ANNALES DE L'I. H. P., SECTION B

LAURENT MICLO

Recuit simulé sur $\mathbb{R}^n$. Étude de l'évolution de l'énergie libre

Annales de l'I. H. P., section B, tome 28, nº 2 (1992), p. 235-266
<http://www.numdam.org/item?id=AIHPB_1992__28_2_235_0>

Ann. Inst. Henri Poincaré,

Vol. 28, n° 2, 1992, p. 235-266. *Probabilités et Statistiques*

Recuit simulé sur $\mathbb{R}^n$. Étude de l'évolution de l'énergie libre

par

Laurent MICLO

60, rue des Cigognes, 67540 Ostwald, France

RÉSUMÉ. — On s'intéresse aux algorithmes du recuit simulé sur $\mathbb{R}^n$, pour lesquels on propose une nouvelle démonstration de la convergence. On montre, en utilisant des inégalités de trou spectral, que l'énergie libre (*i. e.* le gain d'information de la loi du processus en un instant par rapport à la probabilité stationnaire en cet instant) satisfait une inégalité différentielle qui implique, pour certains taux de décroissance de la température vers zéro, la convergence du processus vers les minimums globaux du potentiel. Cet article est le développement d'une note publiée aux *C.R. Acad. Sci. Paris* sur ce sujet.

Mots clés : Recuit simulé, création d'entropie, trou spectral.

ABSTRACT. — We consider simulated annealing algorithms on $\mathbb{R}^n$, for which we propose a new demonstration of the convergence. We prove, by using spectral gap inequalities, that free energy (*i. e.* the relative entropy of the distribution of the process at one instant, with respect to the invariant probability at this moment) satisfies a differential inequality, which implies, under some decreasing evolutions of the temperature to zero, the convergence of the process to the global minima of the potential.

Classification A.M.S. : 60 J 60.

0. INTRODUCTION

On se propose de localiser les minimums globaux d'un certain potentiel U, défini sur $\mathbb{R}^n$, en utilisant une méthode probabiliste: l'algorithme du recuit simulé.

Considérons un processus de diffusion, dont le terme de dérive est le gradient du potentiel, et dont le coefficient de diffusion (qui représente la racine carrée de la «température») décroît vers 0, quand le temps tend vers l'infini. Sous certaines conditions, le système aura tendance à se concentrer au voisinage des minimums du potentiel. Cependant, si la décroissance de la tempéature est trop rapide, il pourra se geler en des minimums locaux. Il s'agit donc de déterminer les bons taux de refroidissement.

Nous montrerons qu'il existe un nombre $c>0$ tel que pour tout $k>c$, une décroissance de la température en $\frac{k}{\ln(t)}$ entraîne le système vers les minimums globaux. Il nous est possible d'estimer la vitesse de cette convergence, et dans certains cas (présentés dans l'appendice), nous obtenons des résultats plus précis, notamment que la convergence de l'algorithme du recuit simulé peut avoir lieu pour un refroidissement en $\frac{c}{\ln(t)-\mathrm{A}\ln(\ln(t))}$, pour une certaine constante $\mathrm{A}>0$.

Géométriquement, la constante c peut être décrite de la manière suivante:

Pour x, $y\in\mathbb{R}^n$, on pose $\mathscr{C}_{x,y}$ l'ensemble des chemins allant de x à y, c'est-à-dire l'ensemble des applications continues $\varphi:[0,1]\to\mathbb{R}^n$ telles que $\varphi(0)=x$ et $\varphi(1)=y$. Un chemin $\varphi\in\mathscr{C}_{x,y}$ étant donné, on définit son élévation par

$$e(\varphi)=\sup_{0\leqq t\leqq 1}\mathrm{U}(\varphi(t))-\mathrm{U}(x)-\mathrm{U}(y)+\mathrm{U}_0$$

où $\mathrm{U}_0=\min_{z\in\mathbb{R}^n}\mathrm{U}(z)$. La constante c est alors donnée par

$$c=\sup_{x,\,y\in\mathbb{R}^n}\left(\inf_{\varphi\in\mathscr{C}_{x,y}}e(\varphi)\right).$$

Elle est également liée au comportement asymptotique de la seconde valeur propre (ou trou spectral) de certains opérateurs (qui sont les générateurs infinitésimaux des diffusions considérées), et c'est cette dernière présentation que nous utiliserons.

Signalons que la convergence de l'algorithme du recuit simulé sur $\mathbb{R}^n$ avait déjà été démontrée, pour k assez grand, par T. S. Chiang, C. R. Hwang et S. J. Sheu [2], puis par G. Royer dans [11] pour $k>c$. Cependant la méthode présentée ici est totalement différente, et semble plus simple et plus précise. Nous étudierons l'évolution d'une quantité appelée « énergie libre », associée naturellement au système considéré, et qui se révélera être un outil très performant pour l'étude de l'algorithme du recuit simulé en général. On démontrera qu'elle satisfait une certaine inégalité différentielle, et tend vers 0 en l'infini, sous certaines hypothèses, ce qui entraîne la convergence du processus vers les minimums globaux du potentiel. Précisons que cette méthode permet également d'obtenir, quand la température est constante, des résultats quantitatifs sur l'ergodicité des processus considérés.

L'avantage de cette approche est d'être très générale et de pouvoir se transposer (se simplifier en fait) pour traiter le cas où l'espace des phases est un ensemble fini ou une variété riemannienne, compacte et connexe (*cf.* [10]), ce qui permet de retrouver des résultats connus (*cf.* [5] et [6]). Mais surtout, on peut la généraliser en dimension infinie, c'est le cas, par exemple, pour des diffusions invariantes par translation sur $M^{\mathbb{Z}^d}$ (M désigne une variété riemannienne, compacte, connexe et orientée), dont l'intensité du terme de diffusion tend vers 0, et dont la dérive est, formellement, le gradient de potentiels d'interactions (de rang fini, et invariants par translation) associés à certaines mesures de Gibbs (*cf.* [10]).

Notons que de nombreux travaux ont été faits sur l'algorithme du recuit simulé à temps discret sur un ensemble fini, souvent motivés par des problèmes concrets (optimisation combinatoire, restauration d'images...). Pour ce sujet on renvoit à la thèse d'Olivier Catoni [1] et aux références qu'elle contient.

REMERCIEMENTS

Je tiens, tout particulièrement, à remercier H. Doss, qui a bien voulu relire cet article et m'a suggéré certaines simplifications, pour son soutien et ses encouragements.

1. PRÉSENTATION DU RÉSULTAT PRINCIPAL

Commençons par préciser les hypothèses que nous ferons sur le potentiel U.

$$(\mathrm{H}_1)\qquad \begin{cases} \mathrm{U} \in \mathrm{C}^2(\mathbb{R}^n, \mathbb{R}) \\ \lim_{|x| \to \infty} \mathrm{U}(x) = \infty \\ \lim_{|x| \to \infty} |\nabla \mathrm{U}|(x) = \infty \\ |\nabla \mathrm{U}|^2 - \Delta \mathrm{U} \text{ est minoré.} \end{cases}$$

$$(H_2)\quad \begin{cases} U \in C^\infty(\mathbb{R}^n, \mathbb{R}) \\ \text{U et chacune de ses dérivées ont une croissance} \\ \qquad\qquad \text{au plus polynômiale en l'infini.} \\ \nabla U \text{ est globalement lipschitzien.} \end{cases}$$

Remarques. — (*a*) Typiquement, l'exemple suivant satisfait les deux hypothèses précédentes:

$$U \in C^\infty(\mathbb{R}^n, \mathbb{R})$$

et il existe $R > 0$, $C_1 > 0$ et $1 < \alpha \leqq 2$ tels que

$$|x| \geqq R \quad \Rightarrow \quad U(x) = C_1 |x|^\alpha$$

(où $|.|$ désigne la norme euclidienne usuelle sur $\mathbb{R}^n$).

(*b*) Les hypothèses (H_1) et (H_2) joueront des rôles bien distincts, et il doit être possible d'affaiblir considérablement (H_2) (voire de l'éliminer?). En effet, elle ne nous servira qu'à obtenir (par le calcul de Malliavin) des résultats de régularité de la loi, à un instant $t > 0$, de la diffusion considérée, et les résultats ainsi obtenus seront disproportionnés par rapport aux estimées dont nous aurons réellement besoin.

Introduisons le modèle stochastique que nous étudierons.

Soient $B = (B_t)_{t \geqq 0}$ un mouvement brownien à valeurs dans $\mathbb{R}^n$. défini sur un espace de probabilité $(\Omega, \mathscr{F}, P)$ et X_0 une variable aléatoire de Ω dans $\mathbb{R}^n$, de loi m et indépendante du processus B.

Soit σ une application de classe C^1 de $\mathbb{R}_+$ dans $\mathbb{R}_+^*$, qui représentera l'évolution de la température en fonction du temps.

