References

[1] V. Bally, Approximation for the solutions of stochastic differential equations: I L^p-Convergence, *Stoch. and Stoch. Rep.* **28**, 209–246.

[2] H. Doss and P. Priouret, Support d'un processus de reflexion, *Z. Wahrsch. verw. Gebiete* **61** (1982), 327–345.

[3] N. Ikeda and S. Watanabe, Stochastic Differential Equations and Diffusion Processes, North Holland-Kodanasha, 1981.

[4] N. Ikeda, S. Nakao, and Y. Yamato, A class of approximations of Brownian motion, *Publ. RIMS. Kyoto Univ.* **13** (1977), 285–300.

[5] K. Ito and M. Niso, On stationary solutions of stochastic differential equations, *J. Math. Kyoto Univ.* **4** (1964), 1–75.

[6] V. Mackevicius, S^p stability of solutions of symmetric stochastic differential equations with discontinuous driving semimartingales, Annales d'Institut Henri Poincare, no. 4 (1987), 575–592.

[7] E. J. McShane, Stochastic differential equations, *J. Multivariate Anal.* **5** (1975), 121–177.

[8] S. Nakao and Y. Yamato, Approximation theorems on stochastic differential equations, *Proc. of Intern. Symp. SDE, Kyoto* (1976), 283–296.

[9] P. Protter, Approximations of solutions of stochastic differential equations driven by semimartingales. *The Annals of Prob.* **13** (1985), 716–743.

[10] E. Wong and M. Zakai, on the relation between ordinary and stochastic differential equations. *Intern. J. Eng. Sci.* **3** (1965), 213–229.

Stochastics and Stochastics Reports, Vol. 41, pp. 23–56
Reprints available directly from the publisher
Photocopying permitted by license only

RECUIT SIMULÉ SANS POTENTIEL SUR UNE VARIÉTÉ RIEMANNIENNE COMPACTE

LAURENT MICLO

Département de mathématiques, Université Louis Pasteur, 7, rue René Descartes, 67084 Strasbourg Cedex, France

(Received 25 July 1991; in final form 12 December 1992)

We consider general simulated annealing algorithms on a connected and compact Riemannian manifold, i.e. algorithms for which the drift is not necessarily the gradient of a given potential. When this is satisfied (or equivalently, when the process considered is instantaneously reversible), Holley, Kusuoka and Stroock proved the convergence of the algorithm, for certain decreasing rate of the temperature. The method presented here is different and is very general (it can be adapted to the situation where the phase space is an infinite dimensional one on which we are given a translation invariant and finite range family of potentials). It is based on the study of the evolution of the entropy of the law of the diffusion at one instant with respect to the instantaneous invariant probability, but we also depend on the Sobolev logarithmic inequalities which can be easily deduced from the results of the previous authors. These inequalities give us a differential inequality for the entropy, which implies that it converges to zero under decreasing rates of the temperature of the kind $k/\ln(t)$, for k strictly larger than one certain constant $c \geq 0$. We can then prove that the stochastic process considered converge to the global minima of the quasi-potential introduced by Wentzell and Freidlin as the action functional for the large deviation principle satisfied by the invariant measure. But apart from the case where the manifold is the circle or the instantaneously reversible situation (for which there exists an explicit formula for the invariant probability), the constant is twice the natural generalisation of the constant which appear when the drift derive from a potential, so the constant c obtained is certainly not optimal.

KEY WORDS Simulated Annealing, Entropy, Harnack's Inequalities, Ergodicity, Sobolev Logarithmic Inequalities.

1 INTRODUCTION

On va étudier, sur une variété riemannienne compacte et connexe, des algorithmes du recuit simulé généraux, i.e. dont la dérive est un champ de vecteurs qui n'est pas nécessairement le gradient d'un potentiel (à un facteur dépendant du temps près) donné a priori. Asymptotiquement, quand la température tend vers 0, le rôle que jouait ce dernier dans les algorithmes instantanément réversibles (i.e. dont la dérive est le gradient d'une fonction) est tenu ici par une quasi-énergie introduite par Wentzell et Freidlin, comme la fonctionnelle d'action pour les grandes déviations satisfaites par la probabilité invariante. Ainsi, pour certaines décroissances de la température, le processus considéré tend à ce concentrer au voisinage des minima globaux de cette quasi-énergie. Cependant, mis à part le cas où la variété est le cercle, les taux de décroissance de la température sont un peu plus lents que la généralisation naturelle de ceux connus pour les algorithmes dérivant d'un potentiel, ce qui suggère qu'on n'a pas obtenu, en général, les taux optimaux. Notons que dans le cas où la dérive est le gradient d'une fonction, la convergence a déjà été démontrée par Holley, Kusuoka et Stroock dans [5]. Néanmoins la méthode présentée ici, basée sur l'étude de l'évolution de l'entropie de la loi de la diffusion à un instant donné par rapport à la probabilité stationnaire instantannée, est très simple et se généralise à de nombreuses situations. En effet, elle s'adapte facilement pour traiter le cas où l'espace des phases est un ensemble fini et permet également d'obtenir des résultats en dimension infinie, pour des algorithmes du recuit simulé associés à certaines measures de Gibbs (cf. [9]).

Soit M une variété riemannienne compacte et connexe (de dimension $N \in \mathbb{N}^*$). Comme d'habitude, on note $\langle\,,\rangle$, $|\cdot|$, ∇, div, Δ et λ, le produit scalaire, la norme, le gradient, la divergence, le laplacien et la probabilité associés à la structure riemannienne.

Soit b un champ de vecteurs C^∞ sur M. Pour $\beta \geq 0$, on considère l'opérateur L_β défini sur les fonctions ϕ de $C^2(M)$ par

$$L_\beta \phi = \frac{1}{2}\Delta\phi + \beta\langle b, \nabla\phi\rangle$$

Il est connu (voir, par exemple, la proposition 4.5 p. 278 de [7]) qu'il existe une unique probabilité invariante μ_β associée à L_β, c'est à dire qui satisfait

$$\forall \phi \in C^2(M), \quad \mu_\beta(L_\beta \phi) = 0$$

De plus, μ_β est équivalente à λ et sa densité est de classe C^∞. On conviendra, dans la suite, de noter de la même manière une probabilité absolument continue par rapport à λ et sa densité.

Les opérateurs L_β seront les générateurs des diffusions qu'on va étudier. Plus précisément, considérons le problème de martingales suivant.

On se place sur l'espace canonique $\Omega = C([0, +\infty[, M)$ et on note $(X_t)_{t\geq 0}$ le processus de ses applications coordonnées. On pose

$$\mathscr{F}_t = \sigma\{X_s, 0 \leq s \leq t\}$$

$$\mathscr{F} = \sigma\{X_t, t \geq 0\}$$

Soient $\beta \in C^1(\mathbb{R}_+; \mathbb{R}_+^*)$ (qui représentera l'évolution de l'inverse de la température), et m une probabilité sur M (qui sera la probabilité initiale). Il est bien connu qu'il existe une unique probabilitié P_m sur $(\Omega, \mathscr{F})$, telle que

- $X_0(P_m) = m$
- $\forall \phi \in C^\infty(M)$,

$$\left(\left(\phi(X_t) - \phi(X_0) - \int_0^t L_{\beta_s}\phi(X_s)\,ds\right)_{t\geq 0}, (\mathscr{F}_t)_{t\geq 0}, P_m\right) \quad \text{soit une martingale.}$$

Heuristiquement, ceci signifie que la diffusion $(X_t)_{t\geq 0}$ est régie par

$$dX_t = dB_t + \beta_t b(X_t)\,dt$$

où $(B_t)_{t\geq 0}$ est un mouvement brownien sur M.

Pour $t \geq 0$, on notera m_t la loi de X_t sous P_m.

Pour étudier le comportement asymptotique des m_t quand t tend vers l'infini, on est amené à introduire la notion suivante.

Pour $t \geq 0$, on définit l'entropie (aussi appelée information de Kullback-Leibler) d'une probabilité p sur M, par rapport à μ_{β_t}, par

$$I_t(p) = \begin{cases} \displaystyle\int_M \ln\left(\frac{p}{\mu_{\beta_t}}\right)(x)\,dp(x), & \text{si } p \ll \lambda \\ +\infty, & \text{sinon.} \end{cases}$$

$I_t(p)$ mesure d'une certaine manière la distance entre p et μ_{β_t}, notamment, on a

$$\|p - \mu_{\beta_t}\| \leq 4\sqrt{2I_t(p)}$$

où $\|\cdot\|$ représente la variation totale (cf. [6]). Notons que d'après l'inégalité de Jensen, on a $I_t(p) \geq 0$.

Remarquons également que μ_{β_t} est la probabilité invariante instantanée, au temps t, de la diffusion $(X_t)_{t\geq 0}$ (si β était constant, μ_{β_0} serait la probabilité stationnaire pour le processus).

On se propose de démontrer le résultat suivant.

THÉORÈME 1 *Il existe une constante $\chi \geq 0$ telle que si l'on suppose que pour t suffisamment grand, on ait $\beta(t) = \ln(t)/K$, avec $K > \chi$, alors,*

$$\lim_{t\to+\infty} I_t(m_t) = 0$$

L'intérêt de ce théorème est qu'on peut en déduire des résultats de recuit simulé. En effet, supposons que $\beta^{-1}\ln(\mu_\beta)$ converge uniformément sur M, quand β tend vers l'infini, vers une fonction W. Alors, sous les hypothèses du théorème, les m_t tendent à se concentrer, en temps grand, au voisinage des minima globaux de $-W$, on a notamment, pour tout $\delta > 0$, $\lim_{t\to+\infty} m_t(-W \geq \delta) = 0$ (on aura remarqué que nécessairement, $\min_{x\in M}(-W(x)) = 0$), ce qui constitue un résultat de recuit simulé pour des diffusions non instantanément réversibles, généralisant ainsi des résultats connus pour les algorithmes dérivant d'un potentiel (voir, par exemple, [5]. Pour ceux-ci on peut supposer que $b = -\frac{1}{2}\nabla U$, où $U \in C^\infty(M)$, et il est facile de voir qu'alors, $W = -(U - \min U)$). Dans certains cas, on peut également obtenir des résultats plus précis sur le comportement asymptotique des m_t en temps grand. Ainsi, si on sait que les μ_β convergent étroitement quand β tend vers $+\infty$, il en sera de même des m_t, quand t tend vers $+\infty$.

Reste donc à voir sous quelles conditions les $\beta^{-1}\ln(\mu_\beta)$ convergent uniformément. Mais un résultat classique d'analyse (voir la section suivante) affirme qu'il existe une constante $K_1 > 0$ telle que pour tout $\beta \geq 0$ on ait,

$$|\nabla\beta^{-1}\ln(\mu_\beta)| \leq K_1$$

Ceci permet de voir que la famille $(\beta^{-1}\ln(\mu_\beta))_{\beta\geq 0}$ est équicontinue et uniformément bornée (car il existe un point de M en lequel $\beta^{-1}\ln(\mu_\beta)$ s'annule et le diamètre de M est fini). Ainsi, puisque M est compact, on peut appliquer le théorème d'Ascoli, qui montre que la famille $(\beta^{-1}\ln(\mu_\beta))_{\beta\geq 0}$ est relativement compacte dans $C(M)$ muni de la norme uniforme.

Or sous certaines hypothèses (la condition (A) donnée par Wentzell et Freidlin à la page 169 de leur livre [4], ou sans hypothèse, si M est le cercle, voir l'appendice 1), les μ_β satisfont un principe de grandes déviations avec pour taux β et relativement à une certaine fonctionnelle d'action W (cf. le théorème 4.3 p. 190 de [4]), ce qui prouve que toute limite uniforme des $\beta^{-1}\ln(\mu_\beta)$ ne peut qu'être W, d'où, par compacité, la convergence uniforme des $\beta^{-1}\ln(\mu_\beta)$ vers W.

