Trous spectraux pour certains algorithmes de Métropolis sur $\mathbb{R}$

Laurent Miclo et Cyril Roberto

Laboratoire de Statistique et Probabilités, UMR C5583
Université Paul Sabatier et CNRS
118, route de Narbonne
31 062 Toulouse cedex
email : miclo@cict.fr, roberto@cict.fr

RÉSUMÉ :

Sur $\mathbb{R}$ considérons le noyau markovien $Q(x,dy) = \mathbb{1}_{]x-\frac{1}{2},x+\frac{1}{2}[}(y)\,dy$ réversible par rapport à la mesure de Lebesgue dy. A partir de celui-ci on effectue une transformation de type Métropolis pour obtenir un noyau $P(x,dy)$ réversible par rapport à la loi de densité proportionnelle à $e^{-U(x)}$ relativement à la mesure de Lebesgue dx où U est une fonction convexe régulière vérifiant $\lim_{x\to\pm\infty} U(x) = +\infty$ et une autre hypothèse de croissance.

Le but de cette note est de prouver l'existence d'un trou spectral pour l'opérateur positif $\mathrm{Id} - P$ (où Id désigne l'identité), la difficulté résidant surtout dans la non-compacité du modèle.

Nous introduirons tout d'abord plus en détail les objets utilisés avant de réduire le problème et de le résoudre, nous utiliserons principalement une technique de multiflots.

1 Introduction

L'existence d'un trou spectral permet d'avoir des évaluations quantitatives de la vitesse de convergence vers l'équilibre d'algorithmes stochastiques comme par exemple ceux de Métropolis. Ce sont ces estimations qui motivent notre étude.

Plus précisément, on se donne sur $\mathbb{R}$ le noyau markovien $Q(x,dy) = \mathbb{1}_{]x-\frac{1}{2},x+\frac{1}{2}[}(y)\,dy$ réversible par rapport à la mesure de Lebesgue dy et U une fonction convexe vérifiant $\lim_{x\to\pm\infty} U(x) = +\infty$.

Sous cette hypothèse et par convexité, on montre facilement qu'il existe deux réels a et α strictement positifs tels que pour tout $|x| \geq a$,

$$U(x) \geq \alpha|x| \quad \text{et} \quad |U'(x)| \geq \alpha\,. \tag{1}$$

On définit la mesure

$$\mu(dx) = K^{-1}e^{-U(x)}dx \quad \text{avec} \quad K = \int_{\mathbb{R}} e^{-U(x)}dx$$

$K < \infty$ par définition de a, et on construit par un procédé de type Métropolis le noyau

$$P(x,dy) = e^{-(U(y)-U(x))_+}\,Q(x,dy) + a_x\delta_x(dy) = f_x(y)\,dy + a_x\delta_x(dy)$$

où

$$x_+ \overset{\text{déf.}}{=} \max(x,0), \quad a_x = 1 - \int_{\mathbb{R}} f_x(y)\,dy \quad \text{et} \quad f_x(y) = \begin{cases} e^{-(U(y)-U(x))_+} & \text{si } |y-x| \leq \frac{1}{2} \\ 0 \ \text{sinon} \end{cases} .$$

Par construction, P est réversible par rapport à μ et est irréductible. On pourra donc se servir du semi-groupe markovien $(e^{t(P-\mathrm{Id})})_{t\geq 0}$ pour approximer μ.

On suppose alors de plus sur U qu'il existe une constante $C > 1$ telle que :

$$(H) \qquad \begin{cases} \forall x \in [a,\infty[& U'(x+1) \leq CU'(x) \\ \forall x \in]-\infty,-a] & |U'(x-1)| \leq C|U'(x)| \end{cases}$$

On remarque que l'hypothèse (H) n'est pas très restrictive car tout polynôme convient. En particulier, la loi gaussienne rentre dans le cadre de notre étude. En revanche, des fonctions "trop rapides" comme e^{x^2} ne conviennent pas. On peut également noter que cette hypothèse est technique et que l'intuition nous invite à penser qu'elle n'est pas nécessaire.

Le but de cette note est de montrer que sous ces hypothèses, le trou spectral de $\mathrm{Id} - P$ est strictement positif :

$$\lambda \overset{\text{déf.}}{=} \inf_{f\in\mathbb{L}^2(\mu)\setminus\mathrm{Vect}(\mathbf{1})} \frac{\mu(f(\mathrm{Id}-P)f)}{\mu((f-\mu(f))^2} > 0$$

Par ce biais, nous pourrons évaluer la vitesse de convergence de $e^{t(P-\mathrm{Id})}$ pour t grand vers μ car nous savons que cette convergence est exponentiellement rapide dans $\mathbb{L}^2(\mu)$, de taux λ.

Nous montrerons tout d'abord dans les deux sections suivantes qu'il existe un trou spectral sur $]-\infty,-a]$, $[-a-1,a+1]$ et $[a,\infty[$ pour le noyau P et la mesure μ restreints à ces différents intervalles, où a est défini plus haut. Par symétrie autour de 0, il nous suffira de traiter les cas $[-a-1,a+1]$ et $[a,\infty[$. Nous commencerons par nous intéresser à l'intervalle $[a,\infty[$ qui représente la plus grosse difficulté, on y développera une technique de multiflots due à Sinclair (voir [12]), nous nous placerons ensuite sur $[-a-1,a+1]$ où l'existence d'un trou spectral strictement positif est relativement aisée à obtenir. Dans la quatrième section, nous recollerons enfin les morceaux pour obtenir le résultat général sur $\mathbb{R}$ tout entier.

Puis dans une dernière section, nous discuterons des propriétés similaires dans le cas où l'on s'intéresse plutôt aux chaînes de Markov en temps discret admettant P pour noyaux de probabilités de transition.

Comme nous l'a pertinemment rappelé le rapporteur, il existe en temps discret d'autres manières d'obtenir une convergence exponentiellement rapide vers la probabilité invariante. Par exemple si P était quasi-compact (cf. [9]), on saurait qu'uniformément en $x \in \mathbb{R}$, la variation totale $\|P^n(x,\cdot\,)-\mu\|_{\mathrm{vt}}$ converge vers 0 pour n grand, mais ceci n'est pas possible ici (considérer des points initiaux x très grands). Une approche alternative consiste à trouver une fonction mesurable de Lyapounov, $V \geq 1$, qui satisfait en dehors d'un compact $P(V) \leq \alpha V$, avec $0 \leq \alpha < 1$, car ceci assure tout d'abord que $V \in \mathbb{L}^1(\mu)$ puis qu'il existe deux constantes $K > 0$ et $0 < \widetilde{\rho} < 1$ telles que

$$(2) \qquad \forall\, x_0 \in \mathbb{R}, \qquad \sup_{\{f \in \mathbb{L}^1(\mu)\,:\,|f| \leq V\}} |P^n(f)(x_0) - \mu(f)| \;\leq\; K \widetilde{\rho}^n V(x_0)$$

(voir le chapitre 16 du livre de Meyn et Tweedie [6], en notant que dans notre situation les ensembles compacts sont des « petite sets »).