On sait qu'il existe une unique diffusion $X = (X_t)_{t \geqq 0}$ à valeurs $\mathbb{R}^n$, régie par

$$(1) \qquad dX_t = \sqrt{\sigma(t)}\, dB_t - \frac{1}{2} \nabla U(X_t)\, dt$$

(X_0 étant l'état initial donné précédemment).

C'est cette diffusion X qui, sous certaines hypothèses, convergera en loi, quand t tend vers l'infini, vers les minimums globaux du potentiel U. Pour $t \geqq 0$, nous noterons m_t la loi de X_t.

On supposera, dans toute la suite, que $\sigma(t)$ tend vers 0 quand t tend vers l'infini. On fera également l'hypothèse que pour tout $t \geqq 0$, $\sigma(t) \in]0, 1/2]$, ce qui ne restreint pas la généralité du problème, quitte à multiplier U par une constante, comme on peut le voir par *scaling* du mouvement brownien et par changement de l'unité de temps.

Présentons une famille de probabilités qui jouera un rôle important.

Notons (*cf.* Jacquot [8]) que l'hypothèse (H_1) implique qu'il existe deux nombres A_1, $A_2 > 0$ tels que pour tout $x \in \mathbb{R}^n$, on ait

$$U(x) \geqq A_1 |x| - A_2$$

On définit alors, pour tout $\varepsilon>0$,

$$g_\varepsilon(.)=\frac{1}{\int \exp(-(1/\varepsilon)\,U(x))\,dx}\exp\left(-\frac{1}{\varepsilon}U(.)\right)$$

g_ε est la densité d'une mesure de probabilité (qui sera encore notée g_ε) sur $\mathbb{R}^n$. Remarquons que $g_{\sigma(t)}$ est la probabilité réversible «instantanée», au temps t, de la diffusion X.

Rappelons quelques propriétés des mesures g_ε:

Quand ε tend vers 0, g_ε tend à se concentrer au voisinage des minimums globaux de U, et en fait, on a même un principe de grandes déviations, avec pour fonctionnelle d'action $U-U_0$, où U_0 est le minimum de U.

Dans certains cas, on a aussi des résultats de convergence étroite des g_ε, quand ε tend vers 0 (*cf.* Hwang [7]):

Soit N l'ensemble des minimums globaux.

- si $\int_N dx>0$, où dx est la mesure de Lebesgue sur $\mathbb{R}^n$, alors

$$g_\varepsilon \overset{(e)}{\rightarrow} \frac{1}{\int_N dx} 1_N\,dx$$

- si $N=\{x_0\}$, alors

$$g_\varepsilon \overset{(e)}{\rightarrow} \delta_{x_0}$$

- si $N=\{x_1,\ldots,x_m\}$, alors

$$g_\varepsilon \overset{(e)}{\rightarrow} \frac{1}{\sum_{i=1}^{i=m}(\det U''(x_i))^{-1/2}}\sum_{i=1}^{i=m}(\det U''(x_i))^{-1/2}\,\delta_{x_i}$$

- Plus généralement, supposons que N soit une réunion finie de variétés de classe C^3, compactes, connexes et de dimensions strictement inférieures à n $\left[\text{on suppose seulement que U est de classe } C^3, \text{ satisfait } \lim_{|x|\to\infty} U(x)=\infty \text{ et } \int \exp(-U(x))\,dx<\infty\right]$. On note M la réunion des composantes de N de plus grande dimension, et $\det\left(\frac{\partial^2 U}{\partial t^2}(x)\right)$ le déterminant du hessien normal à N en $x\in N$. Il existe une mesure intrinsèque $\mathscr{M}$ sur M telle que si on suppose que pour tout $x\in N$, on ait

$\det\left(\frac{\partial^2 U}{\partial t^2}(x)\right)\neq 0$, alors

$$g_\varepsilon(dx) \overset{(e)}{\to} \frac{1}{\int_M \det((\partial^2 U/\partial t^2)(y))^{-1/2}\,\mathcal{M}(dy)} \det\left(\frac{\partial^2 U}{\partial t^2}(x)\right)^{-1/2} \mathcal{M}(dx)$$

On définit ensuite, pour $t \geqq 0$, l'entropie (aussi appelée l'information de Kullback et que l'on désignera comme «l'énergie libre») $I_t(p)$ d'une probabilité p sur $\mathbb{R}^n$, par rapport à $g_{\sigma(t)}$, en posant

$$I_t(p) = \begin{cases} \displaystyle\int \ln\left(\frac{p}{g_{\sigma(t)}}\right) dp, \\ \text{si } p \ll dx, \text{ en notant de la même façon } p \text{ et la densité } \dfrac{dp}{dx}, \\ +\infty, \quad \text{sinon.} \end{cases}$$

$I_t(p)$ mesure, d'une certaine manière, la distance entre p et $g_{\sigma(t)}$. En effet, on a

$$\| g_{\sigma(t)} - p \| \leqq 4\sqrt{2 I_t(p)}$$

où $\|.\|$ désigne la variation totale (*cf.* [5]).

Enfin, décrivons la constante c (pour l'équivalence de cette présentation et de celle donnée dans la section précédente, *voir* [8]).

Sur l'espace de Hilbert $L^2(\mathbb{R}^n, g_{\sigma(t)})$, on considére l'opérateur autoadjoint, positif $-L_t$, qui est l'extension de Friedrichs de l'opérateur symétrique $-\mathcal{L}_t$ défini sur l'espace des fonctions de classe C^2 à support compact par

$$\mathcal{L}_t \varphi = \frac{1}{2}[\sigma(t)\Delta\varphi - \langle \nabla U, \nabla \varphi \rangle]$$

L'hypothèse (H_1) assure que le spectre de L_t est discret. 0 est la première valeur propre et admet pour espace propre les fonctions constantes. Soit λ_t la seconde valeur propre, *i.e.*

$$\lambda_t = \frac{1}{2}\sigma(t) \inf_{\langle f^2 \rangle_t = 1} \frac{\int |\nabla f|^2 g_{\sigma(t)}(dx)}{\int (f - \langle f \rangle_t)^2 g_{\sigma(t)}(dx)}$$

où $\langle . \rangle_t$ désigne l'espérance par rapport à $g_{\sigma(t)}(dx)$.

Il est connu (*voir*, par exemple, [8]) que sous (H_1), la limite suivante existe et qu'elle est positive ou nulle

$$c = \lim_{\sigma(t)\to 0} -\sigma(t)\ln(\lambda_t)$$

Nous sommes maintenant en mesure d'énoncer le résultat principal de cet article.

THÉORÈME 1. — *Supposons que la loi initiale m ait des moments de tous ordres.*

Sous les hypothèses (H_1) *et* (H_2), *si pour tout t assez grand, on a*

$$\sigma(t)=\frac{k}{\ln(t)}, \qquad où \quad k>c$$

alors,

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

Ainsi, sous les hypothèses du théorème, les résultats décrits précédemment pour $g_{\sigma(t)}$, quand $\sigma(t)$ tend vers 0, seront également valables pour m_t, quand t tend vers l'infini [en effet $\lim_{t\to\infty} I_t(m_t)=0$ entraîne que les $g_{\sigma(t)}$ et les m_t ont mêmes points d'adhérence pour la topologie de la convergence étroite, quand t tend vers l'infini].

L'algorithme du recuit simulé est donc justifié dans ce cas.

2. QUELQUES PRÉLIMINAIRES SUR LES LOIS m_t POUR $t>0$

Dans cette section, nous allons rappeler quelques résultats sur la régularité des m_t, pour $t>0$, et nous démontrerons certaines majorations sur les moments de ces lois qui nous seront utiles par la suite.

Commençons par remarquer que l'hypothèse (H_2) est exactement celle dont nous avons besoin pour pouvoir appliquer le calcul de Malliavin tel qu'il est présenté par Kusuoka et Stroock dans [9] [*cf.* les exemples (3.14) qu'ils donnent], à la diffusion X. On obtient, si $m=\delta_x$ pour un certain $x\in\mathbb{R}^n$, que pour tout $t>0$, $(\delta_x)_t$ admet une densité de classe C^∞, par rapport à la mesure de Lebesgue. Cette densité possède des propriétés de régularité très fortes, ainsi, par exemple, pour tout multi-indice $\alpha\in\mathbb{N}^n$ et tout $T>1$, l'application

$$\begin{aligned}]T^{-1},T[\times\mathbb{R}^n &\to \mathbb{R} \\ (t,y) &\mapsto \frac{\partial^{|\alpha|}}{\partial y^\alpha}(\delta_x)_t(y) \end{aligned}$$

est continue et bornée par une expression du type

$$K(T,|\alpha|)(1+|x|)^{\gamma(T,|\alpha|)}$$

où $K(T,|\alpha|)$ et $\gamma(T,|\alpha|)$ sont deux constantes ne dépendant que de $T>1$ et $|\alpha|=\alpha_1+\ldots+\alpha_n$. (Pour la continuité, *voir* théorème (2.23) de [4].)

