow \infty} k^{-1} \sum_{i=1}^k f(X_{it/k}).$$

Now, let consider the Markov kernel $P_{t/k}$ for k fixed. Then for any $r > 0$, by Chebyshev inequality

$$\begin{aligned} P_q \left\{ k^{-1} \sum_{i=1}^k f(X_{it/k}) > \gamma \right\} &\leq e^{-rt\gamma} E_q \left[\exp \left(k^{-1} \sum_{i=1}^k rt f(X_{it/k}) \right) \right] \\ &\leq e^{-rt\gamma} \left\langle q/\pi, P_{t/k}^k(r) \cdot \mathbf{1} \right\rangle, \end{aligned}$$

where $P_{t/k}(r)g(x) := \sum_y P_{t/k}(x, y) \exp(rt k^{-1} f(y)) g(y)$. This operator can be written

$$P_{t/k}(r) = \exp(Lt/k) \exp(rM_f t/k).$$

We conclude the proof by using Fatou lemma and the following result, available for all matrices A and B :

$$\lim_{k \rightarrow \infty} \left(\exp(A/k) \exp(B/k) \right)^k = \exp(A + B).$$

□

Lemma 2.2. *For any $r > 0$,*

$$P_q \{ t^{-1} S_t > \gamma \} \leq N_q \exp \{ -t(r\gamma + \lambda_0(r)) \}, \quad (2.2)$$

where $\lambda_0(r)$ is the smallest eigenvalue of selfadjoint operator

$$\tilde{\Lambda}(r) := -(L + L^*)/2 - rM_f.$$

Proof. Let $P_t(r)$ be the semigroup generated by $-\Lambda(r)$. This Feynman-Kac operator acts on $\ell^2(\pi)$ this way

$$P_t(r)g(x) = E_x(\exp(rS_t)g(X_t)).$$

Set $\varphi(t) = \|P_t(r)g\|^2$. Then

$$\varphi'(t) = -2 \left\langle \tilde{\Lambda}(r)P_t(r) \cdot g, P_t(r) \cdot g \right\rangle.$$

Furthermore, the selfadjointness of $\tilde{\Lambda}(r)$ implies that all its eigenvalues are real, so that, by the minimax principle the smallest eigenvalue of $\tilde{\Lambda}(r)$ may be characterized by

$$\lambda_0(r) = \inf \left\{ \left\langle \tilde{\Lambda}(r)g, g \right\rangle / \|g\|^2, g \neq 0 \right\} = \inf \left\{ \langle Lg, g \rangle - r\pi(fg^2) : \|g\|^2 = 1 \right\}.$$

This gives the inequality $\varphi'(t) \leq -2\lambda_0(r)\|P_t(r)g\|^2$. Thus the function $\varphi(t) \exp(2\lambda_0(r)t)$ is non increasing and it follows that

$$|P_t(r)| \leq \exp(-\lambda_0(r)t), \quad (2.3)$$

since $\varphi(0) = \|g\|^2$. We conclude by applying the Cauchy-Schwarz inequality to (2.1) combining with the previous result. □

Thus, bounding the distribution function of the empirical mean will be achieve by estimating the perturbed eigenvalue $\lambda_0(r)$. Actually, $\lambda_0(r)$ is the perturbation of the simple eigenvalue 0 of $(L + L^*)/2$. So, we can directly

use the identities (A.8) in the setting of selfadjoint operator with linear perturbation. More precisely, if λ_1 is the smallest positive eigenvalue of $-(L + L^*)/2$, then for any $|r| < \lambda_1/2$

$$\lambda_0(r) = - \sum_{n=1}^{\infty} \lambda_0^{(n)} r^n \quad (2.4)$$

$$\lambda_0^{(n)} = -\frac{1}{n} \sum_{k_1 + \dots + k_n = n-1} \text{tr} \left[M_f S^{(k_1)} M_f S^{(k_2)} \dots M_f S^{(k_n)} \right], \quad (2.5)$$

the sum being taken on all combinations of positive integers $k_1, \dots, k_n$ such that $k_1 + \dots + k_n = n - 1$. Here S denotes the reduced resolvent of $-(L + L^*)/2$ with respect the eigenvalue $\lambda_0 = 0$ and

$$S^{(0)}g = -\pi(g), \quad S^{(k)}g = S^k g, \quad g \in \ell^2(\pi).$$

This operator may be expressed as the limit of the resolvent $R(\zeta)$ of $-(L + L^*)/2$ as $\zeta \rightarrow 0$. We obtain the following identity

$$S = \int_0^\infty \exp \{t(L + L^*)/2\} dt.$$

For instance, if L is selfadjoint then S is the fundamental matrix of the Markov process

$$Z := \int_0^\infty P_t dt.$$

Lemma 2.3. *For any $r < \lambda_1/(2\|f\|_\infty)$,*

$$\lambda_0(r) \geq - \left(\frac{\sigma_f^2 \lambda_1^2}{2\|f\|_\infty^2} \right) \Phi \left(\frac{\|f\|_\infty r}{\lambda_1} \right) \geq - \frac{r^2 \sigma_f^2 / 2}{1 - 2 \left(\frac{\|f\|_\infty}{\lambda_1} \right) r},$$

where $\sigma_f^2 := 2\langle f, Sf \rangle$ and

$$\Phi(x) = \left(\frac{1-x}{2} \right) \left\{ 1 - \sqrt{1 - \frac{4x^2}{(1-x)^2}} \right\} \quad \text{for } -1 \leq x \leq 1/3.$$

Proof. Define the following equivalence relation on the nonnegative sequences $(k_1, \dots, k_n)$ by: $(k_1, \dots, k_n) \equiv (m_1, \dots, m_n)$ iff there exists a circular permutation which maps the first sequence on the second one. Denote by $[\cdot]$ the corresponding equivalence classes and notice that at least one $k_i = 0$ when the sequence is such that $k_1 + \dots + k_n = n - 1$. In this case, we will choose the sequence $(0, k_1, \dots, k_{n-1})$ as lifting of $[0, k_1, \dots, k_{n-1}]$. The well-known identity $\text{tr}(AB) = \text{tr}(BA)$, available for all matrices A and B , and a straightforward calculation allow us to rewrite (2.5) by

$$\lambda_0^{(n)} = \sum_{[0, k_1, \dots, k_{n-1}]} \left\langle f, S^{(k_1)} M_f \dots M_f S^{(k_{n-1})} f \right\rangle. \quad (2.6)$$

For instance, we have

$$\begin{aligned} \lambda_0^{(1)} &= 0, & \lambda_0^{(2)} &= \langle f, Sf \rangle, \\ \lambda_0^{(3)} &= -\langle f, (SM_f S)f \rangle, & \lambda_0^{(4)} &= \langle f, (SM_f)^2 Sf \rangle - \langle f, Sf \rangle \langle Sf, Sf \rangle. \end{aligned}$$

Notice that for a reversible Markov process, the coefficient $2\lambda_0^{(2)}$ is nothing else than the asymptotic variance of S_t , *i.e.*

$$2\lambda_0^{(2)} = \lim_{t \rightarrow \infty} t^{-1} \text{Var}_\pi(S_t) := \sigma_f^2.$$