De telles applications V existent pour les cas traités dans cet article (car on peut toujours prendre $V(\,\cdot\,) = \exp(\alpha\,|\cdot|\,/2)$, avec α la constante intervenant dans (1), ceci fournissant d'ailleurs des exemples où $V \notin \mathbb{L}^2(\mu)$, en considérant $U(\,\cdot\,) = \alpha\,|\cdot|$), toutefois l'existence d'un trou spectral permet d'obtenir des résultats de convergence uniforme en un sens différent. Par exemple on peut remplacer dans (2) l'ensemble sur lequel est considéré le maximum par

$$\{f \in \mathbb{L}^2(\mu) \,:\, f(x_0) \leq K_1,\, \mu(f^2) \leq K_2\}$$

où $K_1, K_2 \geq 0$ sont deux constantes (pour plus de précisions, voir la fin de l'article). Cependant l'importance de cette uniformité (et le fait que le trou spectral est relativement stable par perturbations bornées de U sur tout $\mathbb{R}$) apparaît plutôt pour des chaînes de Markov inhomogènes en temps. Mais ce gain se paye, car l'étude du trou spectral est plus délicate que l'obtention d'une bonne fonction de Lyapounov V.

D'autre part la dimension 1 est certainement trop restrictive, cependant notons que le résultat présenté s'étend immédiatement par comparaison et tensorisation sur $\mathbb{R}^n$, $n \geq 2$, pour des potentiels de la forme

$$\forall\, x = (x_1, \cdots, x_n) \in \mathbb{R}^n, \qquad U(x) \;=\; U_1(x_1) + \cdots + U_n(x_n)$$

où les U_i, $1 \leq i \leq n$ satisfont les hypothèses considérées ici, en utilisant les inégalités $(U(y) - U(x))_+ \leq \sum_{1 \leq i \leq n} (U_i(y_i) - U_i(x_i))_+$, valables pour tous $x = (x_1, \cdots, x_n)$, $y = (y_1, \cdots, y_n) \in \mathbb{R}^n$ (on peut d'ailleurs s'interroger sur un résultat plus général pour des fonctions $U : \mathbb{R}^n \to \mathbb{R}$ satisfaisant une bonne condition de stricte convexité en dehors d'un compact, à l'instar des travaux connus pour les diffusions, voir [2] ou [1]). Mais nous pensons que le cas réel donne une bonne illustration de l'utilité des multiflots pour palier la non-bornitude des modèles.

2 Existence du trou spectral sur $[a, \infty[$

On restreint le noyau P et la mesure μ à $[a, \infty[$ pour obtenir le noyau

$$P_1(x, dy) = f_x(y)\, \mathbb{1}_{[a,\infty[}(y)\, dy + a_{1,x} \delta_x(dy)$$

avec

$$a_{1,x} = 1 - \int_a^\infty f_x(y)\, dy\ ,$$

et la mesure

$$\mu_1(dx) = K_1^{-1} e^{-U(x)}\, dx \quad \text{où} \quad K_1 = \int_a^\infty e^{-U(x)}\, dx.$$

On a toujours P_1 réversible par rapport à la mesure μ_1, on veut montrer :

$$\lambda_1 \stackrel{\text{déf.}}{=} \inf_{f \in \mathbb{L}^2(\mu_1) \setminus \mathrm{Vect}(\mathbf{I})} \frac{\mu_1(f(\mathrm{Id} - P_1)f)}{\mu_1((f - \mu_1(f))^2} > 0 \ .$$

Il n'existe que peu de résultats donnant des estimations satisfaisantes du trou spectral d'un noyau évoluant sur un ensemble continu, sauf si celui-ci provient d'un semi-groupe de diffusion considéré par exemple au temps 1. Ainsi pour cette étude, on a commencé par chercher à faire des comparaisons avec le processus d'Ornstein-Uhlenbeck, mais sans véritablement y arriver. Rosenthal (voir [10, exemple 5.4 pages 69-71]) s'intéresse par exemple au cas $U(x) = x^2$ (cas gaussien), par ses méthodes, il obtient une borne du trou spectral sur $[-M, M]$ qui explose lorsque M tend vers l'infini. La difficulté provient essentiellement du caractère non compact de $\mathbb{R}$ (ou de l'intervalle $[a, \infty[$).

Aussi, nous allons chercher, par discrétisation, à nous ramener à un noyau sur $\mathbb{N}$ où les résultats sont plus nombreux. Le noyau sur $[a, \infty[$ (et donc le trou spectral correspondant sur $[a, \infty[$) sera vu comme une limite en un certain sens de noyaux sur $\mathbb{N}$. L'idée directrice du choix de la discrétisation est principalement le souci de simplifier les transitions tout en gardant des comportements proches du noyau initial. Proche étant pris au sens où les trous spectraux sont comparables.

2.1 Réduction du problème

On effectue une discrétisation non homogène de $[a, \infty[$:

Par convexité et par construction de a, il existe une suite croissante $(x_k)_{k \in \mathbb{N}}$ de $[a, \infty[$, définie par

$$x_0 = a \quad \text{et} \quad U(x_{k+1}) = U(x_k) + \delta \quad \text{où} \quad \delta \in]0, 1[\quad \text{est fixé.}$$

Par convexité à nouveau, $\delta_k \stackrel{\text{déf.}}{=} x_{k+1} - x_k$ est une suite décroissante et $x_k \stackrel{k \to \infty}{\longrightarrow} \infty$.

On introduit alors

$$\begin{aligned}
&\text{la tribu} \quad \mathcal{A}_\delta = \sigma(I_k \stackrel{\text{déf.}}{=} [x_k, x_{k+1}[, k \in \mathbb{N}) \ , \\
&\text{la mesure} \quad \tilde{\mu}_\delta(i) = \mu_1(I_i) \ , \ \text{pour } i \in \mathbb{N} \ , \\
&\text{et la probabilité de transition} \quad \tilde{P}_\delta(i, j) = \frac{1}{\tilde{\mu}_\delta(i)} \int_{I_i \times I_j} P_1(x, dy) \mu_1(dx) \ . \qquad (3)
\end{aligned}$$

On se ramène ainsi à $\mathbb{N}$, car en posant pour $\delta > 0$,

$$\tilde{\lambda}_\delta \stackrel{\text{déf.}}{=} \inf_{f \in \mathbb{L}^2(\tilde{\mu}_\delta) \setminus \mathrm{Vect}(\mathbf{I})} \frac{\tilde{\mu}_\delta(f(\mathrm{Id} - \tilde{P}_\delta)f)}{\tilde{\mu}_\delta((f - \tilde{\mu}_\delta(f))^2)} \qquad (4)$$

il est bien connu que $\lim_{n \to \infty} \tilde{\lambda}_{2^{-n}} = \lambda_1$ (ceci découle du fait que les tribus $\mathcal{A}_{2^{-n}}$ convergent en croissant pour $n \in \mathbb{N}$ grand, vers la tribu borélienne de $[a, +\infty[$, pour plus de détails voir par exemple [7]).