Ainsi, si m est une probablité sur $\mathbb{R}^n$ ayant des moments de tous ordres, on voit que m_t admet, pour $t>0$, la densité de classe C^∞

$$m_t(.)=\int m(dx)(\delta_x)_t(.)$$

par rapport à la mesure de Lebesgue, et si on pose, pour tout $T>1$,

$$\mathrm{Cte}(T)=K(T,0)\int(1+|x|)^{\gamma(T,0)}m(dx)$$

alors, pour tout $t\in[T^{-1},T]$,

$$\|m_t(.)\|_\infty\leqq \mathrm{Cte}(T)$$

D'autre part, on sait que m_t est une solution faible de l'E.D.P. parabolique sur $]0,\infty[\times\mathbb{R}^n$

$$\frac{\partial m_t}{\partial t}=\frac{\sigma(t)}{2}\sum_{k=1}^{k=n}\frac{\partial}{\partial x_k}\left(g_{\sigma(t)}\frac{\partial}{\partial x_k}\frac{m_t}{g_{\sigma(t)}}\right) \tag{2}$$

[équation de Fokker-Planck associée à la diffusion décrite par (1)].

Mais les résultats précédents nous permettent de voir qu'en fait, m_t en est une solution forte [*i.e.* que (2) est satisfaite au sens de la dérivation usuelle des fonctions]. On en déduit aussi que l'application

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

est de classe C^2, car σ est C^1.

Remarque (*importante*). — Pour démontrer le théorème 1, on oubliera les résultats présentés ci-dessus ainsi que l'hypothèse (H_2), et on s'efforcera de n'utiliser, outre (H_1) et le fait que m admet des moments de tous ordres $\left[i.e.\ \forall p\geqq 0,\ \int|U|^p m(dx)<+\infty\right]$, que les faits suivants

$$(H_3)\quad\left\{\begin{array}{l}\forall\, t>0,\quad m_t(dx)\text{ admet une densité de classe } C^2,\ m_t(.)\\ m_t(x)\text{, en tant que fonction de }(t,x)\text{, ainsi que ses dérivées jusqu'à}\\ \text{l'ordre 2 (en }x\text{), sont continues sur }]0,+\infty[\times\mathbb{R}^n.\\ \text{Pour tout } T>1\text{, il existe une constante Cte}(T)>0\text{ telle que}\\ \qquad\forall\, t\in[T^{-1}.T],\quad\forall x\in\mathbb{R}^n,\\ \qquad|m_t(x)|\leqq\mathrm{Cte}(T)\end{array}\right.$$

Remarquons, par exemple, que l'assertion de bornitude dans (H_3) peut être facilement montrée à partir de la formule de Girsanov, sous certaines conditions sur m, si on suppose, outre (H_1), que

$$\lim_{|x|\to\infty}\frac{U}{|\nabla U|^2}(x)=0$$

$$\lim_{|x| \to \infty} \frac{\Delta U}{|\nabla U|^2}(x)=0$$

hypothèses qui sont typiquement vérifiées dans le cas où pour $|x|$ assez grand, on a

$$U(x)=|x|^\alpha, \qquad \text{avec} \quad \alpha>2$$

Notons que (H_3) permet de voir que pour tout $x \in \mathbb{R}^n$ fixé, l'application

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

est continûment dérivable (et sa dérivée est donnée par l'équation de Fokker-Planck).

D'autre part, on pourrait également affaiblir l'hypothèse sur m, et ne supposer, par exemple, que

$$\int U^3(x)\, m(dx) < \infty$$

Nous allons maintenant prouver un résultat sur les moments des m_t, pour $t \geqq 0$. On a, en fait, des majorations plus précises, qui sont présentées dans l'appendice et qui sont utiles si on s'intéresse à des résultats plus fins que le théorème 1, notamment pour les taux de convergence, mais dont la démonstration est relativement technique et fastidieuse. La forme suivante m'a été suggérée par H. Doss.

LEMME 2. — *Notons* $U_0 = \min_{x \in \mathbb{R}^n} U(x)$.

Supposons que (H_1) *soit satisfaite et que m ait des moments de tous ordres.*

Alors, pour tout $p \geqq 0$ *et tout* $\varepsilon > 0$, *il existe une constante* $M_{p,\varepsilon} > 0$ *telle que pour tout* $t \geqq 0$, *on ait*

$$\int (U-U_0)^p(x)\, m_t(dx) \leqq M_{p,\varepsilon}(t+1)^\varepsilon$$

Notons que nous pouvons admettre, sans perte de généralité, que $U_0 = 0$. Nous supposerons cette convention faite dans toute la suite.

Démonstration. — Rappelons que nous avons déjà défini, pour $t \geqq 0$, l'opérateur

$$\mathscr{L}_t . = \frac{\sigma(t)}{2}\Delta . - \frac{1}{2}\langle \nabla U, \nabla . \rangle$$

Nous le supposerons ici défini sur l'espace des fonctions de classe C^2.

Soit $p \geqq 2$, alors U^p est de classe C^2, et

$$\mathscr{L}_t U^p = \frac{1}{2}(\sigma(t)[p(p-1)U^{p-2}\langle \nabla U, \nabla U\rangle + pU^{p-1}\Delta U] - pU^{p-1}\langle \nabla U, \nabla U\rangle)$$

Du fait que $|\nabla U|^2 - \Delta U$ est minoré et que $\lim_{|x|\to\infty} |\nabla U|(x) = \infty$, il existe un nombre $R_1 > 0$ tel que si $|x| \geqq R_1$, on ait

$$\frac{3}{2}|\nabla U|^2(x) - \Delta U(x) \geqq 0$$

D'autre part, soit $R_2 > 0$ tel que $|x| \geqq R_2$ implique

$$1 - \frac{p-1}{2U(x)} \geqq \frac{3}{4}$$

Posons $R = R_1 \vee R_2$. Pour $x \in \mathbb{R}^n$ tel que $|x| \geqq R$, on a

$$\begin{aligned}
2\mathscr{L}_t U^p(x) &\\
= \sigma(t)pU^{p-1}(x)\Delta U(x) &+ p\left(\sigma(t)\frac{p-1}{U(x)} - 1\right)U^{p-1}(x)|\nabla U|^2(x)\\
&\leqq \sigma(t)pU^{p-1}(x)\Delta U(x) - \frac{3}{4}pU^{p-1}(x)|\nabla U|^2(x)\\
&\leqq \sigma(t)p\left(\Delta U(x) - \frac{3}{2}|\nabla U|^2(x)\right)U^{p-1}(x) \leqq 0
\end{aligned}$$

[on a utilisé le fait que pour tout $t \geqq 0$, $\sigma(t) \leqq 1/2$].

Cependant, il existe clairement une constante $K > 0$ telle que

$$\forall x \in \mathbb{R}^n / |x| \leqq R,$$
$$[p(p-1)U^{p-2}\langle \nabla U, \nabla U\rangle + pU^{p-1}\Delta U](x) \leqq K$$

On a donc montré que,

$$\forall t \geqq 0, \quad \forall x \in \mathbb{R}^n,$$
$$\mathscr{L}_t U^p(x) \leqq \frac{1}{2}K\sigma(t) \leqq \frac{1}{4}K$$

Appliquons la formule d'Itô à la fonction U^p et à la diffusion décrite par (1).

Pour ceci, on introduit, pour $l \geqq 0$, le temps d'arrêt

$$T_l = \inf\{t \geqq 0 / U(X_t) \geqq l\}$$

On a alors, P-p. s., $\forall t \geqq 0$,

$$\mathrm{U}^p(\mathrm{X}_{t \wedge \mathrm{T}_l}) = \mathrm{U}^p(\mathrm{X}_0) + \int_0^{t \wedge \mathrm{T}_l} \mathscr{L}_s \mathrm{U}^p(\mathrm{X}_s)\,ds + p \int_0^{t \wedge \mathrm{T}_l} \mathrm{U}^{p-1}(\mathrm{X}_s) \langle \nabla \mathrm{U}(\mathrm{X}_s), d\mathrm{B}_s \rangle$$

Mais ce dernier terme est une « véritable » martingale [car $\mathrm{U}(x) \leqq l$ implique que x est dans un certain compact, et sur celui-ci $|\nabla \mathrm{U}|(x)$ est borné], ainsi en prenant l'espérance de l'égalité précédente, on obtient

$$\mathrm{E}[\mathrm{U}^p(\mathrm{X}_{t \wedge \mathrm{T}_l})] = \mathrm{E}[\mathrm{U}^p(\mathrm{X}_0)] + \mathrm{E}\left[\int_0^{t \wedge \mathrm{T}_l} \mathscr{L}_s \mathrm{U}^p(\mathrm{X}_s)\,ds\right] \leqq \int \mathrm{U}^p(x) m(dx) + \frac{\mathrm{K}}{4} t$$

qui est bien fini par hypothèse.