D'autre part, la démonstration suivante permet également de retrouver le théorème ergodique qui affirme que si la température est constante, m_t converge exponentiellement vite (au sens de l'entropie et donc aussi de la variation totale) vers la probabilité stationnaire μ_{β_0} (voir la proposition 7 ci-dessous).

2 INÉGALITÉS DE SOBOLEV LOGARITHMIQUES

On va présenter ici un résultat très important sur les inégalités de Sobolev logarithmiques, qui a été obtenu par Holley, Kusuoka et Stroock (regrouper le théorème (1.14) de [6] et le théorème (3.23) de [5]). On donnera aussi une description de la constante χ apparaissant dans le théorème 1.

A tout $U \in C(M)$ on associe la constante $c(U)$ suivante.

Pour $x, y \in M$, on note $\mathscr{C}_{x,y}$ l'ensemble des applications continues ϕ: $[0, 1] \to M$ telles que $\phi(0) = x$ et $\phi(1) = y$. Pour $\phi \in \mathscr{C}_{x,y}$ on définit l'élévation de ϕ relativement à U par

$$e_U(\phi) = \sup_{0\leq t\leq 1} U(\phi(t)) - U(x) - U(y) + \inf_{z\in M} U(z)$$

puis la constante $c(U)$ par

$$c(U) = \sup_{x,y\in M}\left(\inf_{\phi\in\mathscr{C}_{x,y}} e_U(\phi)\right)$$

Notons que si on suppose que les

$$W_\beta \overset{\text{def}}{=} \beta^{-1}\ln(\mu_\beta)$$

convergent uniformément vers une fonction W, alors

$$\lim_{t\to+\infty} c(-W_\beta) = c(-W)$$

L'intérêt de cette constante provient du résultat suivant. On définit, pour $\beta > 0$, la densité de probabilité ν_β par

$$\nu_\beta(x) = \left(\int \exp(-\beta U(y))\lambda(dy)\right)^{-1} \exp(-\beta U(x))$$

Le résultat de Holley, Kusuoka et Stroock s'énonce alors

Lemme 2 *Il existe un nombre $A > 0$, ne dépendant que de M et de $\|U\|_\infty$, tel que si N est la dimension de M, les inégalités de Sobolev logarithmiques suivantes soient satisfaites,*

$$\forall\beta \geq 0, \quad \forall\phi \in C^1(M), \quad \int \phi^2\ln(\phi^2)\,d\nu_\beta$$
$$\leq A(1 + \beta\|\,|\nabla U|\,\|_\infty)^{5N-1}\exp(c(U)\beta)\int\langle\nabla\phi, \nabla\phi\rangle\,d\nu_\beta + \int\phi^2\,d\nu_\beta\ln\left(\int\phi^2\,d\nu_\beta\right)$$

Ce lemme nous permettra de montrer qu'on peut prendre

$$\chi = 2\limsup_{\beta\to+\infty} c(-W_\beta) < +\infty$$

dans le théorème 1. Ainsi, si les W_β convergent uniformément vers une fonction W, on peut prendre $\chi = 2c(-W)$. Néanmoins dans le cas instantannément réversible, ou dans le cas où M est le cercle, on verra qu'on peut prendre $\chi = c(-W)$, et on se demande si cela n'est pas vrai dans le cas général.

3 QUELQUES PRÉLIMINAIRES SUR LA PROBABILITÉ INVARIANTE

On se propose ici de prouver les résultats suivants, techniques mais importants pour la suite.

Proposition 3 *Il existe une constante $K_1 \geq 0$ telle que*

$$\forall\beta \geq 1, \quad \forall x \in M, \quad |\nabla\beta^{-1}\ln(\mu_\beta)|(x) \leq K_1$$

Proposition 4 *Il existe une constante $c' \geq 0$ telle que pour tout $\tilde{c} > c'$, il existe une constante $K_2 \geq 0$ telle que*

$$\forall\beta \geq 1, \quad \forall x \in M, \quad \left|\frac{\partial}{\partial\beta}\ln(\mu_\beta)\right|(x) \leq K_2\exp(\tilde{c}\beta)$$

On verra, au cours de la démonstration, qu'on peut toujours prendre

$$c' = \limsup_{\beta \to +\infty} c(-W_\beta)$$

dans cette proposition. Mais on montrera, dans l'appendice 1, en utilisant une formule explicite pour la probabilité invariante, que si M est le cercle, on peut prendre $\bar{c} = 0$ dans cette proposition. Notons que dans le cas réversible, elle est immédiate avec $\bar{c} = 0$. Dans le cas général, on aurait bien voulu obtenir la proposition 4 avec $c' = 0$.

Commençons par montrer la proposition 3, qui repose sur le résultat classique suivant (la démonstration est inspiré d'un calcul (dit méthode de Bernstein) p. 5 de [3]. Je remercie Gérard Ben Arous et Benoît Perthame de m'avoir donné des références sur ce sujet. Pour une généralisation de la méthode de Bernstein, cf. [1]).

LEMME 5 *Soit Ω le cube $\{x \in \mathbb{R}^n / \forall 1 \le i \le n, |x_i| < 1\}$. Soient, pour $1 \le i, j \le n$, a_{ij}, b_i, $\tilde{b}_i$, c des fonctions de classe C^∞, telles que les dérivées premières (respectivement, les dérivées secondes) des b_i, des $\tilde{b}_i$ et de c (respectivement, des a_{ij}) soient bornées sur Ω. On suppose que pour tout $x \in \Omega$, $A(x) = (a_{ij}(x))_{1 \le i,j \le n}$ est une matrice symétrique définie positive, uniformément en x, c'est à dire qu'il existe une constante $\theta > 0$ telle que*

$$\forall x \in \Omega, \quad \forall v = (v_i)_{1 \le i \le n} \in \mathbb{R}^n, \quad \sum_{i,j} a_{ij}(x) v_i v_j \ge \theta \sum_i v_i^2$$

Soit $u \in C^2(\Omega)$ une solution sur Ω de

$$\sum_{i,j} a_{ij} u_{ij} + \sum_{i,j} a_{ij} u_i u_j + \beta \sum_i b_i u_i + \sum_i \tilde{b}_i u_i = \beta c$$

où β est un réel positif, et où on a noté

$$u_i = \frac{\partial u}{\partial x_i}$$

On suppose que la fonction $z = \sum_{i,j} a_{ij} u_i u_j$ admet un maximum local dans Ω. Notons m la valeur de z en ce point. Il existe alors une constante $K > 0$, ne dépendant que de θ et des normes $\|a_{ij}\|_{C^2(\Omega)}$, $\|b_i\|_{C^1(\Omega)}$, $\|\tilde{b}_i\|_{C^1(\Omega)}$ et $\|c\|_{C^1(\Omega)}$, pour $1 \le i, j \le n$, telle que pour tout $\beta \ge 0$ on ait

$$m \le K(1 + \beta^2)$$

Démonstration: Soit $x_0 \in \Omega$ un maximum local de z. On a donc en ce point, pour tout $1 \le k \le n$,

$$z_k = \sum_{i,j} a_{ij,k} u_i u_j + a_{ij} u_{ik} u_j + a_{ij} u_i u_{jk} = 0 \tag{$*$}$$

et pour toute matrice $M = (m_{ij})_{1 \le i,j \le n}$ définie positive,

$$\sum_{kl} m_{kl} z_{kl} \le 0$$

Prenons $M = A(x_0)$, on obtient

$$\begin{aligned} 0 &\ge \sum_{k,l} a_{kl} \sum_{i,j} [a_{ij,kl} u_i u_j + a_{ij,k} u_{il} u_j + a_{ij,k} u_i u_{jl} + a_{ij,l} u_{ik} u_j + a_{ij,l} u_i u_{jk} \\ &\quad + a_{ij}(u_{ikl} u_j + u_{ik} u_{jl} + u_{il} u_{jk} + u_i u_{jkl})] \\ &\overset{\text{par symétrie}}{=} \sum_{k,l} \sum_{i,j} a_{kl}[a_{ij,kl} u_i u_j + 4 a_{ij,k} u_{il} u_j + 2 a_{ij}(u_{ikl} u_j + u_{ik} u_{jl})] \end{aligned}$$

Désignons par ($**$) cette expression et intéressons-nous d'abord au terme

$$\begin{aligned} \sum_{k,l} \sum_{i,j} a_{ij} a_{kl} u_{ikl} u_j &= \sum_{i,j} a_{ij} u_j \left[\left(\sum_{k,l} a_{kl} u_{kl} \right)_i - \sum_{kl} a_{kl,i} u_{kl} \right] \\ &= \sum_{i,j} a_{ij} u_j \left[-\left(\sum_{k,l} a_{kl} u_k u_l + \beta \sum_k b_k u_k + \sum_k \tilde{b}_k u_k - \beta c \right)_i - \sum_{kl} a_{kl,i} u_{kl} \right] \\ &= \sum_{i,j} a_{ij} u_j \Bigg[-\beta \sum_k (b_{k,i} u_k + b_k u_{ki}) - \sum_k (\tilde{b}_{k,i} u_k + \tilde{b}_k u_{ki}) \\ &\quad + \beta c_i - \sum_{k,l} a_{kl,i} u_{kl} \Bigg] \end{aligned}$$

d'aprés ($*$), ainsi,

$$\begin{aligned} \left| \sum_{k,l} \sum_{i,j} a_{ij} a_{kl} u_{ikl} u_j \right| &\le \beta \sqrt{\sum_{i,j,k} a_{ij}^2 u_k^2} \sqrt{\sum_{i,j,k} b_{k,i}^2 u_j^2} + \beta \sqrt{\sum_{i,j,k} a_{ij}^2 b_k^2} \sqrt{\sum_{i,j,k} u_j^2 u_{ki}^2} \\ &\quad + \sqrt{\sum_{i,j,k} a_{ij}^2 u_k^2} \sqrt{\sum_{i,j,k} \tilde{b}_{k,i}^2 u_j^2} + \sqrt{\sum_{i,j,k} a_{ij}^2 \tilde{b}_k^2} \sqrt{\sum_{i,j,k} u_j^2 u_{ki}^2} \\ &\quad + \beta \sqrt{\sum_{i,j} a_{ij}^2} \sqrt{\sum_{i,j} c_i^2 u_j^2} + \sqrt{\sum_{i,j,k,l} a_{ij}^2 u_{kl}^2} \sqrt{\sum_{i,j,k,l} a_{kl,i}^2 u_j^2} \\ &\le \beta |A| |\nabla b| |\nabla u|^2 + \beta |A| |b| |\nabla u| |\nabla^2 u| \\ &\quad + |A| |\nabla \tilde{b}| |\nabla u|^2 + |A| |\tilde{b}| |\nabla u| |\nabla^2 u| \\ &\quad + \beta |\nabla c| |A| |\nabla u| + |\nabla A| |A| |\nabla u| |\nabla^2 u| \end{aligned}$$

(si v est une fonction de $C^\infty(\Omega, \mathbb{R}^p)$ (ici $p = 1$, n ou n^2), on a noté ci-dessus ∇v la fonction de $C^\infty(\Omega, \mathbb{R}^{pn})$ obtenue en posant

$$\nabla v = \left(\frac{\partial v}{\partial x_i} \right)_{1 \le i \le n}$$

et en identifiant $(\mathbb{R}^p)^n$ à $\mathbb{R}^{np}$, en mettant, par exemple, les colonnes les une au dessus des autres. ∇^2 est alors l'itéré de ∇. Toutes les normes $|\cdot|$ représentent les normes euclidiennes usuelles dans les différents $\mathbb{R}^q$ considérés ($q = n$, n^2 ou n^3)).