En vue d'estimer $\tilde{\lambda}_\delta$, cherchons à maîtriser $\tilde{\mu}_\delta$ et $\tilde{P}_\delta$:
Par construction de $(x_k)_{k \in \mathbb{N}}$ et par définition de δ_k,

$$K_1^{-1}(x_{k+1} - x_k) e^{-U(x_{k+1})} \ \leq \ \tilde{\mu}_\delta(k) = K_1^{-1} \int_{x_k}^{x_{k+1}} e^{-U(x)} dx \ \leq \ K_1^{-1}(x_{k+1} - x_k) e^{-U(x_k)}$$

d'où

$$(5) \qquad K_1^{-1}e^{-1}\delta_k e^{-U(x_k)} \;\leq\; \tilde{\mu}_\delta(k) \;\leq\; K_1^{-1}\delta_k e^{-U(x_k)}$$

On pose alors

$$\boxed{\hat{\mu}_\delta(k) = Z_\delta^{-1}\delta_k e^{-U(x_k)}}$$

où Z_δ est la constante de renormalisation (qui a priori dépend très peu de $\delta > 0$, car elle vérifie $K_1 \leq Z \leq K_1 e$), et on note que $\hat{\mu}_\delta(k)\#\tilde{\mu}_\delta(k)$, où $\#$ signifie un encadrement à des constantes indépendantes du niveau d'approximation δ près. Par un calcul analogue, on montre que $\tilde{P}_\delta(i,j)\#\delta_j e^{-(U(x_j)-U(x_i))_+}$, pour i et j vérifiant $|i-j| \leq \lfloor \frac{1}{2\delta_{i\wedge j}} - 1 \rfloor$, on pose donc

$$\boxed{\hat{P}_\delta(i,j) = \begin{cases} \delta_j e^{-(U(x_j)-U(x_i))_+}, & \text{si } i \text{ et } j \text{ vérifient } 0 < |i-j| \leq \lfloor \frac{1}{2\delta_{i\wedge j}} - 1 \rfloor \\ 1 - \sum_{k\neq i} \hat{P}_\delta(i,k), & \text{si } i = j \\ 0 & \text{sinon} \end{cases}}$$

où $\lfloor \; \rfloor$ désigne la partie entière.

Par décroissance de δ_k, la condition $|i-j| \leq \lfloor \frac{1}{2\delta_{i\wedge j}} - 1 \rfloor$ assure que $x_{j+1} - x_i \leq \frac{1}{2}$ et $x_{i+1} - x_j \leq \frac{1}{2}$. Ce qui permet de réaliser plus précisément que pour tous $i \neq j \in \mathbb{N}$, on a $\tilde{P}_\delta(i,j) \geq e^{-2}\hat{P}_\delta(i,j)$. Tenant compte de (5) et de résultats généraux de comparaison (voir par exemple [11]), on obtient que

$$\tilde{\lambda}_\delta \geq e^{-3}\hat{\lambda}_\delta$$

où

$$\hat{\lambda}_\delta \overset{\text{déf.}}{=} \inf_{f\in\mathbb{L}^2(\hat{\mu}_\delta)\setminus\mathrm{Vect}(\mathbb{1})} \frac{\hat{\mu}_\delta(f(\mathrm{Id}-\hat{P}_\delta)f)}{\hat{\mu}_\delta(f-\hat{\mu}_\delta(f))^2}$$

Ce dernier résultat et (4) montrent que le trou spectral λ_1 sera strictement positif si nous arrivons à minorer $\hat{\lambda}_\delta$ indépendamment de δ (c'est l'objet de la section 2.2).

Afin de simplifier les notations, introduisons les fonctions $j_{\min}$ et $j_{\max}$ de $\mathbb{N}$ dans $\mathbb{N}$ définies par :

$$\text{pour tout } i, \quad \hat{P}_\delta(i,j) \neq 0 \text{ si et seulement si } j \in \{j_{\min}(i), \ldots, j_{\max}(i)\}$$

On vérifie facilement que $i - j_{\min}(i)$ est croissante et $j_{\min}(j_{\max}(i)) = j_{\max}(j_{\min}(i)) = i$. On remarque enfin qu'on connaît explicitement $j_{\max}(i)$ car $j_{\max}(i) - i = \lfloor \frac{1}{2\delta_i} - 1 \rfloor$ et que $j_{\min}(i)$ n'est défini qu'implicitement.

Au cours de la démonstration, nous aurons besoin des résultats suivants relatifs à ces fonctions :

Lemme 1 *Si C désigne la constante définie dans l'hypothèse (H), pour δ assez petit et pour tout i,*

$$j_{\max}(i) - i - 1 \leq 2C(i - j_{\min}(i)).$$

Lemme 2 *Pour δ assez petit et pour tout i,*

$$\frac{1}{i - j_{\min}(i)} \leq 4C\delta_i.$$

Demonstration du lemme 1 :
Soit $j = j_{\min}(i)$, d'après le théorème des accroissements finis, il existe $x_i^{'} \in [x_i, x_{i+1}]$ et $x_j^{'} \in [x_j, x_{j+1}]$ tels que

$$\delta_i = \frac{\delta}{U'(x_i^{'})} \quad \text{et} \quad \delta_j = \frac{\delta}{U'(x_j^{'})}.$$

D'après l'hypothèse (H) on a

$$\begin{array}{rcl} U'(x_i^{'}) & \leq & CU'(x_j^{'}) \\ \Rightarrow \quad \dfrac{1}{2\delta_i} & \leq & C\dfrac{1}{2\delta_j} \; . \end{array}$$

Par ailleurs $U' \geq \alpha$ sur $[a, \infty[$ d'après (1), d'où

$$\frac{1}{2\delta_j} \geq \frac{\alpha}{2\delta} > 4 \quad \text{pour } \delta \text{ petit.}$$

Ainsi

$$\begin{array}{rcl} \dfrac{1}{2\delta_i} & \leq & 2C(\dfrac{1}{2\delta_j} - 2) \\ \Rightarrow \quad \dfrac{1}{2\delta_i} - 1 - 1 & \leq & 2C(\dfrac{1}{2\delta_j} - 1 - 1) \\ \Rightarrow \quad \lfloor \dfrac{1}{2\delta_i} - 1 \rfloor - 1 & \leq & 2C \lfloor \dfrac{1}{2\delta_j} - 1 \rfloor \end{array}$$

ce qui donne le résultat puisque par définition

$$\lfloor \frac{1}{2\delta_i} - 1 \rfloor = j_{\max}(i) - i \quad \text{et} \quad \lfloor \frac{1}{2\delta_j} - 1 \rfloor = i - j_{\min}(i) \; .$$

Démonstration du lemme 2 :
On pose ici aussi $j = j_{\min}(i)$.
D'après la démonstration précédente on a le résultat intermédiaire

$$\frac{1}{2\delta_i} \leq 2C(\frac{1}{2\delta_j} - 2)$$

donc

$$\frac{1}{2\delta_i} \leq 2C \lfloor \frac{1}{2\delta_j} - 1 \rfloor$$

il vient

$$\begin{array}{rcl} \dfrac{1}{i - j_{\min}(i)} = \lfloor \dfrac{1}{2\delta_j} - 1 \rfloor^{-1} & \leq & 2C \dfrac{1}{\frac{1}{2\delta_i}} \\ & \leq & 4C\delta_i \; . \end{array}$$

2.2 Minoration de $\hat{\lambda}_\delta$

Pour montrer que $\hat{\lambda}_\delta$ est minorée par une constante indépendante de δ, on va utiliser un résultat de Sinclair (voir [12, section 4]).

Sur $\mathbb{N}$ on construit un graphe où les sommets sont les points et $e = (u, v)$ est une arête si et seulement si $\hat{\mu}_\delta(u)\hat{P}_\delta(u, v) > 0$. Pour tout x différent de y, on note Γ_{xy} l'ensemble des chemins joignant x à y. Sur Γ_{xy}, on choisit une probabilité ϕ_{xy}, $\phi_{xy}(\gamma_{xy})$ représente la probabilité de choisir le chemin γ_{xy} parmi l'ensemble des chemins joignant x à y. On notera souvent γ au lieu de γ_{xy} pour ne pas alourdir les notations.