On fait alors tendre l vers l'infini, pour s'apercevoir, par convergence monotone, que

$$\mathrm{E}[\mathrm{U}^p(\mathrm{X}_t)] \leqq \int \mathrm{U}^p(x)\, m(dx) + \frac{\mathrm{K}}{4} t$$

Nous venons donc de démontrer que pour tout $p \geqq 2$, il existe une constante $\mathrm{K}_p > 0$ telle que

$$\forall t \geqq 0, \quad \int \mathrm{U}^p(x)\, m_t(dx) \leqq \mathrm{K}_p(1+t)$$

(d'où le résultat annoncé pour $\varepsilon \geqq 1$).

Soient $p \geqq 2$ et $1 > \varepsilon > 0$, on a alors, pour tout $t \geqq 0$,

$$\int \mathrm{U}^p(x)\, m_t(ds) \leqq \left(\int \mathrm{U}^{p/\varepsilon}(x)\, m_t(dx)\right)^\varepsilon \leqq (\mathrm{K}_{p/\varepsilon})^\varepsilon (1+t)^\varepsilon$$

Il suffit donc de poser $\mathrm{M}_{p,\varepsilon} = (\mathrm{K}_{p/\varepsilon})^\varepsilon$.

Le résultat pour $0 \leqq p < 2$ s'en déduit aisément. □

Remarque. — Les résultats de cette section nous montrent déjà que sous les conditions (H_1) et (H_3) et sous l'hypothèse que m admet des moments de tous ordres, on a pour tout $t > 0$,

$$\mathrm{I}_t(m_t) < \infty$$

$\Big[$en effet, d'après l'hypothèse (H_3) on a $\int \ln(m_t(x))\, m_t(dx) < \infty$ et d'après le lemme 2, on a $\int |\ln g_{\sigma(t)}(x)|\, m_t(dx) < \infty \Big]$.

Ils nous permettent également de voir que pour tout $p \geqq 0$ et tout $t>0$,

$$\int \left(\ln\left(\frac{m_t}{g_{\sigma(t)}}(x)+e \right) \right)^p m_t(dx) < \infty$$

et plus précisément, que cette expression est uniformément majorée, pour t dans les compacts de $]0, +\infty[$, résultat dont nous nous servirons ultérieurement.

3. ÉTUDE DE L'ÉVOLUTION DE L'ÉNERGIE LIBRE

En vue d'obtenir des inégalités différentielles satisfaites par $I_t(m_t)$, nous allons calculer sa dérivée temporelle. Formellement, ce calcul est immédiat.

PROPOSITION 3. — *Supposons que m ait des moments de tous ordres, et que les conditions* (H_1) *et* (H_3) *soient remplies. Alors l'application* $t \mapsto I_t(m_t)$ *est absolument continue sur* $]0, +\infty[$, *et presque sûrement pour* $t>0$,

$$\frac{dI_t(m_t)}{dt} = \frac{1}{\sigma(t)^2} \frac{d\sigma(t)}{dt} \int U(g_{\sigma(t)} - m_t)\, dx - 2\sigma(t) \int \left| \nabla \sqrt{\frac{m_t}{g_{\sigma(t)}}} \right|^2 g_{\sigma(t)}\, dx$$

Démonstration. — Soient $\varepsilon>0$ et $\rho \in C_c^\infty(\mathbb{R}^n)$ (l'espace des fonctions de classe C^∞ et à support compact) positive.

On commence par étudier l'évolution de la quantité suivante, définie pour $t>0$,

$$H_t^\varepsilon = \int \rho\, m_t \ln\left(\frac{m_t}{g_{\sigma(t)}} + \varepsilon \right) dx$$

L'intérêt d'introduire H_t^ε est purement technique; notamment, cela nous permet de dériver sous l'intégrale, sans nous préoccuper des points où, éventuellement, m_t s'annule, et d'intégrer par parties.

Il apparaît, d'après la condition (H_3) (*cf.* les remarques qui la suivent), que l'application $t \mapsto H_t^\varepsilon$ est continûment dérivable sur $]0, +\infty[$, et pour tout $t>0$, sa dérivée est donnée par

$$\begin{aligned}\frac{dH_t^\varepsilon}{dt} = \int \rho \frac{\partial m_t}{\partial t} \ln\left(\frac{m_t}{g_{\sigma(t)}} + \varepsilon \right) dx \\ + \int \rho \left(\frac{m_t}{g_{\sigma(t)}} + \varepsilon \right)^{-1} \frac{m_t}{g_\sigma(t)} \frac{\partial m_t}{\partial t}\, dx \\ - \int \rho \left(\frac{m_t}{g_{\sigma(t)}} + \varepsilon \right)^{-1} \left(\frac{m_t}{g_{\sigma(t)}} \right)^2 \frac{\partial g_{\sigma(t)}}{\partial t}\, dx\end{aligned}$$

Intéressons-nous au premier terme du membre de droite. On remplace $\dfrac{\partial m_t}{\partial t}$ par l'expression donnée par l'équation de Fokker-Planck, puis on intègre par parties, pour voir qu'il vaut

$$-\frac{\sigma(t)}{2}\int\sum_{k=1}^{k=n}\frac{\partial}{\partial x_k}\left(\rho\ln\left(\frac{m_t}{g_{\sigma(t)}}+\varepsilon\right)\right)g_{\sigma(t)}\frac{\partial}{\partial x_k}\frac{m_t}{g_{\sigma(t)}}\,dx$$
$$=-\frac{\sigma(t)}{2}\int\left(\frac{m_t}{g_{\sigma(t)}}+\varepsilon\right)^{-1}\rho\, g_{\sigma(t)}\sum_{k=1}^{k=n}\left(\frac{\partial}{\partial x_k}\frac{m_t}{g_{\sigma(t)}}\right)^2 dx$$
$$-\frac{\sigma(t)}{2}\int\sum_{k=1}^{k=n}\ln\left(\frac{m_t}{g_{\sigma(t)}}+\varepsilon\right)\frac{\partial\rho}{\partial x_k}\frac{\partial}{\partial x_k}\left(\frac{m_t}{g_{\sigma(t)}}\right)g_{\sigma(t)}\,dx$$

Soient $t>s>0$ fixés.

On intègre l'expression obtenue pour $\dfrac{d\mathrm{H}_u^\varepsilon}{du}$, u variant entre s et t, pour trouver que

$$\mathrm{H}_t^\varepsilon-\mathrm{H}_s^\varepsilon=-\int_s^t du\frac{\sigma(u)}{2}\int\rho\left(\frac{m_u}{g_{\sigma(u)}}+\varepsilon\right)^{-1}\left|\nabla\frac{m_u}{g_{\sigma(u)}}\right|^2 dg_{\sigma(u)}$$
$$-\int_s^t du\frac{\sigma(u)}{2}\int\ln\left(\frac{m_u}{g_{\sigma(u)}}+\varepsilon\right)\sum_{k=1}^{k=n}\frac{\partial\rho}{\partial x_k}\frac{\partial}{\partial x_k}\left(\frac{m_u}{g_{\sigma(u)}}\right)dg_{\sigma(u)}$$
$$+\int_s^t du\int\rho\left(\frac{m_u}{g_{\sigma(u)}}+\varepsilon\right)^{-1}\frac{m_u}{g_{\sigma(u)}}\frac{\partial m_u}{\partial u}dx$$
$$-\int_s^t du\int\rho\left(\frac{m_u}{g_{\sigma(u)}}+\varepsilon\right)^{-1}\left(\frac{m_u}{g_{\sigma(u)}}\right)^2\frac{\partial g_{\sigma(u)}}{\partial u}dx$$

On désignera par (*) le membre de droite. Nous allons étudier le comportement de chaque terme de (*), quand ε tend vers 0. Pour ceci, commençons par rappeler le résultat suivant (*voir*, par exemple, le lemme 3.23 de [12]) :

Notons θ le support (compact) de ρ. Alors

$$(\dagger)\qquad\left\{\begin{array}{l}\forall u>0,\quad\forall 1\leq k\leq n,\quad\forall x\in\theta,\\ \left(\dfrac{\partial}{\partial x_k}\dfrac{m_u}{g_{\sigma(u)}}(x)\right)^2\leq 2\left\|\dfrac{m_u}{g_{\sigma(u)}}\right\|_{C^2(\theta)}\dfrac{m_u}{g_{\sigma(u)}}(x)\end{array}\right.$$

[cette majoration découle facilement du fait que l'application $x\mapsto\dfrac{m_u}{g_{\sigma(u)}}(x)$ est de classe C^2, et est positive].

Posons

$$K=2\sup_{s\leqq u\leqq t}\left\|\frac{m_u}{g_{\sigma(u)}}\right\|_{C^2(\theta)}$$

qui est une quantité finie, vu les hypothèses de régularité faites sur m_u dans (H_3).