De même, en appliquant la formule de Schwarz successivement pour chaque indice, on voit que les deux premiers termes de (**) sont bornés par

$$|A||\nabla^2 A||\nabla u|^2 + 4|A||\nabla A||\nabla u||\nabla^2 u|$$

et on obtient donc

$$\sum_{i,j,k,l} a_{ij}a_{kl}u_{ik}u_{jl} \leq c_1|\nabla u| + c_2|\nabla u|^2 + c_3|\nabla u||\nabla^2 u|$$

où,

$$c_1 = \beta|\nabla c||A|$$

$$c_2 = \beta|A||\nabla b| + |A||\nabla\tilde{b}| + \tfrac{1}{2}|A||\nabla^2 A|$$

$$c_3 = \beta|A||\tilde{b}| + |A||b| + 3|A||\nabla A|$$

Soit $\alpha = (\alpha_{ij})_{1\leq i,j\leq n}$ une matrice définie positive telle que

$$\alpha^t\alpha = A$$

c'est à dire que pour tout $1 \leq i, j \leq n$, on a

$$a_{ij} = \sum_p \alpha_{pi}\alpha_{pj}$$

Posons

$$U_{pk} = \sum_i \alpha_{pi}u_{ik}$$

On a alors,

$$\begin{aligned}\sum_{i,j,k,l} a_{ij}a_{kl}u_{ik}u_{jl} &= \sum_{klp} a_{kl}U_{pk}U_{pl}\\ &\geq \sum_p \theta|U_p|^2\\ &= \theta\sum_{p,k} U_{pk}^2\end{aligned}$$

Mais on a pour tout vecteur V de $\mathbb{R}^n$,

$$|\alpha V|^2 = \langle\alpha V, \alpha V\rangle = \langle V, AV\rangle \geq \theta|V|^2$$

d'où,

$$\theta\sum_{p,k} U_{pk}^2 \geq \theta^2\sum_{k,i} u_{ik}^2 = \theta^2|\nabla^2 u|^2$$

On a donc obtenu,

$$\begin{aligned}|\nabla^2 u|^2 &\leq \theta^{-2}[c_1|\nabla u| + c_2|\nabla u|^2 + c_3|\nabla u||\nabla^2 u|]\\ &\leq \theta^{-2}[c_1|\nabla u| + c_2|\nabla u|^2] + \frac{1}{2}\theta^{-4}c_3^2|\nabla u|^2 + \frac{1}{2}|\nabla^2 u|^2\end{aligned}$$

c'est à dire,

$$|\nabla^2 u|^2 \leq c_4|\nabla u| + c_5|\nabla u|^2$$

où,

$$c_4 = 2\theta^{-2}c_1$$

$$c_5 = 2\theta^{-2}c_2 + \theta^{-4}c_3^2$$

Cependant, puisque

$$\sum_{i,j} a_{ij}u_{ij} + \sum_{i,j} a_{ij}u_iu_j + \beta\sum_i b_iu_i + \sum_i \tilde{b}_iu_i = \beta c$$

on en déduit que

$$\begin{aligned}\Big(\sum_{i,j} a_{ij}u_iu_j\Big)^2 &\leq 4\Big[\beta^2c^2 + \beta^2\Big(\sum_i b_iu_i\Big)^2 + \Big(\sum_i \tilde{b}_iu_i\Big)^2 + \Big(\sum_{i,j} a_{ij}u_{ij}\Big)^2\Big]\\ &\leq 4[\beta^2|c|^2 + \beta^2|b|^2|\nabla u|^2 + |\tilde{b}|^2|\nabla u|^2 + |A|^2|\nabla^2 u|^2]\\ &\leq 4\beta^2|c|^2 + 4\beta^2|b|^2|\nabla u|^2 + 4|\tilde{b}|^2|\nabla u|^2 + 4|A|^2(c_4|\nabla u| + c_5|\nabla u|^2)\\ &\leq c_6 + c_7|\nabla u|^2\end{aligned}$$

où,

$$c_6 = 4\beta^2|c|^2 + 2|A|^2c_4$$

$$c_7 = 4\beta^2|b|^2 + 4|\tilde{b}|^2 + 2|A|^2c_4 + 4|A|^2c_5$$

Mais,

$$|\nabla u|^2 \leq \theta^{-1}\sum_{i,j} a_{ij}u_iu_j$$

ainsi, si on note $m = \sum_{i,j} a_{ij}u_iu_j$,

$$m^2 - (c_6 + c_7\theta^{-1}m) \leq 0$$

d'où on déduit

$$m \leq \frac{1}{2}\left(c_7\theta^{-1} + \sqrt{(c_7\theta^{-1})^2 + 4c_6^2}\right)$$

ce qui prouve le résultat annoncé, car les constantes c_6 et c_7 s'écrivent, respectivement, sous la forme $c'_6 + c''_6\beta^2$ et $c'_7 + c''_7\beta^2$ où c'_6, c''_6, c'_7 et c''_7 ne dépendent que de θ, $|\nabla^2 A|$, $|\nabla A|$, $|A|$, $|\nabla b|$, $|b|$, $|\nabla \tilde{b}|$, $|\tilde{b}|$, $|\nabla c|$ et de $|c|$. ∎

L'estimée de la proposition 3 s'obtient alors par localisation: On recouvre M par un nombre fini de cartes (O_k, ϕ_k) dont l'image est Ω. Puis on se place dans une carte (O_k, ϕ_k) telle que le maximum de $|\nabla \ln(\mu_\beta)|$ soit atteint dans O_k. Notons, dans cette carte, $(a_{ij}(x))_{1\leq i,j\leq N}$ l'inverse de la matrice définissant la métrique riemannienne et $(b^i(x))_{1\leq i\leq N}$ le champ de vecteur b. Posons également $u = \ln(\mu_\beta)$. L'équation

$$\frac{1}{2}\Delta\mu_\beta - \beta\,\mathrm{div}(\mu_\beta b) = 0$$

s'écrit aussi

$$\frac{1}{2}\Delta u + \frac{1}{2}|\nabla u|^2 - \beta\langle b, \nabla u\rangle - \beta\,\mathrm{div}(b) = 0$$

Ainsi, si on exprime ceci dans la carte O_k, on est ramené à la situation décrite par le lemme 5 (avec $n = N$, en gardant les mêmes a_{ij} et en prenant

$$b_i = -2b^i, \quad \tilde{b}_i = -D\sum_j\left(\frac{\partial}{\partial x_j}D^{-1}a_{ji}\right)$$

et $c = 2\,\mathrm{div}(b)$, où $D = \sqrt{\det((a_{ij})_{1\leq i,j\leq n})}$), ce qui prouve, en prenant la plus grande des constantes qui apparaissent pour chaque carte, qu'il existe une constante $K_1 > 0$ telle que pour tout $\beta \geq 1$,

$$\|\nabla\beta^{-1}\ln(\mu_\beta)\|_\infty \leq K_1$$

Remarquons que si on note D le diamètre de M, la proposition 3 permet de voir que

$$\forall\beta\geq 1, \quad \forall x\in M, \quad \exp(-K_1 D\beta) \leq \mu_\beta(x) \leq \exp(K_1 D\beta)$$

(inégalité de Harnack)

Nous allons maintenant démontrer la proposition 4.

Il faut tout d'abord voir que pour un $y \in M$ fixé, $\mu_\beta(y)$ est dérivable en β. Ceci se démontre grâce à la théorie de la perturbation asymptotique des opérateurs, développée dans le livre de Kato [8], qui permet de prouver que μ_β est dérivable par rapport à β, dans l'espace de Banach $C(M)$ (voir l'appendice 2). Cette théorie permet également d'obtenir une expression pour la dérivée, mais une fois admise l'existence de cette dernière, on va retrouver cette formule en utilisant une méthode probabiliste.

LEMME 6 *Pour tout $y \in M$ fixé, $\mu_\beta(y)$ est dérivable en β, pour $\beta > 0$, et*

$$\frac{\partial\mu_\beta}{\partial\beta}(y) = -\mu_\beta(y)\int_0^\infty E_y[(\langle b, \nabla\ln(\mu_\beta)\rangle + \mathrm{div}(b))(\tilde{Z}_t)]\,dt$$

où $(\tilde{Z}_t)_{t\geq 0}$ est une diffusion sur M, issue de y, et y, et de générateur

$$\frac{1}{2}\Delta\cdot - \langle\beta b - \nabla\ln(\mu_\beta), \nabla\cdot\rangle$$

(du moins sur $C^2(M)$).

Démonstration: Commençons par faire quelques rappels sur la formule de Girsanov.

Soit $(Z_t)_{t\geq 0}$ une diffusion sur M, homogène dans le temps, de générateur L_β (β étant ici un réel strictement positif fixé), et admettant comme loi initiale une certaine probabilité m sur M.

Soit $(B_t)_{t\geq 0}$ un mouvement brownien sur M (i.e. une diffusion de générateur $\frac{1}{2}\Delta$) admettant également m pour loi initiale. La formule de Girsanov permet d'exprimer la loi de $(Z_t)_{0\leq t\leq T}$ (notée $P_{m,L_\beta,T}$) à partir de celle de $(B_t)_{0\leq t\leq T}$.

Pour toute application borélienne positive F définie sur $C([0, T], M)$, on a

$$E_m[F((Z_t)_{0\leq t\leq T})] = E_m\left[F((B_t)_{0\leq t\leq T})\mathscr{E}\left(\beta\int_0^T\langle b(B_t), dB_t\rangle\right)\right]$$

(les deux termes valant éventuellement $+\infty$) où $\mathscr{E}(\beta\int_0^T\langle b(B_t), dB_t\rangle)$ est l'exponentielle stochastique de l'intégrale d'Itô $\beta\int_0^T\langle b(B_t), dB_t\rangle$,

$$\mathscr{E}\left(\beta\int_0^T\langle b(B_t), dB_t\rangle\right) = \exp\left(\beta\int_0^T\langle b(B_t), dB_t\rangle - \frac{1}{2}\beta^2\int_0^T|b(B_t)|^2\,dt\right)$$

(les m en indice de l'espérance rappellent que les processus considérés ont pour loi initiale m.)

La probabilité μ_β est stationnaire pour la diffusion $(Z_t)_{t\geq 0}$, ainsi pour tout $T \geq 0$ et toute fonction borélinenne bornée F sur M, on a

$$E_{\mu_\beta}[F(Z_T)] = \int\mu_\beta(dy)F(y)$$

mais le terme de gauche s'écrit aussi

$$\int\mu_\beta(dy)E_y\left[\mathscr{E}\left(\beta\int_0^T\langle b(B_t), dB_t\rangle\right)F(B_T)\right]$$

(on note $E_y[\cdot]$ à la place de $E_{\delta_y}[\cdot]$).

Ainsi, en dérivant les deux termes par rapport à β, ce qui est bien permis, on obtient

$$\begin{aligned}\int \lambda(dy)\frac{d\mu_\beta}{d\beta}(y)F(y) &= \int \lambda(dy)\frac{d\mu_\beta}{d\beta}(y)E_y\left[\mathscr{E}\left(\beta\int_0^T \langle b(B_t), dB_t\rangle\right)F(B_T)\right]\\ &+ \int \mu_\beta(dy)E_y\left[\left(\int_0^T \langle b(B_t), dB_t\rangle - \beta\int_0^T |b(B_t)|^2\,dt\right)\right.\\ &\left.\times\, \mathscr{E}\left(\beta\int_0^T \langle b(B_t), dB_t\rangle\right)F(B_T)\right]\end{aligned}$$

Notons (†) cette expression, le premier terme du membre de droite vaut

$$\int \lambda(dy)\frac{d\mu_\beta}{d\beta}(y)E_y[F(Z_T)]$$

Or le théorème ergodique (voir la proposition 7 suivante) affirme que $E_y[F(Z_T)]$ converge vers $\mu_\beta(F)$, quand T tend vers l'infini, et ceci uniformément en y. Ainsi,

$$\begin{aligned}\lim_{T\to+\infty}\int \lambda(dy)\frac{d\mu_\beta}{d\beta}(y)E_y[F(Z_T)] &= \mu_\beta(F)\int \lambda(dy)\frac{d\mu_\beta}{d\beta}(y)\\ &= 0\end{aligned}$$

Quant au second terme du membre de droite de (†), il s'écrit aussi

$$E_{\mu_\beta}\left[\left(\int_0^T \langle b(Z_t), dZ_t\rangle - \beta\int_0^T |b(Z_t)|^2\,dt\right)F(Z_T)\right]$$

Pour estimer cette intégrale, il vaut mieux retourner le temps. Posons, pour $0 \le t \le T$,

$$\tilde{Z}_t = Z_{T-t}$$

et notons que puisque μ_β est la probabilité stationnaire associée à L_β, $(\tilde{Z}_t)_{0\le t\le T}$ est, sous $P_{\mu_\beta, L_\beta, T}$, une diffusion de loi initiale μ_β et de générateur

$$\tilde{L}_\beta\cdot = \frac{1}{2}\Delta\cdot - \langle \beta b - \nabla\ln(\mu_\beta), \nabla\cdot\rangle$$

(voir par exemple [7], p. 280).