Now, we go back to study each coefficient $\lambda^{(n)}$. For that, notice that only the equivalence classes, different from $[0, k_1, \dots, k_{n-2}, 0]$, contribute to the sum in (2.6). Accordingly, the number of nonzero terms in the sum is given by

$$\beta_n := \sum_{m=1}^{\lceil n/2 \rceil} \frac{1}{m} \binom{n-m-1}{m-1} \binom{n-2}{n-m-1},$$

where m is precisely the number of (non adjacent) zeros in the sequence $(0, k_1, \dots, k_{n-1})$. The generating function of the β_n is

$$\Phi(x) = \left(\frac{1-x}{2} \right) \left\{ 1 - \sqrt{1 - \frac{4x^2}{(1-x)^2}} \right\} \quad \text{for } -1 \leq x \leq 1/3.$$

Therefore, equation (2.6) may be rewritten as

$$\lambda_0^{(n)} = \sum_{m=1}^{\lceil n/2 \rceil} \frac{(-1)^m}{m} \binom{n-m-1}{m-1} \binom{n-2}{n-m-1} \lambda_m^{(n)},$$

where

$$\begin{aligned} \lambda_m^{(n)} = & \sum \langle f, S^{k_1} M_f S^{k_2} M_f \dots M_f S^{k_{i_1}} f \rangle \langle f, S^{k_{i_1}+2} M_f \dots M_f S^{k_{i_1}+i_2+1} f \rangle \\ & \dots \langle f, S^{k_{i_1}+\dots+i_{m-1}+m} M_f \dots M_f S^{k_{n-1}} f \rangle, \end{aligned}$$

the sum being taken on all sequences

$$(0, k_1, \dots, k_{i_1}, 0, k_{i_1+2}, \dots, k_{i_1+i_2+1}, 0, k_{i_1+i_2+3}, \dots, 0, k_{i_1+\dots+i_{m-1}+m}, \dots, k_{n-1}),$$

so that each $k_i \geq 1$, $\sum k_i = n-1$ and $i_1 + i_2 + \dots + i_m + m = n$ (m being the number of zeros in the sequence).

The challenge is now to obtain an accurate upper bound on each term $\langle f, S^{k_1} M_f S^{k_2} M_f \dots M_f S^{k_{i_1}} f \rangle$. Before using the Cauchy–Schwarz inequality, let us introduce the square root operator $S^{1/2}$ of the positive operator S . If $\|T\|_{p,q}$ denotes the norm of the linear operator T , from $\ell^p(\pi)$ into $\ell^q(\pi)$, then $\|S^{1/2}\|_{2,2} = \frac{1}{\sqrt{\lambda_1}}$. Let $\sigma_f^2 = 2\|S^{1/2}f\|^2$ which is the asymptotic variance for a reversible Markov process. By this way, we obtain the following bound

$$\begin{aligned} \langle f, S^{k_1} M_f S^{k_2} M_f \dots M_f S^{k_{i_1}} f \rangle & \leq \|S^{1/2}f\| \|S^{k_1-1/2}\|_{2,2} \|f\|_\infty \|S^{k_2} M_f \dots M_f S^{k_{i_1}} f\| \\ & \leq \frac{\sigma_f^2}{2} \left(\frac{1}{\lambda_1} \right)^{k_1+\dots+k_{i_1}-1} \|f\|_\infty^{i_1-1}. \end{aligned} \quad (2.7)$$

Hence,

$$\begin{aligned} |\lambda_0^{(n)}| &\leq \sum_{m=1}^{\lceil n/2 \rceil} \frac{1}{m} \binom{n-m-1}{m-1} \binom{n-2}{n-m-1} \left(\frac{\sigma_f^2}{2}\right)^m \|f\|_\infty^{n-2m} \|S\|_{2,2}^{n-1-m} \\ &\leq \left(\frac{\|f\|_\infty^n}{\lambda_1^{n-1}}\right) \left(\frac{\sigma_f^2 \lambda_1}{2\|f\|_\infty^2}\right) \beta_n, \end{aligned}$$

since $\sigma_f^2 \lambda_1 / (2\|f\|_\infty^2) \leq 1$.

To conclude the proof of the Lemma 2.3, we use the generating function $\Phi(x)$ and the inequality $\Phi(x) \leq x^2/(1-2x)$. $\square$

Theorem 2.4. *Let (P_t, π) be an irreducible Markov process on a finite set E and L its infinitesimal generator. Let $f : E \rightarrow \mathbb{R}$ be such that $\pi(f) = 0$, $\|f\|_\infty \leq a$. Then, for any initial distribution q , all $t > 0$ and all $0 < \gamma \leq a$,*

$$P_q \{t^{-1}S_t \geq \gamma\} \leq N_q \exp \left\{ -\frac{2t\gamma^2}{\sigma_f^2 \left(1 + \sqrt{1 + 4a\gamma/(\lambda_1 \sigma_f^2)}\right)^2} \right\} \quad (2.8)$$

where $N_q = \|q/\pi\|_2$ and λ_1 is the smallest positive eigenvalue of $-(L + L^*)/2$.

The proof of Theorem 2.4 is now easy, since it is enough to maximize the function

$$r \rightarrow r\gamma - \frac{r^2 \sigma_f^2 / 2}{1 - 2 \left(\frac{\|f\|_\infty}{\lambda_1}\right) r}.$$

This optimum is reached for

$$r_0 = \frac{\lambda_1}{2\|f\|_\infty} \left(1 - \left(1 + \frac{4\gamma\|f\|_\infty}{\lambda_1 \sigma_f^2} \right)^{-1/2} \right),$$

hence the following inequality

$$P_q \{t^{-1}S_t \geq \gamma\} \leq N_q \exp \left\{ -\frac{t\lambda_1^2 \sigma_f^2}{8\|f\|_\infty^2} \left(\sqrt{1 + \frac{4\|f\|_\infty \gamma}{\lambda_1 \sigma_f^2}} - 1 \right)^2 \right\}.$$

Remark 2.5. The numbers β_n are known as the Motzkin numbers M_n . More precisely, the Motzkin number generating function is $M(x) = (1 - x - \sqrt{1 - 2x - 3x^2})/(2x^2)$, so $\beta_{n+2} = M_n$. The Motzkin numbers enumerate various combinatorial objects. In particular, they give the number of paths from $(0,0)$ to $(0,n)$ which never dip below $y = 0$ and are made up only of the steps $(1,0), (1,1), (1,-1)$, i.e. $\rightarrow, \nearrow, \searrow$. The first are 1, 2, 4, 9, 21, 51, $\dots$. We have the recurrence relation

$$M_n = M_{n-1} + \sum_{k=0}^{n-2} M_k M_{n-2-k}$$

with $M_0 = 1$. The Motzkin number M_n is also given by

$$\begin{aligned} M_n &= -\frac{1}{2} \sum_{a+b=n+2, a, b \geq 0} (-3)^a \binom{1/2}{a} \binom{1/2}{b} \\ &= \frac{(-1)^{n+1}}{2^{2n+5}} \sum_{a+b=n+2, a, b \geq 0} \frac{(-3)^a}{(2a-1)(2b-1)} \binom{2a}{a} \binom{2b}{b}. \end{aligned}$$

Remark 2.6. A slightly more accurate Chernoff's bound can be obtained by using the function Φ instead. In this case, we get

$$P_q \{t^{-1}S_t \geq \gamma\} \leq N_q \exp \left\{ -\frac{t\lambda_1 b^2}{a^2} h(a\gamma/b^2) \right\},$$

where h is an explicit function such that

$$\frac{y}{4} \log(1+y) - \frac{13y^4}{192} \leq h(y) \leq \frac{y}{4} \log(1+y).$$

Thus, the previous Chernoff's bound is not so far from the following expression

$$N_q \exp \left\{ -\frac{t\lambda_1 \gamma}{4a} \log \left(1 + \frac{a\gamma}{b^2} \right) \right\},$$

which looks like the Bennett's inequality achieved for independent random variables.