Ainsi

$$\forall s\leqq u\leqq t,\quad \forall \varepsilon>0,$$
$$\int\rho\left(\frac{m_u}{g_{\sigma(u)}}+\varepsilon\right)^{-1}\frac{m_u}{g_{\sigma(u)}}\left(\frac{m_u}{g_{\sigma(u)}}\right)^{-1}\left|\nabla\frac{m_u}{g_{\sigma(u)}}\right|^2 dg_{\sigma(u)}\leqq K\int\rho\, dg_{\sigma(u)}$$

Il en découle, par convergence dominée, que le premier terme de $(*)$ tend, quand ε tend vers 0, vers

$$-\int_s^t d\mu\frac{\sigma(u)}{2}\int\rho\, 1_{\{m_u/g_{\sigma(u)}>0\}}\left(\frac{m_u}{g_{\sigma(u)}}\right)^{-1}\left|\nabla\frac{m_u}{g_{\sigma(u)}}\right|^2 dg_{\sigma(u)}$$
$$=-\int_s^t du\,2\,\sigma(u)\int\rho\left|\nabla\sqrt{\frac{m_u}{g_{\sigma(u)}}}\right|^2 dg_{\sigma(u)}$$

avec la convention (que nous supposerons faite dans toute la suite) que $\left|\nabla\sqrt{\frac{m_u}{g_{\sigma(u)}}}\right|(x)=0$, si x est un point en lequel m_u s'annule.

D'autre part, du fait que pour tout $x\in\theta$ et tout $s\leqq u\leqq t$, on ait

$$\left|\ln\left(\frac{m_u}{g_{\sigma(u)}}+\varepsilon\right)\frac{\partial}{\partial x_k}\frac{m_u}{g_{\sigma(u)}}\right|(x)\leqq\sqrt{K}\left|\ln\left(\frac{m_u}{g_{\sigma(u)}}+\varepsilon\right)\sqrt{\frac{m_u}{g_{\sigma(u)}}}\right|(x)$$

le second terme de $(*)$ tend (par convergence dominée), quand ε tend vers 0, vers

$$-\int_s^t du\frac{\sigma(u)}{2}\int\ln\left(\frac{m_u}{g_{\sigma(u)}}\right)\sum_{k=1}^{k=n}\frac{\partial\rho}{\partial x_k}\frac{\partial}{\partial x_k}\left(\frac{m_u}{g_{\sigma(u)}}\right)dg_{\sigma(u)}$$

[avec la convention que $\ln\left(\frac{m_u}{g_{\sigma(u)}}\right)\frac{\partial}{\partial x_k}\frac{m_u}{g_{\sigma(u)}}(x)=0$, si x est un point qui annule m_u].

Quant aux troisième et quatrième termes de $(*)$, ils tendent, quand ε tend vers 0, respectivement (toujours par convergence dominée, grâce à la présence de ρ), vers

$$\int_s^t du\int\rho\, 1_{\{m_u>0\}}\frac{\partial m_u}{\partial u}dx$$

et

$$-\int_s^t du \int \rho \frac{m_u}{g_{\sigma(u)}} \frac{\partial g_{\sigma(u)}}{\partial u} dx$$

Mais, du fait que

$$m_u(x)=0 \quad \Rightarrow \quad \frac{\partial m_u}{\partial u}(x)=0$$

le premier de ces termes s'écrit aussi

$$\int_s^t du \int \rho \frac{\partial m_u}{\partial u} dx = \int_s^t du \frac{d}{du} \int \rho\, m_u\, dx = \left[\int \rho\, m_u\, dx \right]_s^t$$

D'autre part, remarquons que pour tout $t>0$,

$$\begin{aligned}\lim_{\varepsilon \to 0} \mathrm{H}_t^\varepsilon &= \int \rho \frac{m_t}{g_{\sigma(t)}} \ln\left(\frac{m_t}{g_{\sigma(t)}}\right) dg_{\sigma(t)} \\ &= \int \rho\, m_t \ln\left(\frac{m_t}{g_{\sigma(t)}}\right) dx\end{aligned}$$

Notons H_t cette dernière quantité.

En regroupant les résultats précédents, on a donc prouvé que

$$(**) \quad \begin{aligned}\mathrm{H}_t - \mathrm{H}_s = &-2 \int_s^t du\, \sigma(u) \int \rho\, g_{\sigma(u)} \left| \nabla \sqrt{\frac{m_u}{g_{\sigma(u)}}} \right|^2 dx \\ &- \int_s^t du \frac{\sigma(u)}{2} \int g_{\sigma(u)} \ln\left(\frac{m_u}{g_{\sigma(u)}}\right) \left\langle \nabla \rho, \nabla \frac{m_u}{g_{\sigma(u)}} \right\rangle dx \\ &+ \int \rho\, m_t\, dx - \int \rho\, m_s\, dx \\ &+ \int_s^t du \left(\frac{d}{du} \frac{1}{\sigma(u)} \right) \int \rho\, m_u [\mathrm{U} - \langle \mathrm{U} \rangle_t]\, dx\end{aligned}$$

Nous allons choisir des ρ particuliers, que nous ferons tendre vers 1.

Soit $f \in \mathrm{C}^\infty(\mathbb{R}^n)$, décroissante et satisfaisant

$$f(x) = \begin{cases} 1, & \text{si } x \leqq 1 \\ 0, & \text{si } x \geqq 2. \end{cases}$$

Soit $g \in \mathrm{C}^\infty(\mathbb{R}^n)$, convexe et telle que pour $|x| \geqq 1$, on ait $g(x)=|x|$.

On définit, pour $p \in \mathbb{N}$ et $x \in \mathbb{R}^n$,

$$\rho_p(x) = f^2\left(\sum_{k=1}^{k=n} g(x_k) - p\right)$$

Alors, ρ_p est une fonction de classe C^∞ à support compact, et il existe une constante $M_1>0$ telle que

$$\forall p\in\mathbb{N},\quad \forall 1\leqq k\leqq n,\quad \forall x\in\mathbb{R}^n,$$
$$|\rho_p(x)|\vee\left|\frac{\partial}{\partial x_k}\sqrt{\rho_p}(x)\right|\leqq M_1$$

On a également que

$$\lim_{p\to\infty}\rho_p=1$$

et pour tout $1\leqq k\leqq n$,

$$\lim_{p\to\infty}\frac{\partial\sqrt{\rho_p}}{\partial x_k}=0$$

On considère (✱✱) avec $\rho=\rho_p$.

Notons qu'il existe une constante $K>0$ (dépendant de t et de s) telle que pour tout $p\in\mathbb{N}$,

$$H_t-H_s-\int\rho_p(m_t-m_s)\,dx-\int_s^t du\frac{d}{du}\left(\frac{1}{\sigma(u)}\right)\int\rho_p m_u(U-\langle U\rangle_t)\,dx\geqq -K$$

Il s'ensuit que

$$\begin{aligned}
-K\leqq &-2\int_s^t du\,\sigma(u)\int g_{\sigma(u)}\rho_p\left|\nabla\sqrt{\frac{m_u}{g_{\sigma(u)}}}\right|^2dx\\
&-2\int_s^t du\,\sigma(u)\int g_{\sigma(u)}\sqrt{\rho_p}\ln\left(\frac{m_u}{g_{\sigma(u)}}\right)\\
&\times\sqrt{\frac{m_u}{g_{\sigma(u)}}}\left\langle\nabla\sqrt{\rho_p},\frac{1}{2}\left(\sqrt{\frac{m_u}{g_{\sigma(u)}}}\right)^{-1}\nabla\frac{m_u}{g_{\sigma(u)}}\right\rangle dx\\
&\leqq -2\int_s^t du\,\sigma(u)\int g_{\sigma(u)}\rho_p\left|\nabla\sqrt{\frac{m_u}{g_{\sigma(u)}}}\right|^2dx\\
&+2\int_s^t du\,\sigma(u)\sqrt{\int g_{\sigma(u)}\rho_p\left|\nabla\sqrt{\frac{m_u}{g_{\sigma(u)}}}\right|^2dx}\\
&\times\sqrt{\int m_u\left(\ln\frac{m_u}{g_{\sigma(u)}}\right)^2|\nabla\sqrt{\rho_p}|^2\,dx}
\end{aligned}$$

Cependant, on a $|\nabla\sqrt{\rho_p}|^2\leqq nM_1^2$, et, puisque m admet des moments de tous ordres, il existe, d'après la section 2, une constante $M_2>0$ (dépendant de s et de t), telle que

$$\forall u/s\leqq u\leqq t,$$

$$\int m_u \left(\ln \frac{m_u}{g_{\sigma(u)}}\right)^2 dx \leqq M_2$$

Ainsi, si on note

$$J_p(u) = \int g_{\sigma(u)} \rho_p \left| \nabla \sqrt{\frac{m_u}{g_{\sigma(u)}}} \right|^2 dx$$

on a,

$$-K \leqq -2 \int_s^t du\, \sigma(u) (J_p(u) - \sqrt{n M_1^2 M_2} \sqrt{J_p(u)})$$

or,

$$\sqrt{n M_1^2 M_2} \sqrt{J_p(u)} \leqq \frac{n M_1^2 M_2}{2} + \frac{J_p(u)}{2}$$

d'où,

$$n M_1^2 M_2 \int_s^t \sigma(u)\, du + K \geqq \int_s^t J_p(u)\, \sigma(u)\, du$$

Notons que si on s'autorise l'hypothèse (H_2), ce résultat de finitude peut s'obtenir directement, grâce aux majorations en norme C^2 des densités m_u (qui peuvent se déduire des rappels du début de la section 2), et à la relation (†) appliquée avec $\theta = \mathbb{R}^n$.