Sous cette construction, la borne de Sinclair est

$$\eta_\delta(\phi) = \sup_{e=(u,v)} \eta_\delta(\phi)(e) = \sup_{e=(u,v)} \frac{1}{\hat{\mu}_\delta(u)\hat{P}_\delta(u, v)} \sum_{x,y\,:\,\gamma_{xy}\ni e} \hat{\mu}_\delta(x)\hat{\mu}_\delta(y)\phi_{xy}(\gamma)\,.$$

Le théorème s'énonce alors :

$$\hat{\lambda}_\delta \geq \frac{1}{8\eta_\delta^2(\phi)}$$

La clef de la démonstration repose sur le choix de ϕ :

Soit $x < y$ (si $x > y$ on prendra le chemin construit à partir de celui allant de y à x à l'envers, ce qui est permis par réversibilité). Partant de y, on choisit y_1 tel que $y - j_{\min}(y) \leq y_1 < y$ avec probabilité $p_y \stackrel{\text{déf.}}{=} \frac{1}{y - j_{\min}(y)}$ puis on choisit y_2 de manière équiprobable, $y_1 - j_{\min}(y_1) \leq y_2 < y_1$, avec probabilité p_{y_1} et ainsi de suite jusqu'à dépasser x, $y_{n+1} \leq x < y_n$, on pose alors $y_{n+1} = x$. Le chemin ainsi construit est formé des arêtes $(x, y_n), (y_n, y_{n-1}), \ldots, (y_1, y)$.

On va majorer explicitement $\eta_\delta(\phi)$. Pour cela, on peut par construction se restreindre aux arêtes $e = (u, v)$ avec $u < v$. Soit $e = (u, v)$ une telle arête fixée $(u < v)$, on doit calculer en premier lieu

$$\frac{1}{\hat{\mu}_\delta(u)\hat{P}_\delta(u, v)} = \frac{1}{\delta_u \delta_v} e^{U(x_v)} \tag{6}$$

et en deuxième lieu, la somme

$$\sum_{x,y\,:\,\gamma_{xy}\ni e} \hat{\mu}_\delta(x)\hat{\mu}_\delta(y)\phi_{xy}(\gamma) = \sum_{x\leq u} \hat{\mu}_\delta(x) \sum_{y\geq v} \hat{\mu}_\delta(y) \sum_{\gamma_{xy}\ni e} \phi_{xy}(\gamma)$$

Le dernier terme $\sum_{\gamma_{xy}\ni e} \phi_{xy}(\gamma)$ est à comprendre comme la somme, à x et y fixés, sur tous les chemins $\gamma \in \Gamma_{xy}$ tels que $\gamma \ni e$.

D'après le choix de ϕ, on doit distinguer les cas $x < u$ et $x = u$:
On pose

$$\begin{aligned}
\sum_{x,y\,:\,\gamma_{xy}\ni e} \hat{\mu}_\delta(x)\hat{\mu}_\delta(y)\phi_{xy}(\gamma) = \quad & \sum_{x=u} \hat{\mu}_\delta(x) \sum_{y\geq v} \hat{\mu}_\delta(y) \sum_{\gamma_{xy}\ni e} \phi_{xy}(\gamma) \\
& + \sum_{x<u} \hat{\mu}_\delta(x) \sum_{y\geq v} \hat{\mu}_\delta(y) \sum_{\gamma_{xy}\ni e} \phi_{xy}(\gamma) \\
= \quad & Q_1 + Q_2 + Q_3 + Q_4
\end{aligned}$$

avec

$$\begin{aligned}
Q_1 &= \hat{\mu}_\delta(u)\hat{\mu}_\delta(v)\sum_{\gamma_{uv}\ni e}\phi_{uv}(\gamma)\\
Q_2 &= \hat{\mu}_\delta(u)\sum_{y>v}\hat{\mu}_\delta(y)\sum_{\gamma_{uy}\ni e}\phi_{uy}(\gamma)\\
Q_3 &= \hat{\mu}_\delta(v)\sum_{x<u}\hat{\mu}_\delta(x)\sum_{\gamma_{xv}\ni e}\phi_{xv}(\gamma)\\
Q_4 &= \sum_{x<u}\hat{\mu}_\delta(x)\sum_{y>v}\hat{\mu}_\delta(y)\sum_{\gamma_{xy}\ni e}\phi_{xy}(\gamma)
\end{aligned}$$

Considérons chacun des cas séparément :

Pour Q_4 :

On doit s'intéresser à $\sum_{\gamma_{xy}\ni e}\phi_{xy}(\gamma)$ qui représente, à $x < u$ et $y > v$ fixés, la probabilité globale (sous $\phi_{x,y}$) d'emprunter l'arête $e = (u, v)$, elle aussi préalablement fixée, en allant de y à x. Soit m le premier point par lequel le chemin passe entre v et $j_{\max}(v)$ (à partir de ce point, le chemin peut directement sauter en v). A partir de m, le chemin va passer par n points $x_1, x_2, \ldots, x_n$ puis par v et par u et enfin continuer vers x. Pour se rendre de m à x_1 puis de x_1 à $x_2, \ldots$, on a une probabilité $p_m, p_{x_1}, \ldots$ A chaque étape les probabilités vérifient par croissance de la fonction $\mathrm{id} - j_{\min}$, $p_m \le p_{x_1} \le \cdots \le p_v$.

Ainsi on a, à x et y fixés, en sommant sur les chemins γ joignant x à y

$$\sum_{\gamma_{xy}\ni e}\phi_{xy}(\gamma) \le \sum_{n=0}^{m-v-1} C^n_{m-v-1}(p_v)^{n+1}p_v\ , \tag{7}$$

où C^n_{m-v-1} est le nombre de choix de n points entre m et v, p_v^{n+1} est une majoration de la probabilité de passer par ces n points puis par v et où le dernier p_v représente la probabilité de choisir u une fois en v.

Or, d'après le lemme 1, $j_{\max}(v) - v - 1 \le \frac{2C}{p_v}$, donc

$$\sum_{n=0}^{m-v-1} C^n_{m-v-1}(p_v)^{n+1}p_v = (1+p_v)^{m-v-1}(p_v)^2 \le (1+p_v)^{j_{\max}(v)-v-1}(p_v)^2 \le e^{2C}(p_v)^2 \tag{8}$$

Pour finir d'estimer Q_4, il faut calculer

$$\sum_{y>v}\hat{\mu}_\delta(y) = \sum_{y>v}\delta_y e^{-U(x_y)} \le e\int_{x_{v+1}}^{\infty} e^{-U(x)}dx\ .$$

(l'inégalité provient de (5)).