Example 2.7 (Markov process on two-points state-space). Let consider a Markov process on state-space $E = \{a, b\}$, so its infinitesimal generator is a matrix

$$L = \begin{pmatrix} -\lambda & \lambda \\ \mu & -\mu \end{pmatrix},$$

and the stationary reversible distribution is $\pi = (\mu/(\lambda + \mu), \lambda/(\lambda + \mu))$.

The eigenvalues of $-L$ are 0 and $\lambda_1 = \lambda + \mu$, and an eigenvector corresponding to the eigenvalue λ_1 is $\varphi = (\sqrt{\lambda/\mu}, -\sqrt{\lambda/\mu})$.

Let f be a function on E such that $\pi(f) = 0$, i.e. $\mu f(a) + \lambda f(b) = 0$. The asymptotic variance is

$$\sigma_f^2 = 2 \frac{\mu/\lambda}{\lambda + \mu} f(a)^2 = \frac{-f(a)f(b)}{\lambda + \mu} = \frac{|f(a)f(b)|}{\lambda + \mu}$$

and the fundamental matrix

$$Z = \frac{-1}{(\lambda + \mu)^2} L, \quad Z^{1/2} = \frac{-1}{(\lambda + \mu)^{3/2}} L.$$

Now, let consider the perturbed matrix $\Lambda(r) = -L - rf$, whose eigenvalues are

$$\lambda_{\pm}(r) = \frac{1}{2} \left((\lambda + \mu) - r(f(a) + f(b)) \pm \sqrt{\Delta} \right),$$

with $\Delta = (\lambda + \mu - r(f(a) + f(b)))^2 - 4r^2 f(a)f(b) \geq 0$.

Assume now that $r < (\lambda + \mu)/|f(a) + f(b)|$. Then, the eigenvalue $\lambda_0(r)$ may be written as

$$\lambda_0(r) = \frac{1}{2} ((\lambda + \mu) - r(f(a) + f(b))) \left(1 - \sqrt{1 - \frac{4r^2 f(a)f(b)}{(\lambda + \mu - r(f(a) + f(b)))^2}} \right).$$

We deduce the Taylor series expansion of $\lambda_0(r)$:

$$\lambda_0(r) = (\lambda + \mu) \sum_{n=2}^{\infty} \left(\frac{r}{\lambda + \mu} \right)^n \sum_{m=1}^{\lceil n/2 \rceil} \frac{1}{m} \binom{2(m-1)}{m-1} \binom{n-2}{2(m-1)} (f(a)f(b))^m (f(a) + f(b))^{n-2m}.$$

A comparison between this last expression and (2.4) gives

$$\begin{aligned} |\lambda_0^n| &\leq \frac{1}{(\lambda + \mu)^{n-1}} \sum_{m=1}^{\lceil n/2 \rceil} \frac{1}{m} \binom{2(m-1)}{m-1} \binom{n-2}{2(m-1)} \left(\frac{\mu}{\lambda} f(a)^2 \right)^m |f(a) + f(b)|^{n-2m} \\ &= \left(\frac{\sigma_f^2}{2} \right)^{n/2} \sum_{m=1}^{\lceil n/2 \rceil} \frac{1}{m} \binom{2(m-1)}{m-1} \binom{n-2}{2(m-1)} \frac{(\lambda/\mu)^{n/2-m}}{(\lambda + \mu)^{n/2-1}} |1 - \mu/\lambda|^{n-2m} \\ &\leq \left(\frac{\sigma_f^2}{2} \right)^{n/2} \|Z^{1/2}\|_{2,\infty}^{n-2} \beta_n \end{aligned}$$

where $\|Z^{1/2}\|_{2,\infty} = \max(\sqrt{\lambda/\mu}, \sqrt{\mu/\lambda})/\sqrt{\lambda + \mu}$.

Thus, we get that

$$\lambda_0(r) \geq -\frac{1}{\|Z^{1/2}\|_{2,\infty}^2} \Phi \left(\|Z^{1/2}\|_{2,\infty} \left(\frac{\sigma_f^2}{2} \right)^{1/2} r \right).$$

3. BERRY–ESSÉEN BOUND FOR A FINITE-STATE REVERSIBLE MARKOV PROCESSES

In this section, we assume that the Markov process is reversible for the stationary distribution π . To obtain a Berry–Esséen bound, we will follow the approach of [21], which is based on the following inequality (see (3.13) in [12], p. 538)

$$\left| P \left(S_t / (\sigma_f \sqrt{t}) \leq x \right) - F(x) \right| \leq \frac{1}{\pi} \int_{-W}^W \frac{|\varphi_q(\omega) - e^{-\omega^2/2}|}{|\omega|} d\omega + \frac{24}{\pi \sqrt{2\pi} W},$$

where $\varphi_q(\omega) = E_q(i\omega S_t / (\sigma_f t^{1/2}))$, $F(x)$ is the standard Gaussian distribution and W is a positive constant. Mann used perturbation theory to bound $|\varphi_q(\omega) - e^{-\omega^2/2}|$ and obtained a Berry–Esséen bound for Markov chains with finite or countably infinite state space. We extend its result to Markov processes and obtain a sharper bound by using Kato's perturbation theory.

Theorem 3.1. *Let (P_t, π) be an irreducible reversible Markov process on a finite set E and L its infinitesimal generator. Let $f : E \rightarrow \mathbb{R}$ be such that $\pi(f) = 0$, $\|f\|_\infty \leq a$. Then, for any initial distribution q , all $t > 0$,*

$$\left| P_q \left\{ \frac{1}{\sigma_f \sqrt{t}} S_t \leq x \right\} - F(x) \right| \leq N_q \frac{26a}{\lambda_1 \sigma_f \sqrt{t}} + \frac{N_q}{\pi} e^{-\lambda_1 t}, \quad (3.1)$$

where $F(x)$ is the standard Gaussian distribution and $\sigma_f = \lim_{t \rightarrow \infty} t^{-1} \text{Var}_\pi S_t$ is the asymptotic variance of S_t .