On fait alors tendre p vers l'infini, pour s'apercevoir, par convergence monotone, que $\int_s^t du\, \sigma(u) \int g_{\sigma(u)} \left| \nabla \sqrt{\frac{m_u}{g_{\sigma(u)}}} \right|^2 dx$ est une quantité finie $\left[\text{majorée par } n M_1^2 M_2 \int_s^t \sigma(u)\, du + K\right]$.

Revenons à (∗∗) avec $\rho = \rho_p$. On fait tendre p vers l'infini.

Le premier terme du membre de droite tend, par convergence monotone (ou dominée, grâce au résultat ci-dessus), vers

$$2 \int_s^t du\, \sigma(u) \int g_{\sigma(u)} \left| \nabla \sqrt{\frac{m_u}{g_{\sigma(u)}}} \right|^2 dx$$

Le second terme est borné par

$$2 \int_s^t du\, \sigma(u) \sqrt{\int g_{\sigma(u)} \left| \nabla \sqrt{\frac{m_u}{g_{\sigma(u)}}} \right|^2 dx} \sqrt{\int m_u \left(\ln \frac{m_u}{g_{\sigma(u)}}\right)^2 |\nabla \sqrt{\rho_p}|^2 dx}$$

qui est une quantité finie, et qui tend vers 0, quand p tend vers l'infini.

Enfin, le dernier terme tend vers

$$\int_s^t du \left(\frac{d}{du} \frac{1}{\sigma(u)}\right) \int m_u [U - \langle U \rangle_u]\, dx$$

par convergence dominée.

Intéressons-nous au membre de gauche de (∗∗). Il est clair (en considérant séparément les points tels que $m_t \geqq 1$ et les autres), toujours par convergence dominée, que

$$\lim_{p \to \infty} \int \rho_p m_t \ln \frac{m_t}{g_{\sigma(t)}} dx = \mathrm{I}_t(m_t)$$

$$\lim_{p \to \infty} \int \rho_p m_s \ln \frac{m_s}{g_{\sigma(s)}} dx = \mathrm{I}_s(m_s)$$

Finalement, on a donc obtenu que

$$\forall t > s > 0,$$

$$\mathrm{I}_t(m_t) - \mathrm{I}_s(m_s) = \int_s^t du \left[\frac{d}{du}\left(\frac{1}{\sigma(u)}\right) \int m_u [\mathrm{U} - \langle \mathrm{U} \rangle_u] \, dx - 2\sigma(u) \int g_{\sigma(u)} \left| \nabla \sqrt{\frac{m_u}{g_{\sigma(u)}}} \right|^2 dx \right]$$

d'où le résultat annoncé. □

Pour obtenir une inégalité différentielle satisfaite par $\mathrm{I}_t(m_t)$, il nous reste à comparer les termes

$$\int \left| \nabla \sqrt{\frac{m_t}{g_{\sigma(t)}}} \right|^2 g_{\sigma(t)} \, dx$$

et

$$\mathrm{I}_t(m_t) = \int \frac{m_t}{g_{\sigma(t)}} \ln\left(\frac{m_t}{g_{\sigma(t)}}\right) g_{\sigma(t)} \, dx$$

Notons que nous pourrions le faire directement, si on disposait d'inégalités de Sobolev-logarithmiques pour les $g_{\sigma(t)}$, *i.e.* si on savait que pour tout $\sigma(t)$, il existe une constante $\tilde{a}(\sigma((t))) > 0$ telle que

$$\forall f \in \mathrm{C}^1(\mathbb{R}^n), \quad f > 0,$$

$$\int f^2 \ln(f^2) g_{\sigma(t)} \, dx \leqq \tilde{a}(\sigma(t)) \int |\nabla f|^2 g_{\sigma(t)} \, dx + \int f^2 g_{\sigma(t)} \, dx \ln\left[\int f^2 g_{\sigma(t)} \, dx\right]$$

Pour $\sigma(t) > 0$ fixé, soit $a(\sigma(t))$ l'infimum des nombres $\tilde{a}(\sigma(t))$ admissibles pour les inégalités ci-dessus. Pour pouvoir exhiber des taux de décroissance de la température qui entrainent

$$\lim_{t \to \infty} \mathrm{I}_t(m_t) = 0$$

il nous suffirait de connaître le comportement asymptotique de $a(\sigma(t))$ quand $\sigma(t)$ tend vers 0. Cependant, un inconvénient majeur de cette méthode est qu'on n'a pas, en toute généralité sous les hypothèses (H_1) et (H_2), des inégalités de Sobolev-logarithmiques pour les $g_{\sigma(t)}$. Ainsi, par exemple, il est connu que si pour $|x|$ assez grand, on a $U(x)=|x|^\alpha$, avec $\alpha \leqq \frac{3}{2}$, alors les mesures $g_{\sigma(t)}$ ne satisfont pas d'inégalités de Sobolev-logarithmiques.

Par contre, précisons que ce programme est tout à fait adapté, via les inégalités de Sobolev-logarithmiques prouvées par Holley, Kusuoka et Stroock dans [5] et [6], à l'étude de l'algorithme du recuit simulé en temps continu, quand l'espace des phases, au lieu d'être $\mathbb{R}^n$, est une variété riemannienne compacte, ou un ensemble fini (on considère dans ce cas un processus de sauts), *cf.* [10].

Cette démarche est également utilisable, si on suppose, par exemple, que pour $|x|$ assez grand, $U(x)=|x|^2$. En effet, on peut montrer assez facilement, en partant des inégalités de Sobolev-logarithmiques pour les mesures gaussiennes (*cf.* Gross [3]), qu'il existe une constante $c'>0$ telle que pour tout $0<\sigma(t)\leqq 1$, on ait $a(\sigma(t))\leqq \exp\left(\frac{c'}{\sigma(t)}\right)$. Ceci nous permet de conclure (*cf.* le lemme de la section 5), que si pour t assez grand on a

$$\sigma(t)=\frac{k}{\ln(t)}, \qquad \text{avec} \quad k>c',$$

alors,

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

Le théorème 1 est donc démontré, dans ce cas, avec c' à la place de c. Il nous resterait à comparer ces deux constantes.

En fait, pour comparer $\int \left|\nabla \sqrt{\frac{m_t}{g_{\sigma(t)}}}\right|^2 dg_{\sigma(t)}$ et $I_t(m_t)$, nous passerons par l'intermédiaire de $\int\left(\sqrt{\frac{m_t}{g_{\sigma(t)}}}-\left\langle \sqrt{\frac{m_t}{g_{\sigma(t)}}}\right\rangle_t\right)^2 g_{\sigma(t)}\,dx$. En effet, par définition de λ_t, on a

$$\int\left(\sqrt{\frac{m_t}{g_{\sigma(t)}}}-\left\langle \sqrt{\frac{m_t}{g_{\sigma(t)}}}\right\rangle_t\right)^2 g_{\sigma(t)}\,dx \leqq \frac{1}{2}\sigma(t)\lambda_t^{-1}\int\left|\nabla\sqrt{\frac{m_t}{g_{\sigma(t)}}}\right|^2 dg_{\sigma(t)}$$

[Les points qui annulent m_t ne posent pas vraiment de problème, comme on peut s'en persuader en considérant d'abord $\sqrt{\frac{m_t}{g_{\sigma(t)}}+\varepsilon}$, pour $\varepsilon>0$.]

Il ne nous reste donc plus qu'à comparer

$$\mathrm{I}_t(m_t) \qquad \text{et} \qquad \int\left(\sqrt{\frac{m_t}{g_{\sigma(t)}}}-\left\langle\sqrt{\frac{m_t}{g_{\sigma(t)}}}\right\rangle_t\right)^2 g_{\sigma(t)}\,dx,$$

ce qui est le but de la section suivante.

On pourrait penser que cette double comparaison risque de ne pas être très précise du point de vue des estimations. En fait, il n'en est rien (ou presque), puisque ce passage nous permettra de prouver la convergence de l'algorithme du recuit simulé pour la constante c, qui est fortement suspectée d'être la meilleure constante pour laquelle le théorème 1 est satisfait (on sait que c'est le cas si on se place sur une variété riemannienne, compacte et connexe (ou un ensemble fini) *cf.* [6]).