Pour effectuer facilement ce retournement du temps, il est commode de passer par l'intégrale stochastique de Stratonovitch, que l'on désignera par $\oint$. Notons que

$$\int_0^T \langle b(Z_t), dZ_t\rangle = \oint_0^T \langle b(Z_t), dZ_t\rangle - \frac{1}{2}\int_0^T \nabla b(dZ_t, dZ_t)$$

où ∇b est la forme bilinéaire définie en tout point x de M par $\nabla b(l, k) = \langle \nabla_l b, k\rangle$, pour tous les vecteurs l, $k \in T_x M$ (∇ désigne ici la dérivée covariante), la dernière intégrale ci-dessus étant alors la variation quadratique associée. On s'aperçoit (pour les notions et résultats de géométrie stochastique utilisés ici, on renvoit à [2]) qu'ici,

$$\int_0^T \nabla b(dZ_t, dZ_t) = \int_0^T \operatorname{div}(b)(Z_t)\,dt$$

Cependant, on sait bien comment se comporte l'intégrale de Stratonovitch par inversion du temps, on a

$$\oint_0^T \langle b(Z_t), dZ_t\rangle = -\oint_0^T \langle b(\tilde{Z}_t), d\tilde{Z}_t\rangle$$

(en effet, soit $0 = t_1 < t_2 < \cdots < t_n = T$ une partition de pas constant de $[0, T]$. Soit γ: $[0, T] \to M$ une courbe C^1 par morceaux, telle que pour tout $1 \le i \le n-1$, $[t_i, t_{i+1}] \ni t \mapsto \gamma(t)$ soit une des géodésiques minimisantes allant de Z_{t_i} à $Z_{t_{i+1}}$ (on a en fait unicité presque sûre des géodésiques précédentes, car le cut-locus associé à un point est de mesure nulle pour λ et donc pour μ_β, la loi des Z_t, qui lui est équivalente). Alors

$$\int_0^T \langle b(\gamma(t)), d\gamma(t)\rangle$$

converge dans L^2 vers

$$\oint_0^T \langle b(Z_t), dZ_t\rangle$$

quand le pas de la partition tend vers 0. Le résultat annoncé découle immédiatement de cette procédure d'approximation.)

Ainsi,

$$\begin{aligned}&E_{\mu_\beta}\left[\left(\int_0^T \langle b(Z_t), dZ_t\rangle - \beta\int_0^T |b(Z_t)|^2\,dt\right)F(Z_T)\right]\\ &= E_{\mu_\beta}\left[F(\tilde{Z}_0)\left(-\oint_0^T \langle b(\tilde{Z}_t), d\tilde{Z}_t\rangle - \frac{1}{2}\int_0^T \operatorname{div}(b)(\tilde{Z}_t)\,dt - \beta\int_0^T |b(\tilde{Z}_t)|^2\,dt\right)\right]\\ &= E_{\mu_\beta}\left[F(\tilde{Z}_0)\left(-\int_0^T \langle b(\tilde{Z}_t), d\tilde{Z}_t\rangle - \int_0^T \operatorname{div}(b)(\tilde{Z}_t)\,dt - \beta\int_0^T |b(\tilde{Z}_t)|^2\,dt\right)\right]\end{aligned}$$

On va chercher à aire disparaître l'intégrale stochastique

$$\int_0^T \langle b(\tilde{Z}_t), d\tilde{Z}_t\rangle$$

Heuristiquement, puisque $(\tilde{Z}_t)_{0 \le t \le T}$ est régie par

$$d\tilde{Z}_t = d\tilde{B}_t - (\beta b - \nabla \ln(\mu_\beta))(\tilde{Z}_t)\, dt$$

où $(\tilde{B}_t)_{0 \le t \le T}$ est un mouvement brownien sur M, on a envie d'écrire,

$$E_y\left[\int_0^T \langle b(\tilde{Z}_t), d\tilde{Z}_t\rangle\right] = -E_y\left[\int_0^T \langle b(\tilde{Z}_t), (\beta b - \nabla \ln(\mu_\beta))(\tilde{Z}_t)\rangle\, dt\right]$$

ces deux termes ayant bien un sens.

Cette expression est correcte, mais pour la justifier, il faut passer par la formule de Girsanov.

$$E_y\left[\int_0^T \langle b(\tilde{Z}_t), d\tilde{Z}_t\rangle\right] = E_y\left[\int_0^T \langle b(\tilde{B}_t), d\tilde{B}_t\rangle \mathscr{E}\left(\int_0^T \langle(-\beta b + \nabla \ln(\mu_\beta))(\tilde{B}_t), d\tilde{B}_t\rangle\right)\right]$$

Cependant,

$$t \mapsto M_t = \int_0^t \langle b(\tilde{B}_s), d\tilde{B}_s\rangle$$

et

$$t \mapsto N_t = \mathscr{E}\left(\int_0^t \langle(-\beta b + \nabla \ln(\mu_\beta))(\tilde{B}_s), d\tilde{B}_s\rangle\right)$$

sont deux martingales continues de crochet

$$\langle M, N\rangle_t = \int_0^t \left\langle b(\tilde{B}_s), (-\beta b + \nabla \ln(\mu_\beta))(\tilde{B}_s)\mathscr{E}\left(\int_0^s \langle(-\beta b + \nabla \ln(\mu_\beta))(\tilde{B}_u), d\tilde{B}_u\rangle\right)\right\rangle ds$$

ainsi,

$$\begin{aligned}
&E_y\left[\int_0^T \langle b(\tilde{B}_t), d\tilde{B}_t\rangle \mathscr{E}\left(\int_0^T \langle(-\beta b + \nabla \ln(\mu_\beta))(\tilde{B}_t), d\tilde{B}_t\rangle\right)\right] \\
&\quad = E_y\left[\int_0^T \left\langle b(\tilde{B}_t), (-\beta b + \nabla \ln(\mu_\beta))(\tilde{B}_t)\mathscr{E}\left(\int_0^t \langle(-\beta b + \nabla \ln(\mu_\beta))(\tilde{B}_s), d\tilde{B}_s\rangle\right)\right\rangle dt\right] \\
&\quad = \int_0^T E_y\left[\langle b(\tilde{B}_t), (-\beta b + \nabla \ln(\mu_\beta))(\tilde{B}_t)\rangle \mathscr{E}\left(\int_0^t \langle(-\beta b + \nabla \ln(\mu_\beta))(\tilde{B}_s), d\tilde{B}_s\rangle\right)\right] dt \\
&\quad = \int_0^T E_y[\langle b(\tilde{Z}_t), (-\beta b + \nabla \ln(\mu_\beta))(\tilde{Z}_t)\rangle]\, dt
\end{aligned}$$

En résumé, on a donc prouvé que

$$\begin{aligned}
&E_{\mu_\beta}\left[\left(\int_0^T \langle b(Z_t), dZ_t\rangle - \beta \int_0^T |b(Z_t)|^2\, dt\right)F(Z_T)\right] \\
&\quad = -\int \mu_\beta(dy) f(y) \int_0^T E_y[(\langle b, \nabla \ln(\mu_\beta)\rangle + \mathrm{div}(b))(\tilde{Z}_t)]\, dt
\end{aligned}$$

Cependant,

$$\int_0^T E_y[(\langle b, \nabla \ln(\mu_\beta)\rangle + \mathrm{div}(b))(\tilde{Z}_t)]\, dt$$

converge uniformément en y, quand T tend vers l'infini. En effet, d'après le théorème ergodique (voir plus loin la proposition 7)

$$E_y[(\langle b, \nabla \ln(\mu_\beta)\rangle + \mathrm{div}(b))(\tilde{Z}_t)]$$

converge exponentiellement vite (uniformément en y), quand t tend vers l'infini, vers

$$\mu_\beta(\langle b, \nabla \ln(\mu_\beta)\rangle + \mathrm{div}(b))$$

Or

$$\begin{aligned}
\mu_\beta(\langle b, \nabla \ln(\mu_\beta)\rangle + \mathrm{div}(b)) &= \int \langle b, \nabla \mu_\beta\rangle + \mu_\beta \,\mathrm{div}(b)\, d\lambda \\
&= 0
\end{aligned}$$

Finalement, en faisant tendre T vers l'infini dans (†), on a donc établi que pour toute fonction borélienne bornée F sur M,

$$\int \lambda(dy) \frac{d\mu_\beta}{d\beta}(y) F(y) = -\int \lambda(dy)\mu_\beta(y) F(y) \int_0^\infty E_y[(\langle b, \nabla \ln(\mu_\beta)\rangle + \mathrm{div}(b))(\tilde{Z}_t)]\, dt$$

ce qui prouve que λ-p.s. en y,

$$\frac{d\mu_\beta}{d\beta}(y) = -\mu_\beta(y) \int_0^\infty E_y[(\langle b, \nabla \ln(\mu_\beta)\rangle + \mathrm{div}(b))(\tilde{Z}_t)]\, dt$$

Mais les deux termes sont continus en y (en effet, rappelons que dans le terme de gauche, $d\mu_\beta/d\beta$ est la dérivée de μ_β en tant qu'élément de $C(M)$), et on a donc l'égalité pour tout y, ce qui termine la démonstration du lemme 6. ■

Remarque: Il existe une démonstration facile mais formelle du lemme 6.

Supposons en effet que l'on sache que $\partial\mu_\beta/\partial\beta$ appartient à $C^2(M)$ et que l'application

$$]0, \infty[\ni \beta \to \frac{\partial\mu_\beta}{\partial\beta} \in C^2(M)$$

soit continue.

Notons $U_\beta = \ln(\mu_\beta)$. Alors $U_\beta \in C^2(M)$ et l'application

$$]0, \infty[\ni \beta \to \frac{\partial U_\beta}{\partial\beta} \in C^2(M)$$

est continue.

Ainsi, on a

$$\frac{\partial}{\partial\beta}\Delta U_\beta = \Delta \frac{\partial}{\partial\beta} U_\beta$$

$$\frac{\partial}{\partial\beta}\nabla U_\beta = \nabla \frac{\partial}{\partial\beta} U_\beta$$

Cependant, l'équation satisfaite par U_β s'écrit,

$$\frac{\Delta}{2} U_\beta + \frac{1}{2}|\nabla U_\beta|^2 - \beta\langle b, \nabla U_\beta\rangle - \beta\,\mathrm{div}(b) = 0$$

on la dérive pour obtenir, grâce aux remarques précédentes,

$$\frac{\Delta}{2}\frac{\partial U_\beta}{\partial\beta} + \left\langle \nabla U_\beta - \beta b, \nabla \frac{\partial U_\beta}{\partial\beta}\right\rangle = \langle b, \nabla U_\beta\rangle + \mathrm{div}(b)$$

Soit $(\tilde{Z}_t)_{t\geq 0}$ est une diffusion sur M dont le générateur est donné sur $C^2(M)$ par

$$\frac{1}{2}\Delta\cdot - \langle \beta b - \nabla \ln(\mu_\beta), \nabla\cdot\rangle$$

et issue de $y \in M$.