Let L be the infinitesimal generator of the reversible Markov process, so this operator is selfadjoint. For $\omega \in \mathbb{R}$, let $\chi = i\omega/(\sigma_f t^{1/2})$, where σ_f^2 is the asymptotic variance, and consider the perturbed operator $\Lambda(\omega) = -L + \chi M_f$. Using (A.8) gives the following Taylor series expansion for $\lambda_0(\chi)$ noted by $\lambda_0(\omega)$

$$\lambda_0(\omega)t = \frac{\omega^2}{2} + t \sum_{n \geq 3} \lambda_0^{(n)} (-1)^n \left(\frac{i\omega}{\sigma_f \sqrt{t}} \right)^n.$$

Now, assuming that $\|f\|_\infty \leq a$, we get the following inequalities valid if $|\omega| < \lambda_1 \sigma_f \sqrt{t}/(2a)$

$$\begin{aligned} |\lambda_0(\omega) - \frac{\omega^2}{2}| &\leq t \sum_{n \geq 3} |\lambda_0^{(n)}| \left(\frac{\omega}{\sigma_f \sqrt{t}} \right)^n \leq t \lambda_1 \sum_{n \geq 3} \left(\frac{a^n}{\lambda_1^{n-1}} \right) \left(\frac{\sigma_f^2 \lambda_1}{2a^2} \right) \beta_n \\ &= \lambda_1 t \left(\frac{\sigma_f^2 \lambda_1}{2a^2} \right) \left(\Phi \left(\frac{a\omega}{\lambda_1 \sigma_f \sqrt{t}} \right) - \left(\frac{a\omega}{\lambda_1 \sigma_f \sqrt{t}} \right)^2 \right) \leq \frac{a\omega^3}{\lambda_1 \sigma_f \sqrt{t}} \left(1 - \frac{2a\omega}{\lambda_1 \sigma_f \sqrt{t}} \right)^{-1}. \end{aligned}$$

We deduce that

$$\left| \exp(-\lambda_0(\omega)t) - e^{-\omega^2/2} \right| \leq e^{-\omega^2/2} \left(e^{H(|\omega|)} - 1 \right)$$

where

$$H(|\omega|) = \frac{a|\omega|^3}{\lambda_1 \sigma_f \sqrt{t}} \left(1 - \frac{2a|\omega|}{\lambda_1 \sigma_f \sqrt{t}} \right)^{-1}.$$

Let $u(\omega)$ be the eigenvector of $\Lambda(\omega)$ for the eigenvalue $\lambda_0(\omega)$ normalized by $\langle \mathbf{1}, u(\omega) \rangle = 1$. Moreover, we have

$$\begin{aligned} |\varphi_q(\omega) - e^{-\lambda_0(\omega)t}| &\leq N_q \|e^{-\Lambda(\omega)t} \mathbf{1} - e^{-\lambda_0(\omega)t} \mathbf{1}\| \\ &= N_q \|e^{-\Lambda(\omega)t} (\mathbf{1} - u(\omega)) - e^{-\lambda_0(\omega)t} (\mathbf{1} - u(\omega))\| \\ &\leq N_q \|\mathbf{1} - u(\omega)\| \left(e^{-\lambda_1 t} + e^{H(|\omega|) - \omega^2/2} \right), \end{aligned}$$

where we used the inequality (2.3) with

$$\lambda_0(\omega) = \inf \{ \langle (\Lambda(\omega) + \Lambda^*(\omega)/2)g, g \rangle / \|g\|^2, g \neq 0, \pi(g) = 0 \} = \inf \{ \langle -Lg, g \rangle / \|g\|^2, g \neq 0, \pi(g) = 0 \} = \lambda_1.$$

Getting an upper bound of $\|\mathbf{1} - u(\omega)\|$ is not so difficult, since the relation (A.11) gives here

$$\begin{aligned} u(\omega) - \mathbf{1} &= \chi (I - Z[\chi M_f + \lambda(\chi)])^{-1} Zf \\ \|\mathbf{1} - u(\omega)\| &\leq \frac{|\chi| \|Z^{1/2}\|_{2,2} \|Z^{1/2} f\|_2}{1 - |\chi| \|Z M_f\|_{2,2} - |\lambda_0(\chi)| \|Z\|_{2,2}} \leq \frac{|\chi| \left(\sigma_f^2 / (2\lambda_1) \right)^{1/2}}{1 - |\chi| a / \lambda_1 - \Phi(a|\chi|/\lambda_1)} \leq \frac{3|\chi| \sigma_f}{\sqrt{2\lambda_1}} = \frac{3\omega}{\sqrt{2\lambda_1 t}}, \end{aligned}$$

since $\Phi(x) \leq 1/3$ for $x \leq 1/3$. As $\sigma_f \leq a\sqrt{2/\lambda_1}$, we get $\|\mathbf{1} - u(\omega)\| \leq 3|\omega|a/(\lambda_1 \sigma_f t^{1/2})$. Now, assuming that $|\omega|a/(\lambda_1 \sigma_f t^{1/2}) \leq 1/6$ gives $H(|\omega|) \leq \omega^2/4$, thus

$$|\varphi_q(\omega) - e^{-\lambda_0(\omega)t}| \leq N_q \frac{3|\omega|a}{\lambda_1 \sigma_f t^{1/2}} \left(e^{-\lambda_1 t} + e^{-\omega^2/4} \right).$$

Let observe that

$$\begin{aligned} |\varphi_q(\omega) - e^{-\omega^2/2}| &\leq |\varphi_q(\omega) - e^{-\lambda_0(\omega)t}| + |\exp(-\lambda_0(\omega)t) - e^{-\omega^2/2}| \\ &\leq N_q \frac{3|\omega|a}{\lambda_1 \sigma_f t^{1/2}} \left(e^{-\lambda_1 t} + e^{-\omega^2/4} \right) + e^{-\omega^2/2} \left(e^{H(|\omega|)} - 1 \right), \end{aligned}$$

but $N_q \geq 1$ and for $x \geq 0$, $e^x - 1 \leq xe^x$, therefore

$$|\varphi_q(\omega) - e^{-\omega^2/2}| \leq N_q \frac{3|\omega|a}{\lambda_1 \sigma_f t^{1/2}} \left(e^{-\lambda_1 t} + e^{-\omega^2/4} (1 + \omega^2/2) \right).$$

Thus,

$$\left| P \left(S_t / (\sigma_f \sqrt{t}) \leq x \right) - F(x) \right| \leq N_q \frac{3a}{\pi \lambda_1 \sigma_f t^{1/2}} \int_{-W}^W \left(e^{-\lambda_1 t} + e^{-\frac{\omega^2}{4}} \left(1 + \frac{\omega^2}{2} \right) \right) d\omega + \frac{24}{\pi \sqrt{2\pi} W},$$

with $W = (\lambda_1 \sigma_f t^{1/2})/6a$. We obtain after some computations,

$$\left| P \left(S_t / (\sigma_f \sqrt{t}) \leq x \right) - F(x) \right| \leq N_q \frac{26a}{\lambda_1 \sigma_f \sqrt{t}} + \frac{N_q}{\pi} e^{-\lambda_1 t}.$$

Remark 3.2. It is possible to obtain an universal Berry–Esséen bound which does not depend on the function f ; however this bound is less accurate. For that, to avoid the term $\|f\|_\infty$ in the estimate of eigenvalue $\lambda_0(\omega)$ (see (2.7)), we use the quantity $\|Z^{1/2}\|_{2,\infty}$. So we get that

$$|\varphi_q(\omega) - e^{-\omega^2/2}| \leq N_q \frac{3|\omega| \|Z^{1/2}\|_{2,\infty}}{\sqrt{2t}} \left(e^{-\lambda_1 t} + e^{-\omega^2/4} (1 + \omega^2/2) \right),$$

thus

$$\left| P \left(S_t / (\sigma_f \sqrt{t}) \leq x \right) - F(x) \right| \leq N_q \frac{26 \|Z^{1/2}\|_{2,\infty}}{\sqrt{2t}} + \frac{N_q}{\pi} e^{-\lambda_1 t}.$$