4. COMPARAISON DES DEUX TERMES PRÉCÉDENTS

Commençons par un lemme général.

Soient μ une probabilité, et f une fonction de $\mathrm{L}^2(\mu)$,positive et telle que $\int f^2\,d\mu=1$.

On notera $\langle f\rangle=\int f\,d\mu$.

LEMME 4. — *Il existe un nombre* $\delta_0>0$, *indépendant de* μ *et* f, *tel que*

$$\forall\, 0<\delta<\delta_0$$

$$\frac{1}{\delta}\int(f-\langle f\rangle)^2\,d\mu+4\,\delta\int f^2\,[\ln(f+e)+1]^2\,d\mu \geqq \int f^2\ln(f)\,d\mu$$

Démonstration. — Soit G l'application

$$\mathbb{R}_+\to\mathbb{R}$$
$$x\mapsto x^2\ln x$$

Pour $\alpha>0$ et $\rho>-\langle f\rangle$, définissons

$$\mathrm{F}_\alpha(\rho)=\mathrm{G}(\langle f\rangle+\rho)-\mathrm{G}(\langle f\rangle)-\mathrm{G}'(\langle f\rangle)\,\rho-\alpha\rho^2$$

alors,

$$\mathrm{F}_\alpha(0)=\mathrm{F}'_\alpha(0)=0$$

et pour $\rho > -\langle f \rangle$,

$$F''_\alpha(\rho) = 3 + 2\ln(\langle f \rangle + \rho) - 2\alpha$$

Il s'ensuit que pour tout $M > 0$, il existe un $\alpha(M) > 0$ tel que

$$\forall -\langle f \rangle < \rho \leqq M,$$
$$F_{\alpha(M)}(\rho) \leqq 0$$

Remarquons, d'autre part, que

$$\rho \geqq 1 \quad \Rightarrow \quad (\rho + \langle f \rangle)^2 \ln(\rho + \langle f \rangle)$$
$$\leqq G'(\langle f \rangle)\rho + [\ln(\rho + \langle f \rangle + e) + 1](\rho + \langle f \rangle)^2$$

En effet, il suffit de voir que pour $\rho \geqq 1$, on a

$$G'(\langle f \rangle)\rho + \rho^2 \geqq 0$$

ce qui est clairement vérifié, puisque $|G'(\langle f \rangle)| \leqq 1$ (car $0 < \langle f \rangle \leqq 1$).

Prenons $M = 1$ et posons $\alpha = \alpha(1)$, la constante correspondante. On a alors,

$$\forall \rho > -\langle f \rangle,$$
$$(\rho + \langle f \rangle)^2 \ln(\rho + \langle f \rangle)$$
$$\leqq G(\langle f \rangle) 1_{\rho \leqq 1} + \alpha\rho^2 1_{\rho \leqq 1} + G'(\langle f \rangle)\rho$$
$$+ [\ln(\rho + \langle f \rangle + e) + 1](\rho + \langle f \rangle)^2 1_{\rho > 1}$$
$$\leqq \alpha\rho^2 1_{\rho \leqq 1} + G'(\langle f \rangle)\rho$$
$$+ \left[\frac{1}{\delta}\rho^2 + \delta[\ln(\rho + \langle f \rangle + e) + 1]^2 (\rho + \langle f \rangle)^2 \left(\frac{\rho + \langle f \rangle}{\rho}\right)^2\right] 1_{\rho > 1}$$

ceci pour tout $\delta > 0$. [On aura remarqué que $0 < \langle f \rangle \leqq 1 \Rightarrow G(\langle f \rangle) \leqq 0$.]

Or, $\rho > 1$ implique que

$$\frac{\rho + \langle f \rangle}{\rho} \leqq 2$$

ainsi, si on pose $\delta_0 = \alpha^{-1}$, on a pour tout $0 < \delta \leqq \delta_0$,

$$(\rho + \langle f \rangle)^2 \ln(\rho + \langle f \rangle)$$
$$\leqq G'(\langle f \rangle)\rho + \frac{1}{\delta}\rho^2 + 4\delta[\ln(\rho + \langle f \rangle + e) + 1]^2 (\rho + \langle f \rangle)^2$$

On pose alors $\rho = f - \langle f \rangle$, puis on intègre l'inégalité ci-dessus par rapport à $d\mu$ pour obtenir le résultat annoncé. □

Pour pouvoir appliquer avec profit le lemme précédent, nous avons besoin de la majoration suivante.

PROPOSITION 5. — *Supposons que les hypothèses* (H_1) *et* (H_3) *soient satisfaites, que m admette des moments de tous ordres et que pour* $t \geqq t_0 > 1$, *on ait, pour une certaine constante* $l > 0$, $\sigma(t) = \frac{l}{\ln(t)}$. *Alors, pour tout* $\varepsilon > 0$, *il existe une constante* $A_\varepsilon > 0$ *telle que,*

$$\forall t \geqq t_0,$$
$$\int \left(\ln\left(\frac{m_t}{g_{\sigma(t)}} + e\right)\right)^2 m_t(dx) \leqq A_\varepsilon \exp\left(\frac{\varepsilon}{\sigma(t)}\right)$$

Démonstration. — Soit $\rho \in C_c^\infty(\mathbb{R}^n)$. Pour des raisons techniques, on commence par étudier l'évolution de

$$H_t = \int \rho \left(\ln\left(\frac{m_t}{g_{\sigma(t)}} + e\right)\right)^2 m_t\, dx$$

Il est clair que $t \mapsto H_t$ est une application continûment dérivable sur $]0, +\infty[$, et que pour tout $t > 0$,

$$\begin{aligned} \frac{dH_t}{dt} &= 2\int \rho \left(\ln\left(\frac{m_t}{g_{\sigma(t)}} + e\right)\right)\left(\frac{m_t}{g_{\sigma(t)}} + e\right)^{-1} \frac{\partial}{\partial t}\left(\frac{m_t}{g_{\sigma(t)}}\right) m_t\, dx \\ &\quad + \int \rho \left(\ln\left(\frac{m_t}{g_{\sigma(t)}} + e\right)\right)^2 \frac{\partial m_t}{\partial t} dx \\ &= \int \rho\, F\left(\frac{m_t}{g_{\sigma(t)}}\right) \frac{\partial m_t}{\partial t} dx \\ &\quad - 2\int \rho \left(\ln\left(\frac{m_t}{g_{\sigma(t)}} + e\right)\right)\left(\frac{m_t}{g_{\sigma(t)}} + e\right)^{-1} \left(\frac{m_t}{g_{\sigma(t)}}\right)^2 \frac{\partial g_{\sigma(t)}}{\partial t} dx \end{aligned}$$

où on a posé,

$$\begin{aligned} F:\ & [0, +\infty[\to \mathbb{R} \\ & x \mapsto 2\ln(x+e)\frac{x}{x+e} + \ln^2(x+e) \end{aligned}$$

On vérifie immédiatement que pour x positif ou nul, $F'(x)$ est strictement positif.

On intègre par parties le premier terme de l'expression précédente pour $\dfrac{dH_t}{dt}$, en utilisant le fait que

$$\frac{\partial m_t}{\partial t} = \frac{\sigma(t)}{2} \sum_{k=1}^{k=n} \frac{\partial}{\partial x_k}\left(g_{\sigma(t)} \frac{\partial}{\partial x_k}\left(\frac{m_t}{g_{\sigma(t)}}\right)\right)$$

pour obtenir

$$\frac{dH_t}{dt} = -\frac{\sigma(t)}{2}\int \rho \sum_{k=1}^{k=n} F'\left(\frac{m_t}{g_{\sigma(t)}}\right)\left(\frac{\partial}{\partial x_k}\left(\frac{m_t}{g_{\sigma(t)}}\right)\right)^2 g_{\sigma(t)}\, dx$$

$$-\frac{\sigma(t)}{2}\int F\left(\frac{m_t}{g_{\sigma(t)}}\right)\left\langle \nabla\rho, \nabla\frac{m_t}{g_{\sigma(t)}}\right\rangle g_{\sigma(t)}\, dx$$

$$+2\frac{d}{dt}\left(\frac{1}{\sigma(t)}\right)\int \rho\left(\ln\left(\frac{m_t}{g_{\sigma(t)}}+e\right)\right)$$

$$\times\left(\frac{m_t}{g_{\sigma(t)}}+e\right)^{-1}\left(\frac{m_t}{g_{\sigma(t)}}\right)^2 [U-\langle U\rangle_t]\, g_{\sigma(t)}\, dx$$

On considère les $(\rho_p)_{p\in\mathbb{N}}$ qui sont apparus dans la démonstration de la proposition 4, et on prend $\rho=\rho_p$.