Il vient

$$\begin{aligned}
\sum_{y>v}\hat{\mu}_\delta(y) &\le e\, e^{-U(x_{v+1})}\int_0^\infty e^{-(U(x+x_{v+1})-U(x_v))}dx\\
&\le e\,\frac{e^{-U(x_v)}}{\alpha} \qquad (9)
\end{aligned}$$

car par convexité et par l'hypothèse (1),

$$U(x + x_{v+1}) - U(x_{v+1}) \geq \frac{1}{\alpha}(x + x_{v+1} - x_{v+1}) \quad U' \geq \frac{1}{\alpha} \text{ et } U(x_{v+1}) \geq U(x_v).$$

En regroupant les calculs de (7), (8) et (9) et en majorant grossièrement $\sum_{x<u} \hat{\mu}_\delta(x)$ par $\sum_{x\in\mathbb{N}} \hat{\mu}_\delta(x) = 1$, on obtient

$$Q_4 \leq e^{-U(x_v)} \frac{e}{\alpha} e^{2C} {p_v}^2 \tag{10}$$

Pour Q_3 :

On majore le terme $\sum_{\gamma_{xv}\ni e} \phi_{xv}(\gamma)$ par p_v car dans Q_3, on a $x < u$, le chemin doit directement sauter de v en u (avec probabilité p_v) puis se rendre en x. De même que précédemment on majore grossièrement $\sum_{x<u} \hat{\mu}_\delta(x)$ par 1, il vient

$$Q_3 \leq \delta_v e^{-U(x_v)} p_v \tag{11}$$

Pour Q_2 :

On doit s'intéresser à $\sum_{\gamma_{uy}\ni e} \phi_{uy}(\gamma)$. Ici, $x = u$, le chemin se rend de y à v puis saute directement de v à $u = x$. En reprenant les calculs du cas Q_4, les majorations (7) et (8) restent valables sinon le dernier p_v qui ici n'apparaît plus :

$$\sum_{\gamma_{uy}\ni e} \phi_{uy}(\gamma) \leq \sum_{n=0}^{m-v-1} C^n_{m-v-1} (p_v)^{n+1} \leq e^{2C} p_v \ .$$

Ce dernier calcul et (9) nous donne

$$Q_2 \leq \delta_u e^{-U(x_u)} \frac{e}{\alpha} e^{-U(x_v)} e^{2C} p_v \tag{12}$$

Pour Q_1 :

On a

$$\sum_{\gamma_{uv}\ni e} \phi_{uv}(\gamma) = \phi_{uv}(\gamma) \leq 1$$

où γ est ici l'unique chemin joignant u et v en passant directement par l'arête $e = (u, v)$.
Ainsi

$$Q_1 \leq \hat{\mu}_\delta(u)\hat{\mu}_\delta(v) = \delta_u \delta_v e^{-U(x_u)} e^{-U(x_v)} \tag{13}$$

Avant de regrouper les calculs, remarquons que d'après le lemme 2,

$$p_v = \frac{1}{v - j_{\min}(v)} \leq 4C\delta_v$$

Si on pose $K' = \sup_{x\in\mathbb{R}}(e^{-U(x)})$ ($K' < \infty$ par hypothèses sur U), d'après (10), (11), (12), (13) et (6), on a

$$\eta_\delta(\phi)(e) \leq K' + \frac{4CK'e}{\alpha} e^{2C} + 4C + \frac{e}{\alpha} e^{2C} (4C)^2 \ . \tag{14}$$

où on a utilisé $\frac{\delta_v}{\delta_u} \leq 1$ par croissance de δ_k.

Ainsi, $\eta_\delta(\phi)(e)$ est bornée par une constante indépendante de u, v donc de $e = (u, v)$ et de δ, en prenant le sup sur les arêtes et en passant à la limite en δ, on a l'existence du trou spectral sur $[a, \infty[$,

$$\lambda_1 > 0.$$

On remarque que la constante trouvée dans (14) tend vers l'infini lorsque C tend vers l'infini, ainsi nos résultats ne sont pas généralisables par cette démonstration à des densités par exemple de la forme $e^{-e^{x^2}}$ où $U(x) = e^{x^2}$ ne vérifie pas l'hypothèse (H). Cependant, l'intuition nous laisse penser que plus U croit rapidement vers l'infini en l'infini et plus le trou spectral a des chances d'être grand, aussi, l'hypothèse (H) n'est-elle pas très naturelle.

3 Existence du trou spectral sur $[-a-1, a+1]$

Comme précédemment, on restreint le noyau P et la mesure μ à $[-a-1, a+1]$, pour obtenir le noyau

$$P_2(x, dy) = f_x(y)\ \mathbb{1}_{[-a-1,a+1]}(y)\ dy + a_{2,x}\delta_x(dy)$$

où

$$a_{2,x} = 1 - \int_{-a-1}^{a+1} f_x(y)\ dy$$

et la mesure

$$\mu_2(dx) = K_2^{-1} e^{-U(x)}\ dx \quad \text{avec} \quad K_2 = \int_{-a-1}^{a+1} e^{-U(x)}\ dx\ .$$

Pour montrer que le noyau P_2, réversible par rapport à la mesure μ_2 sur $[-a-1, a+1]$, admet un trou spectral non nul, on va utiliser des arguments très simples de comparaison et d'analyse de Fourier. On aurait également pu reprendre une méthode due à Rosenthal ([10, théorème 5]), qui généralise sur un ensemble continu la borne de Sinclair exposée dans la section précédente.

Soit $f \in \mathbb{L}^2(\mu)$, on cherche donc à minorer

$$\begin{aligned}
&\frac{\int_{[-1-a,1+a]^2} \mu_2(dx) P_2(x, dy)(f(y) - f(x))^2}{\int_{[-1-a,1+a]} \mu_2(dx)(f(x) - \mu_2(f))^2} \\
&\geq \frac{\int_{[-1-a,1+a]^2} \mu_2(dx) P_2(x, dy)(f(y) - f(x))^2}{\int_{[-1-a,1+a]} \mu_2(dx)(f(x) - \int_{[-1-a,1+a]} f(y)\, dy/(2(1+a)))^2} \\
&\geq \exp(-A) \frac{\int_{[-1-a,1+a]^2} \mathbb{1}_{[x-1/2,x+1/2]}(y)(f(y) - f(y))^2\, dxdy}{\int_{[-1-a,1+a]} \Big(f(x) - \int_{[-1-a,1+a]} f(y)\, dy/(2(1+a))\Big)^2 dx}
\end{aligned}$$

avec $A = \max_{[-1-a,1+a]} U - \min_{[-1-a,1+a]} U$.

Soit $\widetilde{f}$ la fonction définie sur $T = \mathbb{R}/\mathbb{Z}$ (aussi identifié avec $[0, 1[$) par

$$\forall\ 0 \leq t < 1, \qquad \widetilde{f}(t) \ = \ \begin{cases} f(-1 - a + 4t(a+1)) & , \text{si } 0 \leq t \leq 1/2 \\ \widetilde{f}(1-t) & , \text{si } 1/2 < t < 1 \end{cases}$$

En considérant séparément les interactions entre des couples de points de T de part et d'autre de $1/2$ et de 0 (qui est identifié avec 1 dans T), il apparaît sans difficulté que

$$\begin{aligned}\int_{[-1-a,1+a]^2} & \mathbb{1}_{[x-1/2,x+1/2]}(y)(f(y)-f(x))^2\,dxdy \\ \geq\ & 4(1+a)^2\int_{T^2}\mathbb{1}_{[x-1/(8(1+a)),x+1/(8(1+a))]}(y)(\widetilde{f}(y)-\widetilde{f}(x))^2\,\widetilde{\mu}(dx)\widetilde{\mu}(dy)\end{aligned}$$

où $\widetilde{\mu}$ est la mesure de Lebesgue sur T, et

$$\begin{aligned}\int_{[-1-a,1+a]} & \left(f(x)-\int_{[-1-a,1+a]}f(y)\,dy/(2(1+a))\right)^2 dx \\ =\ & 4(1+a)\int_T(\widetilde{f}(x)-\widetilde{\mu}(f))^2\,\widetilde{\mu}(dx)\end{aligned}$$

Ainsi le trou spectral de P_2 dans $\mathbb{L}^2(\mu_2)$ est minoré par celui de $\widetilde{P}$ dans $\mathbb{L}^2(\widetilde{\mu})$, multiplié par $\exp(-A)/8$, où $\widetilde{P}$ est le noyau de convolution défini sur le tore T par

$$\forall\ x\in T,\qquad \widetilde{P}(x,dy)\ =\ 4(1+a)\mathbb{1}_{[x-1/(8(1+a)),x+1/(8(1+a))]}(y)\,\widetilde{\mu}(dy)$$

le segment ci-dessus étant toujours compris modulo 1.