For example, take back the case of the Markov process on two-points state-space. We have $\|Z^{1/2}\|_{2,\infty} = \max(\sqrt{\lambda/\mu}, \sqrt{\mu/\lambda})/\sqrt{\lambda+\mu}$. For simplicity, let assume that $\lambda \geq \mu$, we get that $\|f\|_\infty = |f(a)|$, so in this case the two bounds are the same;

$$N_q \frac{26\sqrt{\lambda/\mu}}{\sqrt{\lambda+\mu}\sqrt{2t}} + \frac{N_q}{\pi} e^{-\lambda_1 t}.$$

4. CHERNOFF AND BERRY–ESSÉEN INEQUALITIES FOR GENERAL MARKOV PROCESSES WITH SPECTRAL GAP

Let $(E, \mathcal{E}, \pi)$ a probability space. On E is given a Markov semigroup, *i.e.* a family $p_t(x, dy)$ of probability transitions kernels, defined for $t \geq 0$. This family is identified with the family of operators P_t defined on the space of Borel functions f on E which are bounded, where

$$P_t(f)(x) = \int f(y) p_t(x, dy).$$

In the following, we will assume that π is an invariant measure of P_t , *i.e.* that for every f on E

$$\int_E P_t(f)(x)\pi(dx) = \int_E f(x)\pi(dx).$$

The semigroup P_t can be extended into a contraction operator on the Hilbert space $L^2(\pi)$, and we define the infinitesimal generator L of P_t in $L^2(\pi)$ on the dense domain $D_2(L)$ by

$$L(f) = \lim_{t \rightarrow 0} \frac{P_t(f) - f}{t},$$

the domain being the subspace of the functions f for which such a limit exists.

As $P_t(\mathbf{1}) = \mathbf{1}$, we see that $\lambda = 0$ is an eigenvalue of L . In the following, we always assume that λ is a simple isolated eigenvalue of L , so that the Markov process is ergodic. As L is a closed operator densely defined, there is a unique maximal operator L^* adjoint of L (*i.e.* that any other adjoint of L is a restriction of L^*). Moreover, as $L^2(\pi)$ is a reflexive space, L^* is likewise closed and densely defined [17] (Th. 5.29 III, Sect. 5.5) and is the infinitesimal generator of a Markov semigroup (since the negative real axis belongs to the resolvent set of L^*). The operator $L + L^*$ is a selfadjoint closed operator densely defined on $D_2(L) \cap D_2(L^*)$. Using [17] (Th. 2.11 IX, Sect. 2.3), we claim that $(L + L^*)/2$ is likewise the infinitesimal generator of a Markov semigroup with a simple isolated eigenvalue $\lambda = 0$. The spectral gap is given by the infimum of the set of λ_1 for which $\|P_t(f)\| \leq e^{-\lambda_1 t}$ for all functions $f \in D_2(L)$ such that $\pi(f) = 0$. λ_1 is also the isolation distance of the isolated eigenvalue $\lambda = 0$ of the operator $(L + L^*)/2$.

By using results on the perturbation theory of closed operator with isolated eigenvalue, stated in Annex A, we now are able to extend the Chernoff-type bound obtained previously for a finite-state Markov process to the semigroup P_t . Firstly, if A is a bounded operator then $L + A$ is likewise the infinitesimal generator of a contraction semigroup and $e^{t(L-\chi A)}$ is an entire function of the complex variable χ [17] (Th. 2.1 IX, Sect. 2.1). Secondly, the Trotter product formula [25] states that for all $t > 0$

$$e^{t(L-rA)} = \lim_{n \rightarrow \infty} \left(e^{\frac{t}{n}L} e^{-\frac{t}{n}rA} \right)^n.$$

Thus, we get the Lemma 2.1 for the semigroup P_t and so the Theorem 1.1.

Remark 4.1. For simplicity, let assume furthermore that L is selfadjoint and let consider the case of a concentrated initial distribution on a point $x \in E$. Thus, we cannot apply the Theorem 1.1 directly. So, we suppose that the semigroup P_t has a continuous kernel $p_t(x, y)$ and is ultracontractive, *i.e.* for each $t > 0$, P_t is a contraction from $L^p(\pi)$ to $L^p(\pi)$ for each $p \in [1, \infty]$ and $\|P_t\|_{2,\infty} = C_t < \infty$ [15]. Then, the semigroup $P_t(r)$ is also ultracontractif, since for $g \in L^2(\pi)$ and $e_h(t) := \exp \int_0^t h(X_s)ds$

$$P_t(r)g(x) \leq E_x[e_{2rf}(t)]^{1/2} E_x[|g(X_t)|]^{1/2} \leq \sup_x E_x[e_{2rf}(t)]^{1/2} [\sup_{x,y} |p_t(x, y)|]^{1/2} \|g\|_2 \leq e^{r\|f\|_\infty t} C_{t/2} \|g\|_2.$$

So, $\|P_t(r)\|_{2,\infty} \leq C'_t$ and

$$\begin{aligned} \sup_x P_x \{t^{-1}S_t \geq \gamma\} &\leq e^{-r\gamma} \|P_t(r)\mathbf{1}\|_\infty \leq e^{-r\gamma} \|P_t(r)\|_{2,\infty} \leq e^{-r\gamma} \|P_{t_0}(r)\|_{2,\infty} \|P_{t-t_0}(r)\|_{2,2} \\ &\leq C'_{t_0} \exp \left\{ -\frac{2(t-t_0)\gamma^2}{\sigma_f^2 \left(1 + \sqrt{1 + 4a\gamma/(\lambda_1 \sigma_f^2)}\right)^2} \right\}, \end{aligned}$$

where $0 < t_0 < t$.

Now, let assume that the Markov semigroup is reversible. Then, the Berry–Esséen inequality obtained in finite-dimensional case can be extended the same way by using the Chernoff's inequality. Thus, we get Theorem 1.5. Unfortunately, we say nothing about the way to extend these inequalities for unbounded functions. Nevertheless in the following example, we obtain a Chernoff's inequality despite the unboundedness of the function f .

Example 4.2. Let the Ornstein–Uhlenbeck diffusion on $\mathbb{R}$ with drift $\theta \geq 0$, which is the solution of the stochastic differential equation

$$dX_t = -\theta X_t dt + dB_t,$$

where B_t is the standard Brownian motion on $\mathbb{R}$. It's invariant measure is the Gauss measure $\mathcal{N}(0, 1/(2\theta))$ and the generator is the operator $L(g) = g''/2 - \theta xg'$.

Let consider the unbounded function $f(x) = x^2$ and the operator $L(r)(g) = L(g) + rf g$. A straightforward calculation gives the perturbed eigenvalue

$$\lambda(r) = \frac{1}{2} \left(-\theta + \sqrt{\theta^2 - 2r} \right),$$

with $|2r| \leq \theta^2$. Furthermore, the spectral gap of L is $\lambda_1 = \theta$, the mean of f for the invariant measure is $1/(2\theta)$ and the asymptotic variance is $\sigma_f^2 = 1/(2\theta^3)$.