Soit $T > 0$, on a

$$E_y\left[\frac{\partial U_\beta}{\partial\beta}(\tilde{Z}_T)\right] - \frac{\partial U_\beta}{\partial\beta}(y) = \int_0^T \left(\frac{\Delta}{2}\frac{\partial U_\beta}{\partial\beta} + \left\langle \nabla U_\beta - \beta b, \nabla\frac{\partial U_\beta}{\partial\beta}\right\rangle\right)(\tilde{Z}_t)\, dt$$
$$= \int_0^T (\langle b, \nabla U_\beta\rangle + \mathrm{div}(b))(\tilde{Z}_t)\, dt$$

Mais $E_y[\partial U_\beta/\partial\beta(\tilde{Z}_T)]$ tend vers 0 quand T tend vers l'infini, car μ_β est la probabilité invariante pour le générateur de la diffusion $(\tilde{Z}_t)_{t\geq 0}$, ainsi en faisant tendre T vers l'infini on obtient le résultat énoncé dans le lemme.

Reste donc à évaluer

$$\int_0^\infty E_y[(\langle b, \nabla \ln(\mu_\beta)\rangle + \mathrm{div}(b))(\tilde{Z}_t)]\, dt$$

Commençons par vérifier que dans le cas où b dérive d'un potentiel, i.e. $b = \nabla U$, où U est un élément de $C^\infty(M)$, on retrouve bien le résultat alors trivial

$$\frac{d\ln(\mu_\beta)}{d\beta} = 2\left(U - \int U\, d\mu_\beta\right)$$

(ainsi on peut prendre $\tilde{c} = 0$ dans la proposition 4 si on est dans la situation instantanément réversible.)

En effet, on a dans ce cas,

$$\langle b, \nabla \ln(\mu_\beta)\rangle + \mathrm{div}(b) = \frac{1}{2}2\Delta U + 2\beta\langle b, \nabla U\rangle$$
$$= 2\tilde{L}_\beta U (\overset{\text{ici}}{=} 2L_\beta U)$$

ainsi,

$$E_y[(\langle b, \nabla \ln(\mu_\beta)\rangle + \mathrm{div}(b))(\tilde{Z}_t)] = 2\frac{d}{dt}E_y[U(\tilde{Z}_t)]$$

d'où le résultat annoncé car

$$E_y[U(\tilde{Z}_0)] = U(y)$$

$$\lim_{t\to+\infty} E_y[U(\tilde{Z}_t)] = \mu_\beta(U)$$

Mais dans le cas général, il faut utiliser le théorème ergodique pour évaluer

$$\int_0^\infty E_y[(\langle b, \nabla \ln(\mu_\beta)\rangle + \mathrm{div}(b))(\tilde{Z}_t)]\, dt$$

et c'est par sa faute que va apparaître une constante c', en général non nulle, dans la proposition 4 (cependant le cas du cercle suggère que l'évaluation précédente n'est pas optimale).

Présentons maintenant le résultat ergodique qu'on va utiliser.

On considère, pour $\beta \geq 0$ fixé, la diffusion $(\tilde{Z}_t)_{t\geq 0}$ issue d'une probabilité initiale

m et de générateur $\tilde{L}_\beta \cdot = \frac{1}{2}\Delta \cdot + \langle b_\beta, \nabla \cdot \rangle$, où $b_\beta = -\beta b + \nabla \ln(\mu_\beta)$. Pour $t \geq 0$, on notera m_t la loi de $\tilde{Z}_t$ sur M.

On va étudier l'évolution de l'entropie $I(m_t)$ de m_t par rapport à μ_β, qui est aussi la probabilité invariante pour le générateur $\tilde{L}_\beta$. On a fait ici disparaître le t en indice de l'entropie car la diffusion considérée est homogène dans le temps, mais $I(\cdot)$ dépend évidemment de β.

Notons $c_\beta = c(\beta^{-1} \ln(\mu_\beta))$, l'application $c(\cdot)$, définie sur $C(M)$, étant présentée dans la section 2. Remarquons que l'inégalité de Harnack montre que c_β est uniformément borné pour $\beta \geq 1$. On a alors la

PROPOSITION 7 *Il existe une constante $K_3 > 0$ et un entier $p \geq 2$, tels que pour tout $\beta \geq 1$ et tout $t \geq 1$ (et toute probabilité initiale m), on ait*

$$I(m_t) \leq \exp(-K_3 \beta^{-p} \exp(-\beta c_\beta) t)$$

Démonstration: En utilisant le lemme 2, les propriétés de régularités des m_t pour $t > 0$, et la proposition 3, on voit qu'il existe une constante $K_4 > 0$ et un entier q tels que pour tout $\beta \geq 1$, $I(m_t)$ satisfasse pour $t > 0$ l'inégalité différentielle

$$\frac{d}{dt} I(m_t) \leq -K_4 \beta^{-q} \exp(-\beta c_\beta) I(m_t)$$

(pour plus de détails voir la démonstration du théorème 1 dans la section suivante).

Pour obtenir le résultat annoncé, il suffit donc de majorer convenablement $I(m_1)$. Rappelons que

$$I(m_1) = \int m_1 \ln(m_1)\, d\lambda - \int \ln(\mu_\beta)\, dm_1$$

Vu l'inégalité de Harnack, il existe une constante $K_5 \geq 0$ telle que pour tout $\beta \geq 0$, le dernier terme soit majoré par $K_5 \beta$. Intéressons-nous maintenant au premier terme du membre de droite.

D'après la formule de Girsanov, on a

$$m_1(x) = \int m(dy) E_y\left[\mathscr{E}\left(\int_0^1 \langle b_\beta(B_t), dB_t \rangle\right) \middle| B_1 = x\right] p_1(y, x)$$

où $(B_t)_{t \geq 0}$ est un mouvement brownien sur M et où $p_1(y, x)$ est le noyau de la chaleur associé.

Ainsi, par l'inégalité de Jensen, on a

$$\begin{aligned} m_1(x) \ln(m_1(x)) \leq & \int m(dy) E_y\left[\mathscr{E}\left(\int_0^1 \langle b_\beta(B_t), dB_t \rangle\right) \middle| B_1 = x\right] p_1(y, x) \\ & \times \ln\left(E_y\left[\mathscr{E}\left(\int_0^1 \langle b_\beta(B_t), dB_t \rangle\right) \middle| B_1 = x\right] p_1(y, x)\right) \end{aligned}$$

$$\begin{aligned} = & \int m(dy) E_y\left[\mathscr{E}\left(\int_0^1 \langle b_\beta(B_t), dB_t \rangle\right) \middle| B_1 = x\right] p_1(y, x) \\ & \times \ln\left(E_y\left[\mathscr{E}\left(\int_0^1 \langle b_\beta(B_t), dB_t \rangle\right) \middle| B_1 = x\right]\right) \\ & + \int m(dy) E_y\left[\mathscr{E}\left(\int_0^1 \langle b_\beta(B_t), dB_t \rangle\right) \middle| B_1 = x\right] p_1(y, x) \ln(p_1(y, x)) \end{aligned}$$

Cependant ce dernier terme est majoré par

$$\ln\left(\max_{y, z \in M} p_1(y, z)\right) m_1(x)$$

Considérons le premier terme. En appliquant l'inégalité de Jensen aux espérances conditionnelles, on a

$$\begin{aligned} & E_y\left[\mathscr{E}\left(\int_0^1 \langle b_\beta(B_t), dB_t \rangle\right) \middle| B_1 = x\right] \ln\left(E_y\left[\mathscr{E}\left(\int_0^1 \langle b_\beta(B_t), dB_t \rangle\right) \middle| B_1 = x\right]\right) \\ & \quad \leq E_y\left[\left(\int_0^1 \langle b_\beta(B_t), dB_t \rangle - \frac{1}{2}\int_0^1 |b_\beta|^2(B_t)\, dt\right) \mathscr{E}\left(\int_0^1 \langle b_\beta(B_t), dB_t \rangle\right) \middle| B_1 = x\right] \end{aligned}$$

ainsi,

$$\begin{aligned} \int m_1(x) \ln(m_1(x))\, d\lambda(x) \leq & E_m\left[\left(\int_0^1 \langle b_\beta(B_t), dB_t \rangle - \frac{1}{2}\int_0^1 |b_\beta|^2(B_t)\, dt\right) \right. \\ & \left. \times \mathscr{E}\left(\int_0^1 \langle b_\beta(B_t), dB_t \rangle\right)\right] + \ln(\|p_1\|_{C(M^2)}) \end{aligned}$$

Mais, en utilisant, comme précédemment, des propriétés de martingales, on voit que

$$E_m\left[\int_0^1 \langle b_\beta(B_t), dB_t \rangle \mathscr{E}\left(\int_0^1 \langle b_\beta(B_t), dB_t \rangle\right)\right] = E_m\left[\left(\int_0^1 |b_\beta|^2(B_t)\, dt\right) \mathscr{E}\left(\int_0^1 \langle b_\beta(B_t), dB_t \rangle\right)\right]$$

d'où,

$$\begin{aligned} \int m_1(x) \ln(m_1(x))\, d\lambda(x) \leq & E_m\left[\frac{1}{2}\left(\int_0^1 |b_\beta|^2(B_t)\, dt\right) \mathscr{E}\left(\int_0^1 \langle b_\beta(B_t), dB_t \rangle\right)\right] \\ & + \ln(\|p_1\|_{C(M^2)}) \\ = & E_m\left[\frac{1}{2}\int_0^1 |b_\beta|^2(\tilde{Z}_t)\, dt\right] + \ln(\|p_1\|_{C(M^2)}) \\ \leq & \beta^2 \| |b|^2 \|_\infty + \| |\nabla \ln(\mu_\beta)|^2 \|_\infty + \ln(\|p_1\|_{C(M^2)}) \end{aligned}$$

et d'après la proposition 3, il existe une constante $K_6 > 0$ telle que ceci soit majoré, pour $\beta \geq 1$, par $K_6 \beta^2$.

On en déduit que pour $\beta \geq 1$,

$$I(m_1) \leq (K_5 + K_6)\beta^2$$

Vu l'inégalité différentielle satisfaite par l'entropie, on peut bien trouver une constante $K_3 > 0$ telle que la proposition soit vérifiée. ■

Remarque: Le calcul précédent, pour obtenir une majoration de $I(m_1)$, est bien adapté si on considère des produits d'espaces, et permet en fait de montrer que l'énergie libre spécifique, associée à certaines diffusions sur des espaces du type $M^{\mathbb{Z}^d}$, est finie à tout instant strictement positif.

Pour appliquer la proposition 7, on va utiliser la relation suivante, rappelée avant le théorème 1, valable pour tout application f measurable et bornée sur M,

$$|m_t(f) - \mu_\beta(f)| \leq 4\sqrt{2}\|f\|_\infty \sqrt{I(m_t)}$$

En effet, en prenant $f = \langle b, \nabla \ln(\mu_\beta)\rangle + \operatorname{div}(b)$, on obtient, puisque $\mu_\beta(f) = 0$,

$$\forall t \geq 1, \quad |m_t(f)| \leq 4\sqrt{2}\|\langle b, \nabla \ln(\mu_\beta)\rangle + \operatorname{div}(b)\|_\infty \exp(-K_3\beta^{-p}\exp(-\beta c_\beta)t)$$

En utilisant la proposition 3, on s'aperçoit donc qu'il existe une constante $K_7 > 0$ et un entier $n \geq 2$, tels que pour tout $\beta \geq 1$,

$$\int_1^\infty E_y[(\langle b, \nabla \ln(\mu_\beta)\rangle + \operatorname{div}(b))(\bar{Z}_t)]\, dt \leq K_7 \beta^n \exp(\beta c_\beta)$$

cependant, il existe une constante $K_8 > 0$ telle que pour tout $\beta \geq 1$,

$$\int_0^1 E_y[(\langle b, \nabla \ln(\mu_\beta)\rangle + \operatorname{div}(b))(\bar{Z}_t)]\, dt \leq K_8 \beta$$

ce qui termine la démonstration de la proposition 4, en prenant

$$c' = \limsup_{\beta \to +\infty} c_\beta$$ ■

4 DÉMONSTRATION DU THÉORÈME 1

Une loi initiale m sur M et une évolution de l'inverse de la température $\beta \in C^1(\mathbb{R}_+; \mathbb{R}_+^*)$ étant donné, on considère la diffusion $(X_t)_{t \geq 0}$, non nécessairement homogène, qui leur est associée comme avant le théorème 1.