Or il est facile d'obtenir par une décomposition en série de Fourier (voir [7]) une minoration du trou spectral $\lambda(\widetilde{P})$ de $\widetilde{P}$

$$\lambda(\widetilde{P})\ \geq\ 1-\frac{|\exp(i4\pi/(8(1+a))-1|}{4\pi/(8(1+a))}\ >\ 0$$

d'où le résultat annoncé.

4 Existence du trou spectral sur $\mathbb{R}$ ou comment recoller les morceaux

Afin de montrer l'existence du trou spectral sur $\mathbb{R}$, on va chercher à se ramener à 5 points. On introduit pour cela les intervalles

$$I_1=]-\infty,-a],\ \ I_2=[-a-1,a+1],\ \ I_3=[a,\infty[,$$

et une partition de $\mathbb{R}$ en 5 ensembles

$$A_1=]-\infty,-a-1],\ \ A_2=]-a-1,-a],\ \ A_3=]-a,a[,\ \ A_4=[a,a+1[,\ \ A_5=[a+1,\infty[$$

qui seront représentés par les 5 points auxquels on veut se ramener.

Sur chaque I_i on dispose du noyau $P_i(x,dy)=f_x(y)\mathbb{1}_{I_i}(y)dy+a_{i,x}\delta_x(dy)$ réversible par rapport à la mesure $\mu_i(dx)=K_i^{-1}\mu(dx)$ avec $K_i=\int_{I_i}\mu(dx)$. D'après les sections précédentes, sur chaque I_i, le noyau P_i admet un trou spectral $\lambda_i>0$ sous μ_i, on a donc pour tout $i=1$ à 3, les inégalités

$$(15)\qquad \mu_i(f(\mathrm{Id}-P_i)f)\geq\lambda_i\mu_i(f-\mu_i(f))^2\ .$$

Le recollement des morceaux s'effectue en plusieurs étapes. La première d'entre elles consiste à construire une nouvelle forme à partir de laquelle on se ramènera à 5 points (deuxième étape). On finit la démonstration en montrant que notre nouveau noyau sur 5 points admet un trou spectral, on utilisera pour cela un résultat de Cheeger.

Première étape : construction d'une nouvelle forme

Il est bien connu que la forme de Dirichlet $\mathcal{E}_{P,\mu}(f,f) \stackrel{\text{déf.}}{=} \mu(f(\mathrm{Id}-P)f)$ s'écrit sous la forme $\mathcal{E}_{P,\mu}(f,f) = \frac{1}{2}\int_{\mathbb{R}} d\mu(x)\int_{\mathbb{R}} P(x,dy)(f(x)-f(y))^2$, on montre alors facilement

$$\begin{aligned}
\mathcal{E}(f,f) &= \frac{1}{2}\int_{\mathbb{R}} d\mu(x)\int_{\mathbb{R}} P(x,dy)(f(x)-f(y))^2 \\
&= \frac{1}{2}\sum_{i=1}^{5}\int_{A_i} d\mu(x)\int_{\mathbb{R}} P(x,dy)(f(x)-f(y))^2 \\
&\geq \frac{1}{4}\sum_{i=1}^{3}\int_{I_i\times I_i} d\mu(x)P(x,dy)(f(x)-f(y))^2 \\
&= \frac{1}{4}\sum_{i=1}^{3}\int_{I_i\times I_i} K_i d\mu_i(x)P_i(x,dy)(f(x)-f(y))^2 \\
&= \frac{1}{2}\sum_{i=1}^{3} K_i\mathcal{E}_{P_i,\mu_i}(f_i,f_i) \quad \text{où } f_i = f\mathbb{1}_{I_i} \\
&\geq \frac{1}{2}\sum_{i=1}^{3} K_i\lambda_i\mu_i((f_i-\mu_i(f_i))^2)
\end{aligned}$$

Pour la troisième inégalité, on compte au plus deux fois chaque passage de A_i à A_j ($j \neq i$), pour la dernière inégalité, on utilise l'inégalité (15) trois fois.

On construit alors la nouvelle forme

$$\tilde{\mathcal{E}}(f,f) \stackrel{\text{déf.}}{=} \frac{1}{2}\sum_{i=1}^{3} K_i\lambda_i\mu_i(f_i-\mu_i(f_i))^2 \,.$$

Il existe un générateur de sauts L, vu comme un noyau d'intensités de transition $L(x,dy)$, tel que

$$\tilde{\mathcal{E}}(f,f) = -\mu(fLf).$$

Un calcul élémentaire où l'on décompose $\tilde{\mathcal{E}}(f,f)$ en $\sum_{i,j=1}^{5}\int_{A_i}\int_{A_j}\mu(dx)L(x,dy)f(y)$, et une identification nous donnent

$$\begin{aligned}
&L(x,\mathbb{1}_{\mathbb{R}\setminus\{x\}}(y)dy) = \\
&\quad \mathbb{1}_{A_1}(x)\frac{1}{2}\lambda_1\mu_1(dy) + \mathbb{1}_{A_2}(x)\left[\frac{1}{2}\lambda_1\mu_1(dy)\mathbb{1}_{A_1}(y) + \left[\frac{1}{2}\lambda_1\mu_1(dy) + \frac{1}{2}\lambda_2\mu_2(dy)\right]\mathbb{1}_{A_2}(y)\right. \\
&\quad \left. + \frac{1}{2}\lambda_2\mu_2(dy)\mathbb{1}_{A_3}(y)\right] + \mathbb{1}_{A_3}(x)\frac{1}{2}\lambda_2\mu_2(dy) + \mathbb{1}_{A_4}(x)\left[\frac{1}{2}\lambda_2\mu_2(dy)\mathbb{1}_{A_3}(y)\right. \\
&\quad \left. + \left[\frac{1}{2}\lambda_2\mu_2(dy) + \frac{1}{2}\lambda_3\mu_3(dy)\right]\mathbb{1}_{A_4}(y) + \frac{1}{2}\lambda_3\mu_3(dy)\mathbb{1}_{A_5}(y)\right] + \mathbb{1}_{A_5}(x)\frac{1}{2}\lambda_3\mu_3(dy)
\end{aligned}$$

et

$$L(x,\{x\}) = -L(x,\mathbb{R}\backslash\{x\})$$

Le générateur s'écrit

$$L(x,dy) = L(x,\mathbb{1}_{\mathbb{R}\backslash\{x\}}(y)dy) + L(x,\{x\})\delta_x(dy)$$

En résumé, on a montré

$$\mathcal{E}_{P,\mu}(f,f) \geq \tilde{\mathcal{E}}(f,f)\ . \tag{16}$$

On va désormais travailler avec cette nouvelle forme.