Using the Chernoff inequality, we get

$$P_q \left\{ t^{-1} \int_0^t X_s^2 ds - \frac{1}{2\theta} \geq \gamma \right\} \leq e^{-rt(\gamma + \frac{1}{2\theta})} E_q \left(\exp \left(r \int_0^t X_s^2 ds \right) \right) \leq N_q e^{-rt(\gamma + \frac{1}{2\theta}) - \lambda(r)} := N_q e^{-tQ(r)},$$

where $Q(r) = r(\gamma + \frac{1}{2\theta}) + \frac{1}{2}(-\theta + \sqrt{\theta^2 - 2r})$ is maximized when

$$r = \frac{1}{2} \left(\theta^2 - \frac{1}{\gamma + 1/(2\theta)} \right).$$

Simple computations give the inequality

$$P_q \left\{ t^{-1} \int_0^t X_s^2 ds - \frac{1}{2\theta} \geq \gamma \right\} \leq N_q \exp \left(-t \frac{\theta^2 \gamma^2}{2(\gamma + 1/(2|\theta|))} \right) = N_q \exp \left(-t \frac{\gamma^2}{2\sigma_f^2(1 + 2\lambda_1 \gamma)} \right).$$

We also obtain the following inequality

$$P_q \left\{ t^{-1} \int_0^t X_s^2 ds \geq \gamma \right\} \leq N_q \exp \left(-t \frac{(1 - 2\theta\gamma)^2}{8\gamma} \right),$$

which is in accordance with the large deviations principle proven in [4] with the good rate function

$$I(\gamma) = \frac{(1 - 2\theta\gamma)^2}{8\gamma}, \quad \text{if } \gamma > 0.$$

Bercu and Rouault obtained in [2] an asymptotic expansions of $e^{tI(\gamma)} P_0(S_t \geq \gamma t)$ for the Ornstein–Uhlenbeck process starting from 0.

APPENDIX A. PERTURBATION THEORY OF LINEAR OPERATORS

A.1. Finite-dimensional space

Let T be a linear operator on some finite-dimensional vector space X . The Laurent series of its resolvent $R(\zeta)$ at a simple eigenvalue λ takes the form ([17], p. 38)

$$R(\zeta) = -(\zeta - \lambda)^{-1}P + \sum_{n=0}^{\infty} (\zeta - \lambda)^n S^{n+1}, \quad (\text{A.1})$$

where P is the eigenprojection operator for the eigenvalue λ , and S , called the reduced resolvent of T with respect the eigenvalue λ , is the inverse of the restriction of $T - \lambda$ in the subspace $(I - P)X$. We deduce that P is the residue of $-R(\zeta)$ in λ , and

$$P = -\frac{1}{2\pi i} \int_{\Gamma} R(\zeta) d\zeta, \quad (\text{A.2})$$

where Γ is a positively-oriented small circle enclosing λ , but excluding other eigenvalues of T .

Consider a family of operator-valued functions with the form

$$T(\chi) = T + \chi T^{(1)} + \chi^2 T^{(2)} + \dots, \quad \chi \in \mathbb{C}. \quad (\text{A.3})$$

Then the resolvent $R(\zeta, \chi) = (T(\chi) - \zeta)^{-1}$ of $T(\chi)$ is analytic in the two variables ζ, χ in each domain in which ζ is not equal to any of the eigenvalues of $T(\chi)$ ([17], p. 66). So it can be expanded into the following power series in χ with coefficients depending on ζ :

$$R(\zeta, \chi) = R(\zeta) + \sum_{n=1}^{\infty} \chi^n R^{(n)}(\zeta) \quad (\text{A.4})$$

where each $R^{(n)}$ is an operator-valued function. This series is uniformly convergent if

$$\sum_{n=1}^{\infty} |\chi|^n \|T^{(n)} R(\zeta)\| < 1. \quad (\text{A.5})$$

Let $r(\zeta)$ be the value of $|\chi|$ such that the left member of (A.5) is equal to 1. Then (A.5) is satisfied for $|\chi| < r(\zeta)$. Let λ be one of the eigenvalues of $T = T(0)$ with multiplicity $m = 1$, and Γ be a positively-oriented circle, in the resolvent set of T , enclosing λ but no other eigenvalues of T . The series (A.4) is then uniformly convergent for $\zeta \in \Gamma$ if

$$|\chi| < r_0 = \min_{\zeta \in \Gamma} r(\zeta). \quad (\text{A.6})$$

In the special case in which X is a Hilbert space and T is normal (*i.e.* $T^*T = TT^*$), we get

$$\|R(\zeta)\| = 1/\text{dist}(\zeta, \Sigma(T)), \quad r_0 = \min_{\zeta \in \Gamma} \left(\frac{a}{\text{dist}(\zeta, \Sigma(T))} + c \right)^{-1}$$

for every ζ in the resolvent set of T , where $a = \|T^{(1)}\|$ and c is such that $\|T^{(n)}\| \leq ac^{n-1}$ for $n \geq 1$. If we choose

as Γ the circle $|\zeta - \lambda| = d/2$ where

$$d = \min_{\mu \in \Sigma(T) \setminus \{\lambda\}} |\lambda - \mu|,$$

we obtain $r_0 = (2ad^{-1} + c)^{-1}$.

The existence of the resolvent $R(\zeta, \chi)$, for $\zeta \in \Gamma$, implies that there are no eigenvalues of $T(\chi)$ on Γ . The operator

$$P(\chi) = -\frac{1}{2\pi i} \int_{\Gamma} R(\zeta, \chi) d\zeta \quad (\text{A.7})$$

is the eigenprojection for all the eigenvalues of $T(\chi)$ lying inside Γ . In particular ([17], p. 68) for all $|\chi|$ sufficiently small, we have

$$\dim P(\chi)X = \dim PX = 1.$$

Therefore, only the eigenvalue $\lambda(\chi)$ of $T(\chi)$ lies inside Γ , and $P(\chi)$ is the eigenprojection for this eigenvalue.

As the only eigenvalues of $T(\chi)P(\chi)$ are 0 and $\lambda(\chi)$, we will consider

$$\lambda(\chi) - \lambda = \text{tr}((T(\chi) - \lambda)P(\chi)).$$

Combining (A.7) and substitution for $R(\zeta, \chi)$ from (A.4) give the Taylor series expansion ([17], p. 79)

$$\lambda(\chi) - \lambda = -\frac{1}{2\pi i} \text{tr} \int_{\Gamma} (\zeta - \lambda) R(\zeta, \chi) d\zeta = \sum_{n=1}^{\infty} \chi^n \lambda^{(n)}$$

where

$$\lambda^{(n)} = \sum_{p=1}^n \frac{(-1)^p}{p} \sum_{\substack{\nu_1 + \dots + \nu_p = n \\ k_1 + \dots + k_p = p-1 \\ \nu_i \geq 1, k_j \geq 0}} \text{tr} [T^{\nu_1} S^{(k_1)} \dots T^{\nu_p} S^{(k_p)}] \quad (\text{A.8})$$

with $S^{(0)} = -P$, $S^{(n)} = S^n$. Here S is the reduced resolvent of T with respect the eigenvalue λ .