Commençons par faire quelques rappels sur la régularité des m_t. Grâce à la théorie des E.D.P. paraboliques (ou au calcul de Malliavin), on obtient les résultats suivants:

Pour tout $t > 0$, m_t est absolument continu par rapport à λ, et

$$\begin{aligned}]0, +\infty[\times M &\to \mathbb{R}_+ \\ (t, x) &\mapsto m_t(x) \end{aligned}$$

est une application de classe C^2 (de plus, pour tout $t > 0$ fixé,

$$\begin{aligned} M &\to \mathbb{R}_+ \\ x &\mapsto m_t(x) \end{aligned}$$

est de classe C^∞).

Ainsi, $m_t(x)$ satisfait au sens fort, sur $]0, +\infty[\times M$, l'équation de Fokker-Planck

$$\frac{\partial m_t}{\partial t}(x) = \frac{1}{2}\Delta m_t(x) - \beta_t\langle b, \nabla m_t\rangle(x) - \beta_t \operatorname{div}(b)(x) m_t(x)$$

D'autre part, une application de la formule de Girsanov nous permet de voir que tout $t > 0$, la densité m_t est strictement positive.

Ces résultats et l'inégalité de Harnack nous montrent que $I_t(m_t)$ est finie, pour tout $t > 0$ (mais on aurait pu le montrer en s'inspirant de la fin de la démonstration de la proposition 7), et plus précisément, que l'application $t \mapsto I_t(m_t)$ est continûment dérivable sur $]0, +\infty[$, de dérivée

$$\frac{dI_t(m_t)}{dt} = \int \frac{\partial m_t}{dt}\, d\lambda - \int \frac{\partial \ln(\mu_{\beta_t})}{\partial t}\, dm_t + \int \ln\left(\frac{m_t}{\mu_{\beta_t}}\right)\frac{\partial m_t}{\partial t}\, d\lambda$$

Mais le premier terme s'écrit aussi $(d/dt)\int m_t\, d\lambda$, expression qui est nulle. Quant au second, on peut le borner par

$$K_2 \left|\frac{d\beta_t}{dt}\right| \exp(\tilde{c}\beta_t)$$

où $\tilde{c}$ et K_2 sont comme dans la proposition 4.

Il nous reste à évaluer le dernier terme. Pour ceci, on considère l'expression

$$R_t = \frac{\partial m_t}{\partial t} - \frac{1}{2}\operatorname{div}\left(\mu_{\beta_t}\nabla \frac{m_t}{\mu_{\beta_t}}\right)$$

En général, ce terme est non nul (supposons la température constante, si ce terme est nul pour tout $t > 0$ et toute probabilité initiale m, alors b dérive d'un potentiel.

En effet, posons $\tilde b = \beta_0 b - \frac12 \nabla \ln(\mu_{\beta_0})$, alors

$$\begin{aligned} R_t = 0 &\Leftrightarrow \mathrm{div}(m_t \tilde b) = 0 \\ &\Leftrightarrow \forall \phi \in C^1(M), \quad \int \phi\, \mathrm{div}(m_t \tilde b)\, d\lambda = 0 \\ &\Leftrightarrow \forall \phi \in C^1(M), \quad \int \langle \nabla\phi, \tilde b\rangle\, dm_t = 0 \end{aligned}$$

On fait tendre t vers 0, pour obtenir, en prenant $m = \delta_x$,

$$\forall \phi \in C^1(M), \quad \forall x \in M, \quad \langle \nabla\phi, \tilde b\rangle(x) = 0$$

c'est à dire $\tilde b = 0$), mais le "miracle" est que

$$\int \ln\left(\frac{m_t}{\mu_{\beta_t}}\right) R_t\, d\lambda = 0$$

En effet, on a,

$$\frac{\partial m_t}{\partial t} = \frac12 \Delta m_t - \beta_t\, \mathrm{div}(m_t b)$$

$$\frac12 \mathrm{div}\left(\mu_{\beta_t} \nabla \frac{m_t}{\mu_{\beta_t}}\right) = \frac12 \Delta m_t - \frac12 \mathrm{div}(m_t \nabla \ln(\mu_{\beta_t}))$$

Ainsi,

$$\begin{aligned} R_t &= \mathrm{div}\left(m_t\left(\frac12 \nabla \ln(\mu_{\beta_t}) - \beta_t b\right)\right) \\ &= \left\langle \nabla m_t, \frac12 \nabla \ln(\mu_{\beta_t}) - \beta_t b \right\rangle + m_t\, \mathrm{div}\left(\frac12 \nabla \ln(\mu_{\beta_t}) - \beta_t b\right) \\ &= \left\langle \frac{1}{\mu_{\beta_t}} \nabla m_t, \frac12 \nabla \mu_{\beta_t} - \mu_{\beta_t}\beta_t b \right\rangle + m_t\, \mathrm{div}\left(\frac12 \nabla \ln(\mu_{\beta_t}) - \beta_t b\right) \end{aligned}$$

Cependant, du fait que μ_{β_t} est la probabilité stationnaire, elle satisfait

$$\mathrm{div}\left(\frac12 \nabla \mu_{\beta_t} - \mu_{\beta_t}\beta_t b\right) = 0$$

c'est à dire,

$$\mathrm{div}\left(\mu_{\beta_t}\left(\frac12 \nabla \ln(\mu_{\beta_t}) - \beta_t b\right)\right) = 0$$

d'où,

$$\mu_{\beta_t}\, \mathrm{div}\left(\frac12 \nabla \ln(\mu_{\beta_t}) - \beta_t b\right) = -\left\langle \nabla \mu_{\beta_t}, \frac12 \nabla \ln(\mu_{\beta_t}) - \beta_t b \right\rangle$$

On obtient donc

$$R_t = \left\langle \nabla \frac{m_t}{\mu_{\beta_t}}, \frac12 \nabla \mu_{\beta_t} - \mu_{\beta_t}\beta_t b \right\rangle$$

Or, si on note $F(x) = x\ln(x) - x$, pour $x > 0$, on a

$$F'(x) = \ln(x)$$

et

$$\begin{aligned} \int F'\left(\frac{m_t}{\mu_{\beta_t}}\right) R_t\, d\lambda &= \frac12 \int F'\left(\frac{m_t}{\mu_{\beta_t}}\right) \left\langle \nabla \frac{m_t}{\mu_{\beta_t}}, \frac12 \nabla \mu_{\beta_t} - \mu_{\beta_t}\beta_t b \right\rangle d\lambda \\ &= \frac12 \int \left\langle \nabla F\left(\frac{m_t}{\mu_{\beta_t}}\right), \frac12 \nabla \mu_{\beta_t} - \mu_{\beta_t}\beta_t b \right\rangle d\lambda \\ &= \frac12 \int F\left(\frac{m_t}{\mu_{\beta_t}}\right) \mathrm{div}\left(\frac12 \nabla \mu_{\beta_t} - \mu_{\beta_t}\beta_t b\right) d\lambda \\ &= 0 \end{aligned}$$

car $\mathrm{div}(\frac12 \nabla \mu_{\beta_t} - \mu_{\beta_t}\beta_t b) = 0$, d'où le résultat annoncé.
Ainsi,

$$\int \ln\left(\frac{m_t}{\mu_{\beta_t}}\right) \frac{\partial m_t}{\partial t}\, d\lambda = \frac12 \int \ln\left(\frac{m_t}{\mu_{\beta_t}}\right) \mathrm{div}\left(\mu_{\beta_t} \nabla \frac{m_t}{\mu_{\beta_t}}\right) d\lambda$$

cependant, on peut intégrer par parties le terme de droite pour voir qu'il vaut aussi

$$-2 \int \left| \nabla \sqrt{\frac{m_t}{\mu_{\beta_t}}} \right|^2 \mu_{\beta_t}\, d\lambda$$

En résumé on a donc prouvé que

$$\frac{dI_t(m_t)}{dt} \le K_2 \left|\frac{d\beta_t}{dt}\right| \exp(\tilde c \beta_t) - 2 \int \left| \nabla \sqrt{\frac{m_t}{\mu_{\beta_t}}} \right|^2 \mu_{\beta_t}\, d\lambda$$

Appliquons le lemme 2 sur les inégalités de Sobolev logarithmiques démontrées par Holley, Kusuoka et Stroock, avec $\beta = \beta_t$, $U = -W_{\beta_t}$ et $\phi = \sqrt{m_t/\mu_{\beta_t}}$, pour obtenir, en utilisant la proposition 3 (et l'inégalité de Harnack qui s'en déduit), qu'il existe une constante $A > 0$ telle que pour tout $t > 0$, on ait

$$I_t(m_t) \leq A(1+\beta_t)^{5N-1}\exp(\beta_t c_{\beta_t})\int\left|\nabla\sqrt{\frac{m_t}{\mu_{\beta_t}}}\right|^2 d\mu_{\beta_t}$$

(rappelons que $c_{\beta_t} = c(-W_{\beta_t})$).

Ainsi, on s'aperçoit que pour $t > 0$, $I_t(m_t)$ satisfait l'inégalité différentielle

$$\frac{dI_t(m_t)}{dt} \leq a_t - b_t I_t(m_t)$$

où,

$$a_t = K_2\left|\frac{d\beta_t}{dt}\right|\exp(\tilde{c}\beta_t)$$

$$b_t = 2A^{-1}(1+\beta_t)^{1-5N}\exp(-\beta_t c_{\beta_t})$$

Mais, d'après un lemme classique sur les inégalités différentielles, pour montrer que $\lim_{t\to+\infty} I_t(m_t) = 0$, il suffit de voir que

$$\begin{cases}\displaystyle\int^{\infty} b_t\,dt = +\infty \\ \displaystyle\lim_{t\to+\infty}\frac{a_t}{b_t} = 0\end{cases} \tag{C}$$

Prenons

$$\chi = c' + \limsup_{\beta\to+\infty} c(-W_\beta) = 2\limsup_{\beta\to+\infty} c(-W_\beta)$$

(où il apparait que pour obtenir la constante χ supposée optimale, il eut fallu prouver la proposition 4 avec $c' = 0$).

Supposons que pour t assez grand, on ait $\beta_t = K^{-1}\ln(t)$, avec $K > \chi$.

Soit $\varepsilon = K - \chi > 0$. On prend dans les calculs précédents $\tilde{c} = c' + \varepsilon/2$ et K_2 la constante qui lui est associée par la proposition 4.

Il est facile de vérifier que la condition (C) est alors bien satisfaite, ce qui termine la démonstration du théorème 1.

APPENDICE 1 REPRÉSENTATION DE LA PROBABILITÉ INVARIANTE SUR LE CERCLE

Il est connu que sur le cercle T (qu'on identifiera aussi avec $[0, 2\pi[$), il existe une formule explicite pour la probabilité invariante. En effet, il suffit de se placer sur $\mathbb{R}$ (en prolongeant par périodicité le champ b) et de trouver une solution périodique à

une équation linéaire du second degré, ce qui est ici toujours possible (cf. [10], dans cet article, Nevel'son étudie, sous certaines conditions sur le champ b, le comportement, quand la température tend vers 0, de la probabilité invariante sur T, ce qui nous intéresse directement, vu la remarque suivant le théorème 1).