Deuxième étape : on se ramène à 5 points

Soit maintenant $\mathcal{F} = \sigma(A_1, A_2, \ldots, A_5)$ et $F = \mu(f|\mathcal{F})$ (si $x \in A_i$, $F(x) = \frac{\mu(f\mathbf{1}_{A_i})}{\mu(A_i)}$). Notons également, pour $1 \leq i \leq 5$, $F_i = \mathbb{1}_{A_i}(f - F)$ et $l_i = -L(x,\{x\}) > 0$, qui est une quantité ne dépendant pas du choix du point $x \in A_i$.

On calcule alors que

$$\widetilde{\mathcal{E}}(f,f) \;=\; \widetilde{\mathcal{E}}(F,F) + \sum_{1\leq i\leq 5} l_i\mu(F_i^2)$$

et que

$$\mu((f-\mu(f))^2) \;=\; \mu((F-\mu(F))^2) + \sum_{1\leq i\leq 5} \mu(F_i^2)$$

Il apparaît ainsi que

$$\lambda \geq \inf_{f\in\mathbb{L}^2(\mu)\backslash\mathrm{Vect}(\mathbf{1})} \frac{\widetilde{\mathcal{E}}(f,f)}{\mu(f-\mu(f))^2} \;=\; \lambda_{\check{L}} \wedge \min_{1\leq i\leq 5} l_i$$

où

$$\lambda_{\check{L}} \overset{\text{déf.}}{=} \inf_{f\in\mathbb{L}^2(\mathcal{F},\mu)\backslash\mathrm{Vect}(\mathbf{1})} \frac{\widetilde{\mathcal{E}}(f,f)}{\mu(f-\mu(f))^2}$$

Ceci nous ramène au calcul du trou specral pour un processus irréductible à valeurs dans un ensemble à 5 points (d'autant plus qu'en considérant l'indicatrice de A_i, pour $1 \leq i \leq 5$, on se persuade facilement que $l_i \geq \lambda_{\check{L}}$), et on va rappeler dans l'étape suivante une manière de minorer cette quantité par un nombre strictement positif, ce qui nous permettra de conclure.

Troisième étape : existence du trou spectral sur les 5 points

Sur les 5 points $A_1, A_2, \ldots, A_5$, on dispose d'une matrice $\check{L}$, 5×5, tridiagonale, symétrique dans $\mathbb{L}^2(\check{\mu})$ où $\check{\mu}$ est la mesure définie sur A_i par $\check{\mu}(A_i) = \mu(A_i)$ pour tout $i = 1$ à 5 . Les coefficients de cette matrice nous sont donnés explicitement par L.

Il est bien connu que ce noyau d'intensités irréductible admet un trou spectral. Pour l'évaluer, on a choisit d'utiliser la constante de Cheeger (voir [3], [5], ou [8])

$$I(\check{L},\check{\mu}) \overset{\text{déf.}}{=} \inf_{\substack{0<\check{\mu}(A)\leq\frac{1}{2}\\ A\text{ connexe}\\ A^c\text{ connexe}}} \frac{\check{\mu}(\mathbb{1}_A\check{L}\mathbb{1}_{A^c})}{\check{\mu}(A)}$$

où la connexité est celle relative à la matrice d'incidence de $\check{L}$. On sait que dans le cas d'un espace d'état à 5 points, le trou spectral $\lambda_{\check{L}}$ du noyau $\check{L}$ associé à la mesure $\check{\mu}$ est directement lié à la constante de Cheeger par la relation

$$2I(\check{L},\check{\mu}) \geq \lambda_{\check{L}} \geq \frac{I(\check{L},\check{\mu})}{5}$$

La connexité de A et A^c réduit l'étude à

$$A = [A_1, A_i] \ \text{ et } \ A^c = [A_{i+1}, A_5] \quad \text{ou} \quad A^c = [A_1, A_i] \ \text{ et } \ A = [A_{i+1}, A_5]$$

On en déduit immédiatement

$$I(\check{L},\check{\mu}) \quad \geq \quad \min_{i,j\,:\,\check{L}(A_i,A_j)>0} \check{L}(A_i,A_j) \quad = \quad \min_{i,j\,:\,\mu(\mathbb{1}_{A_i}L(\mathbb{1}_{A_j}))>0} \frac{\mu(\mathbb{1}_{A_i}L(\mathbb{1}_{A_j}))}{\mu(A_i)}$$

5 Temps discret

Comme on l'a déjà mentionné, le trou spectral λ fournit une vitesse de convergence vers l'équilibre μ du processus de Markov de sauts associé au semi-groupe qui admet $P - \mathrm{Id}$ comme générateur. Mais on peut aussi s'intéresser en temps discret à la chaîne de Markov qui admet P pour noyau de probabilités de transition. Il est bien connu (cf. par exemple [4]) que pour une telle chaîne, une vitesse de convergence exponentiellement rapide en le temps (dans $\mathbb{L}^2(\mu)$) est notamment impliquée par l'existence d'un trou spectral $\widetilde{\lambda}$ non nul pour $\mathrm{Id} - PP^*$ (où P^* est l'adjoint de P dans $\mathbb{L}^2(\mu)$), cette constante permettant d'expliciter une minoration du taux : plus précisément, pour tout $f \in \mathbb{L}^2(\mu)$, on a

$$\|P^n(f) - \mu(f)\|_2 \quad \leq \quad \rho^n \sqrt{\mu(f^2) - \mu(f)^2}$$

avec $\rho = \sqrt{1 - \widetilde{\lambda}}$ et où $\|\cdot\|_2$ désigne la norme dans $\mathbb{L}^2(\mu)$.

Par réversibilité, on a ici $P^* = P$, et $\widetilde{\lambda}$ est donc le trou spectral de $\mathrm{Id} - P^2$. Sans vouloir non plus présenter une minoration quantitative, on va se contenter d'indiquer le résultat suivant :

Proposition 3 *Sous les hypothèses précédentes, on est assuré que* $\widetilde{\lambda} > 0$.

En effet, remarquons que pour $x \geq a$, on a

$$\begin{aligned} a_x &= 1 - \int_{[x-1/2,x+1/2]} \exp(-(U(y) - U(x))_+)\, dy \\ &\geq \frac{1}{2} - \int_{[x,x+1/2]} \exp(-\alpha(y-x))\, dy \\ &= \frac{1}{\alpha}(\exp(-\alpha/2) - 1 + \alpha/2) \end{aligned}$$

De même pour $x \leq -a$, ce qui nous conduit à poser

$$\eta \stackrel{\text{déf.}}{=} \frac{1}{\alpha}(\exp(-\alpha/2) - 1 + \alpha/2) \ > \ 0$$

Comme dans la section 2, considérons R_1 la restriction du noyau P^2 à $[a, +\infty[$, et rappelons que P_1 et μ_1 désignent respectivement la restriction de P et la renormalisation de la restriction de μ à cet intervalle. Pour tout $f \in \mathbb{L}^2(\mu)$, on constate immédiatement que

$$\begin{aligned}\int \mu_1(x) R_1(x, dy)(f(y) - f(x))^2 &\geq \int \mu_1(x) P_1^2(x, dy)(f(y) - f(x))^2 \\ &\geq \eta \int \mu_1(x) P_1(x, dy)(f(y) - f(x))^2\end{aligned}$$

ainsi d'après les calculs de la section 2, le couple réversible (R_1, μ_1) admet un trou spectral strictement positif.