An estimation on an eigenvector will be required to establish the Berry–Esséen bound. Since the eigenvectors are not uniquely determined, there are no definite formulas for the eigenvectors of $T(\chi)$ as functions of χ . Let again λ be one of the eigenvalue of T with multiplicity $m = 1$. Following ([17], Sect. II-3), a convenient form of the eigenvector $\varphi(\chi)$ of $T(\chi)$ corresponding to the eigenvalue $\lambda(\chi)$ is given by

$$\varphi(\chi) = (P(\chi)\varphi, \psi)^{-1} P(\chi)\varphi, \quad (\text{A.9})$$

where φ is an unperturbed eigenvector of T for the eigenvalue λ and ψ is an eigenvector of T^* for the eigenvalue $\bar{\lambda}$ normalized by $(\varphi, \psi) = 1$. The choice of the factor in (A.9) is equivalent to each of the following normalization conditions:

$$(\varphi(\chi), \psi) = 1, \quad (\varphi(\chi) - \varphi, \psi) = 0, \quad P(\varphi(\chi) - \varphi) = 0. \quad (\text{A.10})$$

Let $A(\chi) := T(\chi) - T$. Noting further that $S\varphi = 0$, we obtain the following convenient formula ([17], p. 92):

$$\begin{aligned} \varphi(\chi) - \varphi &= -(I + S[A(\chi) - \lambda(\chi) + \lambda])^{-1} S A(\chi) \varphi \\ &= -S [I + A(\chi)S - (\lambda(\chi) - \lambda)S]^{-1} A(\chi) \varphi. \end{aligned} \quad (\text{A.11})$$

A.2. Operators in Banach spaces

As we aim to study general state-space Markov processes, we will have to be able to estimate the lower bound of the spectrum of a perturbed infinitesimal generator of a Markovian semigroup. We saw in the previous section, the essential part of the spectral gap, *i.e.* the distance between the eigenvalue 0 and the other eigenvalues. So following [17], we will introduce the notion of the isolated eigenvalue and of holomorphic dependence of an operator on a parameter and we will show that the results obtained in the finite-dimensional case can be extended, at least for isolated eigenvalues, without essential modification.

Let T be an operator from X to Y , $D(T)$ its domain of definition and $R(T)$ its range. T is said to be closed if for each sequence $u_n \in D(T)$ such that $u_n \rightarrow u$ and $Tu_n \rightarrow v$, u belongs to $D(T)$ and $Tu = v$. The set of closed operators from X to Y will be denoted by $\mathcal{C}(X, Y)$. T is said to be decomposed according to $X = M \oplus N$, if

$$PD(T) \subset D(T), \quad TM \subset M, \quad TN \subset N,$$

where P is the projection on M along N . When T is decomposed as above, the parts T_M, T_N of T in M, N , respectively can be defined. T_M is an operator in the Banach space M with $D(T_M) = D(T) \cap M$ such that $T_M u = Tu \in M$. T_N is defined similarly. Of course, if T is closed, the same is true for T_M and T_N .

The resolvent set of $T \in \mathcal{C}(X, X)$ is the set of the complex number ζ such that $T - \zeta$ is invertible with $R(\zeta) = (T - \zeta)^{-1}$ a bounded operator. The operator-valued function $R(\zeta)$ thus defined on the resolvent set is called the resolvent of T . Thus $R(\zeta)$ has domain X and range $D(T)$. The complementary set $\Sigma(T)$ (in the complex plane) of the resolvent set is called the spectrum of T .

Sometimes the spectrum of a closed operator T can be splitted in a bounded part Σ' separated from the other part Σ'' , in such a way that a rectifiable simple closed curve Γ can be drawn so as to enclose an open set containing Σ' in its interior and Σ'' in its exterior. Under such circumstance we have the following decomposition result (see Th. 6.17 [17], p. 178). Let $\Sigma(T)$ be separated in two parts in the way described above. Then we have a decomposition of T according to a decomposition $X = M' \oplus M''$ of the space in such a way that the spectra of the parts $T_{M'}, T_{M''}$ coincide with Σ', Σ'' respectively and $T_{M'}$ is a bounded operator in M' .

When Σ' consists of the single point λ , that means $\Sigma(T)$ has an isolated point λ , the operator $T_{M'}$ has spectrum consisting of the single point λ . Now, let assume that λ is a simple eigenvalue of T , then M' is finite-dimensional and we obtain the same Laurent series of $R(\zeta)$ that in (A.1). The eigenprojection is given by the expression (A.2), where Γ is any closed curve enclosing λ but no other point of $\Sigma(T)$. In that case, we can define the isolation distance $d = \text{dist}(\lambda, \Sigma(T) \setminus \{\lambda\})$.

A.3. Analytic families of operators

We have previously studied the behavior of isolated eigenvalues of a finite-rank operator depending on a parameter holomorphically. Nevertheless, it is not sufficient to consider bounded operators only, and it should be noted that the series (A.3) is now meaningless. So a generalized definition of holomorphic families of operators will be given, following [17] (VII, Sect. 1.2), and it will be shown that for an holomorphic family of closed operators with a spectral gap, the results obtained in the finite-dimensional case are always available.

Firstly, let recall that a bounded operator $T(\chi)$ from X to Y is bounded-holomorphic if it is differentiable in norm for all χ in a complex domain. Now, let $T(\chi) \in \mathcal{C}(X, Y)$ a family of operators defined in a neighborhood of $\chi = 0$. $T(\chi)$ is said to be holomorphic at $\chi = 0$ (in a generalized sense) if there is a third Banach space Z and two families of operators $U(\chi)$ from Z in X and $V(\chi)$ from Z in Y which are bounded-holomorphic at $\chi = 0$ such that $U(\chi)$ maps Z onto $D(T(\chi))$ one to one and

$$T(\chi)U(\chi) = V(\chi).$$

$T(\chi)$ is holomorphic in a domain D of the complex plane if it is holomorphic at every χ of D . The new notion is a generalization of the old one, since a bounded families $T(\chi)$ is holomorphic in the new sense if and only if it is bounded-holomorphic.

Let $T(\chi) \in \mathcal{C}(X, X)$ be holomorphic in χ near $\chi = 0$. Then the resolvent $R(\zeta, \chi)$ is bounded-holomorphic in the two variables on the set of all ζ, χ such that ζ is in the resolvent set of $T(0)$ and $|\chi|$ is sufficiently small (depending on ζ) [17] (Th. 1.3, VII, Sect. 1.2). Moreover, let assume that the spectrum $\Sigma(T(0))$ of $T(0)$ can be separated in two parts $\Sigma'(0), \Sigma''(0)$ by a closed curve Γ in the manner previously described. Then for sufficiently small $|\chi|$, the curve Γ is in the resolvent set of the $T(\chi)$ and the spectra $\Sigma(T(\chi))$ is likewise separated by Γ into two parts $\Sigma'(\chi), \Sigma''(\chi)$ with the associated decomposition $X = M'(\chi) \oplus M''(\chi)$ of the space [17] (Th. 1.7, VII, Sect. 1.3). The projection on $M'(\chi)$ along $M''(\chi)$ is given by

$$P(\chi) = -\frac{1}{2\pi i} \int_{\Gamma} R(\zeta, \chi) d\zeta.$$

Since $R(\zeta, \chi)$ is bounded-holomorphic in ζ and χ jointly, it can be proved that $P(\chi)$ is bounded-holomorphic near $\chi = 0$. So, when $\Sigma(T(\chi))$ is separated, the eigenvalue problem for $T(\chi)$ is reduced to the eigenvalue problems for the parts of $T(\chi)$ in the two subspaces $M'(\chi)$ and $M''(\chi)$.

If one is interested only in the case of an isolated eigenvalue of $T(0)$, it is convenient to choose as Γ a circle enclosing the single eigenvalue λ of T ; then $P(\chi)$ coincides with the eigenprojection operator for the eigenvalue $\lambda(\chi)$ and is of finite rank. So the operator $T_r(\chi) = T(\chi)P(\chi)$, whose rank is finite, is bounded-holomorphic and can be expanded in a power series in χ . Thus, the Taylor series expansion for the eigenvalue $\lambda(\chi)$ can be obtained by the method of the previous subsection. For instance, the traces of various operators $\text{tr}(T(\chi)P(\chi))$ are also defined.