Soit b: $T \to \mathbb{R}$ un champ de vecteurs de classe C^∞. On lui associe l'opérateur L agissant sur les fonctions $\phi \in C^2(T)$ par

$$L\phi = \frac{1}{2}\phi'' + b\phi'$$

et l'unique probabilité μ invariante pour cet opérateur.

Pour donner une expression explicite de μ en fonction de b, introduisons les notations suivantes.

Pour $x, y \in T$, soit

$$H(x, y) = 2\left(\int_{\gamma(y,x)}(-b(t))_+\,dt + \int_{\gamma(x,y)}(b(t))_+\,dt\right)$$

où, pour $x \neq y$, on a noté $\gamma(x, y)$ l'arc de cercle (identifié à un sous ensemble de $[0, 2\pi[$) obtenu en allant de x à y en tournant dans le sens trigonométrique, et où on a convenu de prendre $\gamma(x, x) = 0$.

Pour $x \in T$, on pose

$$v(x) = \int_T \exp(H(x, y))\lambda(dy)$$

où λ désigne la probabilité de Lebesgue sur T (T étant muni de sa structure riemannienne usuelle).

On a alors

PROPOSITION 8 *La probabilité invariante μ est absolument continue par rapport à λ et sa densité est donnée par*

$$\mu(x) = \left(\int v(y)\lambda(dy)\right)^{-1} v(x)$$

On peut également montrer cette proposition d'une manière plus probabiliste, qui fait peut être mieux comprendre la forme obtenue pour la probabilité invariante:

On approche L par les générateurs L_N de certains processus de sauts sur l'ensemble fini $T_N = \{(i/N)2\pi,\ 1 \leq i \leq N\}$, ne permettant que des sauts aux plus proches voisins. Soit μ_N la probabilité invariante associée à L_N, on peut voir facilement que les μ_N convergent étroitement vers μ. Cependant, il existe une description explicite de μ_N, donnée par Wentzell et Freidlin (cf. le lemme 3.1 p. 177 de [4]), qui est une formule combinatoire faisant intervenir certains graphes de T_N. Or, sur le cercle, on peut énumérer simplement ces graphes, ce qui permet d'étudier la limite des μ_N directement et d'obtenir la proposition.

On peut généraliser cette procédure d'approximation par triangularisation de la variété (du moins facilement sur T^n, si on ne veut pas d'ennuis avec la géométrie), et on se demande si on peut, sans obtenir évidemment de formule explicite pour la probabilité invariante, étudier directement le comportement des quantités apparaissant dans la proposition suivante. Mais dès $M = T^2$, la combinatoire semble exploser.

Soit $\beta \geq 0$. On considère maintenant l'opérateur L_β et la probabilité invariante μ_β associés au champ de vecteurs βb.

On va déduire de la proposition 8 les résultats suivants.

PROPOSITION 9 *L'application* $T \times]0, \infty[\ni (x, \beta) \mapsto \mu_\beta(x)$ *est de classe* C^∞. *De plus, pour tout* $(x, \beta) \in T \times]0, \infty[$, *on a*

$$\left|\frac{d}{d\beta}\ln(\mu_\beta(x))\right| \leq 4\pi\|b\|_\infty$$

$$\left|\frac{d}{dx}\beta^{-1}\ln(\mu_\beta(x))\right| \leq 4\|b\|_\infty$$

Démonstration: Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction C^∞ et définissons, pour $(x, \beta) \in T \times]0, \infty[$, l'application

$$F(x, \beta) = \int_T f(H_\beta(x, z))\lambda(dz)$$

(on note H_β l'application H sur $T \times T$ définie comme avant la proposition 8, mais où βb remplace b. Il est clair que pour $\beta \in \mathbb{R}$, on a $H_\beta = |\beta| H_1$.)

Vérifions que $F \in C^1(T \times]0, \infty[)$.

Soit $0 < h < 1$,

$$\begin{aligned} h^{-1}(F(x+h, \beta) - F(x, \beta)) &= h^{-1}\int_T f(H_\beta(x+h, z)) - f(H_\beta(x, z))\lambda(dz) \\ &= h^{-1}\int_{\gamma(x, x+h)} f(H_\beta(x+h, z)) - f(H_\beta(x, z))\lambda(dz) \\ &\quad + h^{-1}\int_{\gamma(x+h, x)} f(H_\beta(x+h, z)) - f(H_\beta(x, z))\lambda(dz) \end{aligned}$$

Cependant, quand $h > 0$ tend vers 0, uniformément en $z \in \gamma(x, x+h)$, $H_\beta(x, z)$ tend vers

$$H_\beta^+ = 2\beta\int_T (-b(t))_+ \, dt$$

et $H_\beta(x+h, z)$ tend vers

$$H_\beta^- = 2\beta\int_T (b(t))_+ \, dt$$

On en déduit que

$$\lim_{h \to 0_+} h^{-1}\int_{\gamma(x, x+h)} f(H_\beta(x+h, z)) - f(H_\beta(x, z))\lambda(dz) = \frac{1}{2\pi}[f(H_\beta^+) - f(H_\beta^-)]$$

D'autre part, pour $z \notin \gamma(x, x+h)$,

$$H_\beta(x+h, z) - H_\beta(x, z) = -2\beta\int_x^{x+h} b(t)\, dt$$

et il s'en suit que

$$\lim_{h \to 0_+} h^{-1}\int_{\gamma(x+h, x)} f(H_\beta(x+h, z)) - f(H_\beta(x, z))\lambda(dz) = -2\beta b(x)\int_T f'(H_\beta(x, z))\lambda(dz)$$

De la même manière on calcule la dérivée à gauche en x de $F(x, \beta)$ et on s'aperçoit qu'elle vaut la dérivée à droite. Ainsi F est dérivable en x de dérivée

$$\frac{d}{dx}F(x, \beta) = \frac{1}{2\pi}[f(H_\beta^+) - f(H_\beta^-)] - 2\beta b(x)\int_T f'(H_\beta(x, z))\lambda(dz)$$

expression qui est bien continue en (x, β).

D'autre part, F est clairement dérivable en β de dérivée

$$\frac{d}{d\beta}F(x, \beta) = \int_T H_1(x, z) f'((H_\beta(x, z))\lambda(dz)$$

qui est également continue en (x, β).

Ainsi, l'application F est bien de classe C^1 sur $T \times]0, \infty[$.

On en déduit que l'application

$$T \times]0, \infty[\ni (x, \beta) \mapsto \int_T \exp(H_\beta(x, z))\lambda(dz)$$

est de classe C^{k+1}, si b appartient à $C^k(T)$, d'où le premier résultat annoncé.

De plus, en appliquant le calcul précédent avec $f = \exp$, on obtient pour tout $x \in T$, $\beta > 0$,

$$\frac{d}{dx}\mu_\beta(x) = -2\beta b(x) Z_\beta^{-1} \int_T \exp(H_\beta(x, z))\lambda(dz) + \frac{1}{2\pi Z_\beta}[\exp(H_\beta^+) - \exp(H_\beta^-)]$$

où $Z_\beta = \int_{T \times T} \exp(H_\beta(y, z))\lambda(dz)\lambda(dy)$.

Mais remarquons que pour $z \neq x$,

$$\frac{\partial}{\partial z}\exp(H_\beta(x, z)) = 2\beta b(z)\exp(H_\beta(x, z))$$

ainsi,

$$\begin{aligned}\exp(H_\beta^+) - \exp(H_\beta^-) &= \int_x^{x+2\pi} 2\beta b(z)\exp(H_\beta(x, z))\,dz\\ &= 2\pi\int_T 2\beta b(z)\exp(H_\beta(x, z))\lambda(dz)\end{aligned}$$

On a donc obtenu

$$\frac{d}{dx}\mu_\beta(x) = 2\beta Z_\beta^{-1}\int_T (b(z) - b(x))\exp(H_\beta(x, z))\lambda(dz)$$

d'où,

$$\begin{aligned}\left|\frac{d}{dx}\ln(\mu_\beta(x))\right| &= 2\beta\mu_\beta^{-1}(x)Z_\beta^{-1}\left|\int_T (b(z) - b(x))\exp(H_\beta(x, z))\lambda(dz)\right|\\ &\leq 4\beta\|b\|_\infty\end{aligned}$$

D'autre part, on a

$$\begin{aligned}\frac{d}{d\beta}\mu_\beta(x) &= Z_\beta^{-1}\int_T H_1(x, z)\exp(H_\beta(x, z))\lambda(dz)\\ &\quad - \mu_\beta(x)Z_\beta^{-1}\int_{T\times T} H_1(y, z)\exp(H_\beta(y, z))\lambda(dz)\lambda(dy)\end{aligned}$$

or, pour tout $x, z \in T$,

$$0 \leq H_1(x, z) \leq 4\pi\|b\|_\infty$$

d'où,

$$\left|\frac{d}{d\beta}\ln(\mu_\beta(x))\right| \leq 4\pi\|b\|_\infty$$

ce qui termine la démonstration de la proposition. ■

On va s'intéresser au comportement de $\beta^{-1}\ln(\mu_\beta)$ quand β tend vers $+\infty$. Pour $x \in T$, posons

$$\tilde{W}(x) = \sup_{z\in T} H_1(x, z)$$

puis,

$$W(x) = \tilde{W}(x) - \sup_{y\in T}\tilde{W}(y)$$

Alors,

PROPOSITION 10 *$\beta^{-1}\ln(\mu_\beta)$ converge uniformément sur T vers W, quand β tend vers $+\infty$.*

Démonstration: Il est clair que pour $x \in T$ fixé,

$$\lim_{\beta\to+\infty}\beta^{-1}\ln\left(\int_T \exp(H_\beta(x, z))\lambda(dz)\right) = \tilde{W}(x)$$

et que

$$\begin{aligned}\lim_{\beta\to+\infty}\beta^{-1}\ln\left(\int_{T^2}\exp(H_\beta(x, z))\lambda(dz)\lambda(dx)\right) &= \sup_{x,z\in T} H_1(x, z)\\ &= \sup_{y\in T}\tilde{W}(y)\end{aligned}$$

Ainsi, $\beta^{-1}\ln(\mu_\beta)$ converge ponctuellement vers W. Cependant, d'après la proposition 9, les $\beta^{-1}\ln(\mu_\beta)$ satisfont, uniformément en $\beta > 0$, une condition de Lipschitz et sont également uniformément bornés par $4\pi\|b\|_\infty$. Le théorème d'Ascoli montre alors qu'ils forment un ensemble relativement compact dans $C(T)$ muni de la convergence uniforme. Or d'après le résultat précédent, le seul point limite possible, quand β tend vers l'infini est W, d'où la convergence annoncée. ■

Remarque La proposition 10 précise, dans le cas du cercle, le théorème 4.3 p. 190 de [4], qui affirme que les probabilités μ_β satisfont un principe de grandes déviations quand β tend vers $+\infty$. Wentzell et Freidlin donnent une autre formulation de W, qui est valable sur une variété riemannienne compacte quelconque, mais qui nécessite une condition supplémentaire (cf. l'hypothèse (A) donnée p. 169 de [4]).