De la même manière, on traite la restriction de P^2 à $]-\infty, -a]$. Le fait que la restriction de P^2 à $[-1-a, 1+a]$ admette également un trou spectral se prouve comme dans la section 3, et on conclut de manière identique aux considérations de la section précédente, à l'existence d'un trou spectral $\widetilde{\lambda}$ strictement positif.

Voyons maintenant comment à partir de la stricte positivité de $\widetilde{\lambda}$, on peut obtenir les estimées indiquées à la fin de l'introduction :

Soit $x_0 \in \mathbb{R}$ fixé, et notons m la probabilité admettant la densité $\widetilde{f}_{x_0}(\,\cdot\,) \stackrel{\text{déf.}}{=} K f_{x_0}(\,\cdot\,) \exp(U(\,\cdot\,))/(1 - a_{x_0})$ (sous-entendu par rapport à μ).

Pour $n \geq 0$, la probabilité mP^n admet alors pour densité $P^{*n}(\widetilde{f}_{x_0}) = P^n(\widetilde{f}_{x_0})$.

Soit une fonction $g \in \mathbb{L}^2(\mu)$ telle que $\mu(g) = 0$, il apparaît que pour tout $n \geq 0$,

$$P^n(x_0, g) = \sum_{1 \leq k \leq n} a_{x_0}^{n-k}(1 - a_{x_0})(mP^{k-1})(g) + a_{x_0}^n g(x_0)$$

or d'après les considérations précédentes, pour $k \geq 1$,

$$\begin{aligned}(mP^{k-1})(g) &= \mu(P^{k-1}(\widetilde{f}_{x_0})g) \\ &= \mu[(P^{k-1}(\widetilde{f}_{x_0}) - \mu(\widetilde{f}_{x_0}))g] \\ &\leq \sqrt{\mu[(P^{k-1}(\widetilde{f}_{x_0}) - \mu(\widetilde{f}_{x_0}))^2]}\sqrt{\mu(g^2)} \\ &\leq \rho^{k-1}\sqrt{\mu(\widetilde{f}_{x_0}^2) - \mu(\widetilde{f}_{x_0})^2}\sqrt{\mu(g^2)} \\ &\leq \rho^{k-1}\left\|\widetilde{f}_{x_0}\right\|_2 \|g\|_2\end{aligned}$$

Notons par ailleurs que

$$\mu(\widetilde{f}_{x_0}^2) = \frac{K}{(1 - a_{x_0})^2}\int_{[x_0 - 1/2, x_0 + 1/2]} \exp(-U(x) + 2U(y) - 2(U(y) - U(x))_+)\, dy$$

ainsi il apparaît facilement que pour $x_0 \geq 1 + a$,

$$\mu(\widetilde{f}_{x_0}^2) \leq \frac{K}{(1 - a_{x_0})^2}\exp(U(x_0))$$

Il en découle qu'il existe une constante $A \geq K$ telle que

$$\forall\, x_0 \in \mathbb{R}, \qquad \mu(\widetilde{f}_{x_0}^2) \leq \frac{A}{(1 - a_{x_0})^2}\exp(U(x_0))$$

d'où en fait

$$
\begin{aligned}
|P^n(x_0,g)| &\leq A\exp(U(x_0))\sum_{1\leq k\leq n} a_{x_0}^{n-k}\rho^k\left\|g\right\|_2 + a_{x_0}^n g(x_0)\\
&\leq \begin{cases} A\exp(U(x_0))a_{x_0}\dfrac{a_{x_0}^n-\rho^n}{a_{x_0}-\rho}\left\|g\right\|_2 + a_{x_0}^n g(x_0) & \text{, si } \rho\neq a_{x_0}\\ (An\exp(U(x_0))\left\|g\right\|_2+\rho g(x_0))\rho^n & \text{, si } \rho = a_{x_0}\end{cases}
\end{aligned}
$$

ce qui permet d'établir le résultat annoncé, en remplaçant g par $g-\mu(g)$, pour $g\in \mathbb{L}^2(\mu)$ quelconque.

Un autre avantage du trou spectral par rapport aux choix d'une fonction de Lyapounov est que le taux de convergence est plus explicite, ici $a_{x_0}\vee\rho\leq(1/2)\vee\rho$.

Ainsi par exemple avec $U\,:\,\mathbb{R}\ni x\mapsto x^2/2$, en itérant les noyaux P on finit par approximer la loi gaussienne standard relativement vite.

Références

[1] D. Bakry. L'hypercontractivité et son utilisation en théorie des semigroupes. In P. Bernard, editor, *Lectures on Probability Theory. Ecole d'Eté de Probabilités de Saint-Flour XXII-1992*, Lecture Notes in Mathematics 1581. Springer-Verlag, 1994.

[2] H.J Brascamp and E.H. Lieb. On extensions of the Brunn-Minkowski and Prékopa-Leindler theorems, including inequalities for log concave functions, and with an application to the diffusion equation. *Journal of Functional Analysis*, 22:366–389, 1976.

[3] Cheeger J., *A lower bound for smallest eigenvalue of the Laplacian*, Problems in Analysis, Symposium in honor of S. Bochner, Princeton University Press, (1970), 195-199.

[4] Fill J.A., *Eigenvalue bounds on convergence to stationarity for nonreversible Markov chains, with an application to the exclusion process*, The Annals of Applied Probability, (1991) 1 (1), 62-87.

[5] Lawler G. and Sokal A., *Bounds on the L^2 spectrum for Markov chains and Markov processes : a generalization of Cheeger inequality*, Transactions of the American Mathematical Society, (1988) 309 (2), 557-580.

[6] S.P. Meyn and R.L. Tweedie. *Markov Chains and Stochastic Stability.* Springer-Verlag, 1993.

[7] Miclo L., *Trous spectraux à basse température : un contre-exemple à un comportement asymptotique escompté*, Séminaire de Probabilités XXXII, Lecture Notes in Mathematics 1686, Springer-Verlag, (1998), 36-55.

[8] Miclo L., *Une variante de l'inégalité de Cheeger pour les chaînes de Markov finies*, ESAIM : P&S, URL : http://www.emath.fr/ps/, (1998) 2, 1-21.

[9] J. Neveu. *Bases mathématiques du calcul des probabilités.* Masson, 1970.

[10] Rosenthal J.S., *Markov chain convergence : from finite to infinite*, Stochastic Process and their applications, (1996) 62, 55-72.

[11] L. Saloff-Coste. Lectures on finite Markov chains. In P. Bernard, editor, *Lectures on Probability Theory and Statistics. Ecole d'Eté de Probabilités de Saint-Flour XXVI-1996*, Lecture Notes in Mathematics 1665. Springer-Verlag, Berlin, 1997.

[12] Sinclair A., *Improved bounds for mixing rates of Markov chains and multicommodity flow*, Combinatories, proba. comput., (1992) 1, 351-370.