We now give a criterion for an important special case of holomorphic family $T(\chi)$ of operators, called holomorphic family of type (A) [17] (VII, Sect. 2.1). A family $T(\chi) \in \mathcal{C}(X, Y)$, defined for χ in a domain D_0 of the complex plane, is said to be holomorphic of type (A) if i) $D(T(\chi)) = D$ is independent of χ and ii) $T(\chi)u$ is holomorphic for $\chi \in D_0$ for every $u \in D$. In this case $T(\chi)u$ has a Taylor expansion at each $\chi \in D_0$. A holomorphic family $T(\chi)$ of type (A) is actually holomorphic in the generalized sense and we have the following criterion [17] (Th. 2.6).

Let T be a closable operator from X to Y with $D(T) = D$. Let $T^{(n)}$, $n = 1, 2, \dots$, be operators from X to Y with domains containing D , and let there be constants $a, b, c \geq 0$ such that

$$\|T^{(n)}u\| \leq c^{n-1} (a\|u\| + b\|Tu\|), \quad u \in D, \quad n = 1, 2, \dots$$

Then the series

$$T(\chi)u = Tu + \chi T^{(1)}u + \chi^2 T^{(2)}u + \dots, \quad u \in D$$

defines an operator $T(\chi)$ with domain D for $|\chi| < 1/c$. If $|\chi| < (b+c)^{-1}$, $T(\chi)$ is closable and the closure $\tilde{T}(\chi)$ for such χ form a holomorphic family of type (A). Note also that we can choose $c = 0$ if $T^{(n)} = 0$ for all $n \geq 2$, and

$$\|T^{(1)}u\| \leq a\|u\| + b\|Tu\|, \quad u \in D.$$

For a holomorphic family $T(\chi)$ with a separated spectrum, we saw that the spectrum of $T(\chi)$ is likewise separated for sufficiently small χ . Here, we can estimate how small $|\chi|$ should be, since the separation of the spectrum occurs at least if

$$|\chi| < r_0 = \min_{\zeta \in \Gamma} (a\|R(\zeta, T)\| + b\|TR(\zeta, T)\| + c)^{-1}.$$

When X is a Hilbert space and $T = T(0)$ is a normal operator, the foregoing result becomes

$$r_0 = \min_{\zeta \in \Gamma} \left[a \sup_{\lambda' \in \Sigma(T)} |\lambda' - \zeta|^{-1} + b \sup_{\lambda' \in \Sigma(T)} |\lambda'| |\lambda' - \zeta|^{-1} + c \right]^{-1}.$$

If we consider an isolated eigenvalue λ of T with isolation distance d , it is convenient to take Γ as the circle $|\zeta - \lambda| = d/2$. Hence, we have

$$r_0 \geq [(2(a + b|\lambda|)d^{-1} + 2b + c)^{-1}].$$

Here, the existence of the formal power series for $T(\chi)$ makes it possible to take over directly the results of the finite-dimensional case without replacing $T(\chi)$ by $T_r(\chi)$, in spite of the unboundedness of the $T^{(n)}$; in fact, these operators appear in the formulas there only in such combinations as $T^{(n)}R(\zeta)$, $T^{(n)}S$, $T^{(n)}P$ which all are bounded operators.

I would like to thank the Referees, the Associate Editor and Stéphane Puechmorel for suggesting a number of improvements.

REFERENCES

- [1] D. Aldous and J. Fill, *Reversible Markov Chains and Random Walks on Graphs*. Monograph in preparation. Available from the Aldous's home page at <http://www.stat.berkeley.edu/users/aldous/book.html>
- [2] B. Bercu and A. Rouault, Sharp large deviations for the Ornstein–Uhlenbeck process (to appear).
- [3] E. Bolthausen, The Berry–Esseen Theorem for Functionals of Discrete Markov Chains. *Z. Wahrscheinlichkeitstheorie Verw.* **54** (1980) 59–73.
- [4] W. Bryc and A. Dembo, Large deviations for quadratic functionals of gaussian processes. *J. Theoret. Probab.* **10** (1997) 307–332.
- [5] M.F. Cheng and F.Y. Wang, Estimation of spectral gap for elliptic operators. *Trans. AMS* **349** (1997) 1239–1267.
- [6] K.L. Chung. *Markov chains with stationnary transition probabilities*. Springer-Verlag (1960).
- [7] J.D. Deuschel and D.W. Stroock, *Large Deviations*. Academic Press, Boston (1989).
- [8] P. Diaconis, S. Holmes and R.M. Neal, *Analysis of a non-reversible markov chain sampler*, Technical Report. Cornell University, BU-1385-M, Biometrics Unit (1997).
- [9] I.H. Dinwoodie, A probability inequality for the occupation measure of a reversible Markov chain. *Ann. Appl. Probab* **5** (1995) 37–43.
- [10] I.H. Dinwoodie, Expectations for nonreversible Markov chains. *J. Math. Ann. App.* **220** (1998) 585–596.
- [11] I.H. Dinwoodie and P. Ney, Occupation measures for Markov chains. *J. Theoret. Probab.* **8** (1995) 679–691.
- [12] W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol. 2. Wiley & Sons, 2nd Edition (1971).
- [13] S. Gallot and D. Hulin and J. Lafontaine, *Riemannian Geometry*. Springer-Verlag (1990).
- [14] D. Gillman, *Hidden Markov Chains: Rates of Convergence and the Complexity of Inference*, Ph.D. Thesis. Massachusetts Institute of Technology (1993).
- [15] L. Gross, Logarithmic Sobolev Inequalities and Contractivity Properties of Semigroups, in *Dirichlet forms*, Varenna (Italy). Springer-Verlag, *Lecture Notes in Math.* **1563** (1992) 54–88.
- [16] J.L. Jensen, *Saddlepoint Approximations*. *Oxford Statist. Sci. Ser.* **16**.
- [17] T. Kato, *Perturbation theory for linear operators*. Springer (1966).
- [18] D. Landers and L. Rogge, On the rate of convergence in the central limit theorem for Markov chains. *Z. Wahrscheinlichkeitstheorie Verw.* **35** (1976) 169–183.
- [19] G.F. Lawler and A.D. Sokal, Bounds on the L^2 spectrum for Markov chains and Markov processes: A generalization of Cheeger's inequality. *Trans. Amer. Math. Soc.* **309** (1988) 557–580.
- [20] P. Lezaud, Chernoff-type Bound for Finite Markov Chains. *Ann. Appl. Probab* **8** (1998) 849–867.
- [21] B. Mann, *Berry–Esseen Central Limit Theorem for Markov chains*, Ph.D. Thesis. Harvard University (1996).
- [22] K. Marton, A measure concentration inequality for contracting Markov chains. *Geom. Funct. Anal.* **6** (1996) 556–571.
- [23] S.V. Nagaev, Some limit theorems for stationary Markov chains. *Theory Probab. Appl.* **2** (1957) 378–406.
- [24] P.M. Samson, Concentration of measure inequalities for Markov chains and ϕ -mixing processes, *Ann. Probab.* **28** (2000) 416–461.
- [25] H.F. Trotter, On the product of semi-groups of operators. *Proc. Amer. Math. Soc.* **10** (1959) 545–551.
- [26] F.Y. Wang, Existence of spectral gap for elliptic operators. *Math. Sci. Res. Inst.* (1998).