APPENDICE 2 PERTURBATION ASYMPTOTIQUE D'OPÉRATEURS LINÉAIRES

On va utiliser des résultats de la théorie des opérateurs, présentés par Kato dans [8], pour donner une expression de

$$\frac{\partial}{\partial \beta} \ln(\mu_\beta), \quad \beta \in \mathbb{R}$$

Pour ceci, on cherche à interpréter μ_β comme un vecteur propre. Plaçons-nous, a priori, sur l'espace de Hilbert $L^2(\lambda)$ et intéressons-nous à l'opérateur L_β^*, l'opérateur adjoint de L_β (L_β est défini sur $C^2(M)$, mais on peut le considérer comme le prégénérateur d'un semi-groupe de diffusion sur $L^2(\lambda)$, et il admet alors une fermeture, qu'on note encore L_β), il est bien connu que L_β^* est de spectre discret et que 0 en est une valeur propre isolée, admettant pour espace propre la droite engendrée par μ_β. Notons que L_β^* s'exprime sur les fonctions $\phi \in C^2(M)$ par

$$L_\beta^* \phi = \frac{1}{2} \Delta\phi - \beta\langle b, \nabla\phi\rangle - \beta \operatorname{div}(b)\phi$$

Mais il existe une procédure classique pour ramener cet opérateur, de type Feynman-Kac, à un opérateur de diffusion:

Posons

$$\tilde{L}_\beta \cdot = \mu_\beta^{-1} \circ L_\beta^* \circ \mu_\beta \cdot$$

avec pour domaine $D(\tilde{L}_\beta) = \mu_\beta^{-1} D(L_\beta^*)$, où $D(L_\beta^*)$ est le domaine de L_β^* (μ_β est considéré ici comme un opérateur de multiplication dans $L^2(\lambda)$). On reconnait l'opérateur obtenu à partir de L_β par retournement du temps.

Sur $C^2(M)$, $\tilde{L}_\beta$ vaut

$$\tilde{L}_\beta \cdot = \frac{1}{2} \Delta \cdot - \langle \beta b - \nabla \ln(\mu_\beta), \nabla \cdot \rangle$$

Mais il est plus intéressant, pour appliquer les résultats présentés par Kato, de se placer sur l'espace de Banach $C(M)$ (muni de la norme sup $\|\ \|_\infty$). La restriction de $\tilde{L}_\beta$ à $C^2(M)$ peut s'interpréter comme le prégénérateur d'un semi-groupe de diffusion sur $C(M)$, et dans la suite, $\tilde{L}_\beta$ désignera la fermeture, dans $C(M)$, de cette restriction (avec le domaine associé $D(\tilde{L}_\beta)$). On se place résolument sur $C(M)$, et désormais, par définition, on prend

$$L_\beta^* \cdot = \mu_\beta \circ \tilde{L}_\beta \circ \mu_\beta^{-1} \cdot$$

avec pour domaine $D(L_\beta^*) = \mu_\beta D(\tilde{L}_\beta)$. Cet opérateur L_β^* (fermé sur $C(M)$) s'exprime toujours de la même façon sur $C^2(M)$.

Faisons quelques rappels sur la théorie des opérateurs. Pour $\zeta \in \mathbb{C}$ n'étant pas un élément du spectre de L_β^* (i.e. ici n'étant pas une valeur propre de L_β^* car ce dernier est de spectre discret), on peut définir le résolvent

$$R_\beta^*(\zeta) = (L_\beta^* + \zeta)^{-1}$$

(voir [8] p. 172).

Le résolvent réduit en la valeur propre isolée 0, S_β^*, s'exprime alors comme

$$S_\beta^* = \frac{1}{2\pi i} \int_\Gamma R_\beta^*(\zeta) \frac{1}{\zeta} \, d\zeta$$

où Γ est une courbe fermée entourant 0, mais aucun autre élément du spectre (cf. [8] formule 6.33 p. 180).

$\tilde{L}_\beta$ et L_β^* ont évidemment le même spectre. Le résolvent $\tilde{R}_\beta(\zeta)$ et le résolvent réduit en 0, $\tilde{S}_\beta$, associés à $\tilde{L}_\beta$, s'expriment en fonction de leur équivalent pour L_β^*;

$$\tilde{R}_\beta(\zeta) = \mu_\beta^{-1} \circ R_\beta^*(\zeta) \circ \mu_\beta$$
$$\tilde{S}_\beta(\zeta) = \mu_\beta^{-1} \circ S_\beta^*(\zeta) \circ \mu_\beta$$

Cependant, on sait exprimer, sur le demi-plan des nombres complexes de partie réelle strictement négative, le résolvent d'un semi-groupe de contractions. Soit $(\tilde{P}_t)_{t \geq 0}$ le semi-groupe de diffusion de générateur $\tilde{L}_\beta$. On a

$$\tilde{R}_\beta(\zeta) = -\int_0^\infty \exp(\zeta t)\tilde{P}_t \, dt$$

pour tous les nombres complexes ζ de partie réelle strictement négative.

Mais d'après la proposition 7, il est facile de voir qu'il existe $r > 0$, suffisamment petit, tel que cette formule reste vraie pour tout ζ tel que $|\zeta| < r$, si on l'applique à des éléments de $C(M)$ qui sont dans l'orthogonal de 1 (dans $L^2(\mu_\beta)$). En effet, pour de tels éléments, le théorème ergodique assure la convergence absolue du membre de droite.

D'autre part, pour $\zeta \neq 0$, $\tilde{R}_\beta(\zeta)$ s'exprime facilement sur les constantes, on a

$$\tilde{R}_\beta(\zeta)1 = \zeta^{-1}1$$

(mais attention, puisque $\tilde{P}_t 1 = 1$, on voit que la formule précédente n'est pas valable, pour tous les nombres complexes ζ de partie réelle positive, si on l'applique sur la fonction constante 1, car le membre de droite n'est pas convergent).

Présentons maintenant la perturbation que l'on va considérer.

Soit l'opérateur $F \cdot = -\langle b, \nabla \cdot \rangle - \operatorname{div}(b) \cdot$, avec pour domaine $D(F) = C^2(M)$.

Pour $h \in \mathbb{R}$ "petit", $L_{\beta+h}^*$ peut être vu comme une perturbation de L_β^*, au sens donné par Kato, p. 441 de [8],

$$L^*_{\beta+h} = L^*_\beta + hF$$

sur $C^2(M)$, et $C^2(M)$ est un core pour L^*_β.

Les résultats précédent permettent de voir que 0 est une valeur propre isolée de L^*_β qui est stable pour la perturbation (voir la définition présentée par Kato p. 437 et le théorème 1.3 p. 428 de [8]).

De plus, 0 est une valeur propre simple, on peut alors appliquer le théorème 2.6 p. 445 de [8], qui affirme que pour h positif et suffisamment petit, on peut choisir un vecteur propre $\tilde{\mu}_{\beta+h}$ associé à la valeur propre 0 de $L^*_{\beta+h}$, tel que

$$\tilde{\mu}_{\beta+h} = \mu_\beta - hS^*_\beta F\mu_\beta + \circ(h) \tag{‡}$$

où $\lim_{h\to 0_+} \|h^{-1}\circ(h)\|_\infty = 0$.

Ce résultat va nous permettre d'obtenir un développement en h de $\mu_{\beta+h}$.

Commençons par donner une expression plus parlante de $S^*_\beta F\mu_\beta$.

Prenons pour Γ, dans la définition de $\tilde{S}_\beta$, le cercle centré en 0 et de rayon $r/2$, et rappelons que

$$\tilde{S}_\beta = \frac{1}{2\pi i}\int_\Gamma \tilde{R}_\beta(\zeta)\,\frac{1}{\zeta}\,d\zeta$$

Il faut alors considérer les $0 \le h \le 1$, suffisamment petits, tels que $L^*_{\beta+h}$ ait une seule valeur propre à l'intérieur de Γ.

Sous ces hypothèses, on a, sur les éléments de $C(M)$ qui sont dans l'orthogonal de 1 (dans $L^2(\mu_\beta)$),

$$\begin{aligned}\tilde{S}_\beta &= -\frac{1}{2\pi i}\int_\Gamma\left(\int_0^\infty \exp(\zeta t)\tilde{P}_t\,dt\right)\frac{1}{\zeta}\,d\zeta\\ &= -\int_0^\infty\left(\frac{1}{2\pi i}\int_\Gamma \exp(\zeta t)\,\frac{1}{\zeta}\,d\zeta\right)\tilde{P}_t\,dt\\ &= -\int_0^\infty \tilde{P}_t\,dt\end{aligned}$$

Il faut évaluer

$$S^*_\beta\circ F(\mu_\beta) = \mu_\beta\circ\tilde{S}_\beta\circ\mu_\beta^{-1}\circ F(\mu_\beta)$$

Mais

$$\mu_\beta^{-1}\circ F(\mu_\beta) = -\operatorname{div}(b) - \langle b, \nabla\ln(\mu_\beta)\rangle$$

qui est bien dans l'orthogonal de 1 dans $L^2(\mu_\beta)$, d'où,

$$S^*_\beta\circ F(\mu_\beta) = \left(\int_0^\infty \tilde{P}_t(\operatorname{div}(b) + \langle b, \nabla\ln(\mu_\beta)\rangle)\,\mathrm{dt}\right)\mu_\beta$$

Ceci permet de voir que

$$\int S^*_\beta\circ F(\mu_\beta)\,d\lambda = 0$$

(en effet, μ_β est la probabilité stationnaire associée au semi-groupe $(\tilde{P}_t)_{t\ge 0}$, ainsi

$$\begin{aligned}\int S^*_\beta\circ F(\mu_\beta)\,d\lambda &= \int_0^\infty \mu_\beta(\operatorname{div}(b) + \langle b, \nabla\ln(\mu_\beta)\rangle)\,dt\\ &= 0).\end{aligned}$$

Montrons maintenant que $\mu_{\beta+h}$ est dérivable à droite en $h = 0$.

Puisque 0 est une valeur propre simple, il existe une constante $k_{\beta+h}\in\mathbb{R}$, telle que

$$\tilde{\mu}_{\beta+h} = k_{\beta+h}\mu_{\beta+h}$$

ainsi, en intégrant (‡) par rapport à $d\lambda$, on obtient

$$k_{\beta+h} = 1 + \circ(h)$$

($\circ(h)$ est ici un réel), ce qui permet de voir que $k_{\beta+h}$ est dérivable à droite en $h = 0$ et

$$\frac{d}{dh}k_{\beta+h}(0_+) = 0$$

d'où, dans $C(M)$,

$$\frac{d}{dh}\mu_{\beta+h}(0_+) = \frac{d}{dh}\tilde{\mu}_{\beta+h}(0_+)$$

Ainsi la dérivée à droite en β de μ_β vaut

$$-\left(\int_0^\infty \tilde{P}_t(\operatorname{div}(b) + \langle b, \nabla\ln(\mu_\beta)\rangle)\,dt\right)\mu_\beta$$

et comme on trouve la même expression pour la dérivée à gauche, on en déduit que

$$\frac{\partial}{\partial\beta}\ln(\mu_\beta) = -\int_0^\infty \tilde{P}_t(\operatorname{div}(b) + \langle b, \nabla\ln(\mu_\beta)\rangle)\,dt$$

expression obtenue par une méthode probabiliste dans la démonstration de la proposition 4.

Références

[1] M. Bardi and B. Perthame, Uniform estimates for some degenerating quasilinear elliptic equations and a bound on the Harnack constant for linear equations; *Asymptotic Analysis* **4** (1991), 1–16.

[2] M. Emery, *Stochastic Calculus in Manifolds*; Springer-Verlag (1989).

[3] L. C. Evans and H. Ishii, A PDE approach to some asymptotic problems concerning random differential equations with small noise intensities; *Ann. Inst. Henri Poincaré, Anal. Non Linéaire* **2** (1985), 1–20.

[4] M. I. Freidlin and A. D. Wentzell, *Random Perturbations of Dynamical Systems*; Springer-Verlag (1984).

[5] R. Holley, S. Kusuoka and D. Stroock, Asymptotics of the Spectral Gap with Applications to the Theory of Simulated Annealing; *J.F.A.* 83 (1989), 333–347.

[6] R. Holley and D. Stroock, Annealing via Sobolev Inequalities; *C.M.P.* 115 (1988), 553–569.

[7] N. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*; North Holland (1981).

[8] T. Kato, *Perturbation Theory for Linear Operators*; Springer-Verlag (2^{de} édition de 1980).

[9] L. Miclo, Thèse de doctorat de l'Université Paris 6, Laboratoire de probabilités (1991).

[10] M. B. Nevel'son, On the behavior of the invariant measure of a diffusion process with small diffusion on a circle; *Theory Proba. Appl.* **1** (1964), 125–131.