Remarquons que

$$\left|\int F\left(\frac{m_t}{g_{\sigma(t)}}\right)\left\langle \nabla\rho_p, \nabla\frac{m_t}{g_{\sigma(t)}}\right\rangle g_{\sigma(t)}\, dx\right|$$

$$=\left|\int F\left(\frac{m_t}{g_{\sigma(t)}}\right)\left\langle \nabla\rho_p, \nabla\frac{m_t}{g_{\sigma(t)}}\right\rangle\right.$$

$$\left.\times\sqrt{\rho_p F'\left(\frac{m_t}{g_{\sigma(t)}}\right)}\left(\sqrt{\rho_p F'\left(\frac{m_t}{g_{\sigma(t)}}\right)}\right)^{-1} g_{\sigma(t)}\, dx\right|$$

$$\leqq 2\left(\int \rho_p F'\left(\frac{m_t}{g_{\sigma(t)}}\right)\left|\nabla\frac{m_t}{g_{\sigma(t)}}\right|^2 g_{\sigma(t)}\, dx\right)^{1/2}$$

$$\times\left(\int\frac{F^2}{F'}\left(\frac{m_t}{g_{\sigma(t)}}\right)|\nabla\sqrt{\rho_p}|^2 g_{\sigma(t)}\, dx\right)^{1/2}$$

Mais on voit facilement qu'il existe une constante $C_1>0$ telle que

$$\forall x\geqq 0,$$
$$\frac{F^2}{F'}(x)\leqq C_1\,[x\ln^3(x+e)+1]$$

ce qui implique qu'il existe une constante $C_2>0$ (indépendante de p) telle que

$$\int\frac{F^2}{F'}\left(\frac{m_t}{g_{\sigma(t)}}\right)|\nabla\sqrt{\rho_p}|^2 g_{\sigma(t)}\, dx\leqq C_2\left[1+\int\ln^3\left(\frac{m_t}{g_{\sigma(t)}}+e\right)m_t\, dx\right]$$

En appliquant la remarque de la fin de la section 2, on se rend compte que cette dernière expression est bornée, en tant que fonction de t, sur les compacts de $]0, +\infty[$.

Il en est de même de

$$t \mapsto \frac{d}{dt}\left(\frac{1}{\sigma(t)}\right)\int \rho_p\left(\ln\left(\frac{m_t}{g_{\sigma(t)}}+e\right)\right)\left(\frac{m_t}{g_{\sigma(t)}}+e\right)^{-1}\left(\frac{m_t}{g_{\sigma(t)}}\right)^2 [\mathrm{U}-\langle \mathrm{U}\rangle_t]\, g_{\sigma(t)}\, dx$$

uniformément en p, comme on peut le voir grâce à une application des inégalités de Hölder, de la remarque de la fin de la section 2 et du lemme 2.

Il s'en suit, en passant à la limite quand p tend vers l'infini et en s'inspirant de la démonstration de la proposition 3, que pour tout $s>r>0$, on a

$$\int_r^s dt \int \mathrm{F}'\left(\frac{m_t}{g_{\sigma(t)}}\right)\left|\nabla \frac{m_t}{g_{\sigma(t)}}\right|^2 g_{\sigma(t)}\, dx < \infty$$

puis que l'application

$$\mathrm{J}: \quad t \mapsto \int \ln^2\left(\frac{m_t}{g_{\sigma(t)}}+e\right) m_t\, dx$$

est absolument continue sur $]0, +\infty[$ et que p. s. pour $t>0$,

$$\begin{aligned}\frac{d\mathrm{J}_t}{dt} = &-\frac{\sigma(t)}{2}\int \mathrm{F}'\left(\frac{m_t}{g_{\sigma(t)}}\right)\left|\nabla \frac{m_t}{g_{\sigma(t)}}\right|^2 g_{\sigma(t)}\, dx \\ &+ 2\frac{d}{dt}\left(\frac{1}{\sigma(t)}\right)\int \ln\left(\frac{m_t}{g_{\sigma(t)}}+e\right)\left(\frac{m_t}{g_{\sigma(t)}}+e\right)^{-1}\frac{m_t}{g_{\sigma(t)}}[\mathrm{U}-\langle \mathrm{U}\rangle_t]\, m_t\, dx\end{aligned}$$

Ainsi, puisque F' et U sont positives et que pour $t \geqq t_0$, $\sigma(t)$ est décroissante, on a

$$\forall t \geqq t_0, \quad \frac{d\mathrm{J}_t}{dt} \leqq 2\frac{d}{dt}\left(\frac{1}{\sigma(t)}\right)\int \ln\left(\frac{m_t}{g_{\sigma(t)}}+e\right)\left(\frac{m_t}{g_{\sigma(t)}}+e\right)^{-1}\frac{m_t}{g_{\sigma(t)}}\mathrm{U}\, m_t\, dx$$

Cependant,

$$\begin{aligned}\int \ln\left(\frac{m_t}{g_{\sigma(t)}}+e\right)&\left(\frac{m_t}{g_{\sigma(t)}}+e\right)^{-1}\frac{m_t}{g_{\sigma(t)}}\mathrm{U}\, m_t\, dx \\ &\leqq \left(\int \ln^2\left(\frac{m_t}{g_{\sigma(t)}}+e\right) m_t\, dx\right)^{1/2}\left(\int \mathrm{U}^2\, m_t\, dx\right)^{1/2}\end{aligned}$$

et, d'après le lemme 2, si $\varepsilon>0$ est donné, vu la forme de $\sigma(t)$, il existe une constante $\mathrm{C}_\varepsilon>0$ telle que pour tout $t \geqq t_0>1$, on ait

$$\int \mathrm{U}^2\, m_t\, dx \leqq \mathrm{C}_\varepsilon \exp\left(\frac{2\varepsilon}{\sigma(t)}\right)$$

ainsi,

$$\forall t \geqq t_0,$$
$$\frac{dJ_t}{dt} \leqq 2 C_\varepsilon^{1/2} \exp\left(\frac{\varepsilon}{\sigma(t)}\right) \frac{d}{dt}\left(\frac{1}{\sigma(t)}\right) J_t^{1/2}$$

ce qui s'écrit encore

$$\frac{dJ_t^{1/2}}{dt} \leqq 2 \frac{C_\varepsilon^{1/2}}{\varepsilon} \frac{d}{dt}\left(\exp\left(\frac{\varepsilon}{\sigma(t)}\right)\right)$$

On en déduit aisément le résultat annoncé. □

Appliquons les résultats obtenus dans cette section.

Supposons que pour $t \geqq t_0 > 1$, on ait $\sigma(t) = \frac{l}{\ln(t)}$, pour un certain nombre $l > 0$. D'après la proposition 5, si les hypothèses (H_1) et (H_3) sont satisfaites et si m admet des moments de tous ordres, alors pour tout $\varepsilon > 0$, il existe une constante $A_\varepsilon > 0$ telle que

$$\forall t \geqq t_0,$$
$$\int \ln^2\left(\frac{m_t}{g_{\sigma(t)}} + e\right) m_t \, dx \leqq A_\varepsilon \exp\left(\frac{\varepsilon}{\sigma(t)}\right)$$

Cependant, le lemme 4 appliqué avec $\mu = g_{\sigma(t)}$ et $f = \sqrt{\frac{m_t}{g_{\sigma(t)}}}$, nous indique que

$$\forall 0 < \delta < \delta_0, \quad \forall t \geqq 0,$$
$$\int \left(\sqrt{\frac{m_t}{g_{\sigma(t)}}} - \left\langle \sqrt{\frac{m_t}{g_{\sigma(t)}}} \right\rangle_t\right)^2 g_{\sigma(t)} \, dx \geqq \frac{\delta}{2} I_t(m_t) - 4\delta^2 \int \left(\ln\left(\frac{m_t}{g_{\sigma(t)}} + e\right) + 1\right)^2 m_t \, dx$$

Soit $T_\varepsilon > t_0$ tel que pour tout $t \geqq T_\varepsilon$, on ait $\exp\left(-2 \frac{\varepsilon}{\sigma(t)}\right) \leqq \delta_0$.

On prend alors $\delta = \exp\left(-2 \frac{\varepsilon}{\sigma(t)}\right)$ dans l'inégalité précédente et on applique la proposition 5, pour obtenir que pour tout $t \geqq T_\varepsilon$,

$$\int \left(\sqrt{\frac{m_t}{g_{\sigma(t)}}} - \left\langle \sqrt{\frac{m_t}{g_{\sigma(t)}}} \right\rangle_t\right)^2 g_{\sigma(t)} \, dx \geqq \frac{1}{2} \exp\left(-2 \frac{\varepsilon}{\sigma(t)}\right) I_t(m_t) - 8 \exp\left(-4 \frac{\varepsilon}{\sigma(t)}\right) \left[1 + A_\varepsilon \exp\left(\frac{\varepsilon}{\sigma(t)}\right)\right]$$

On a donc prouvé l'inégalité suivante, entre les deux termes que nous cherchions à comparer; pour tout $\varepsilon>0$, il existe deux constantes positives $\tilde{A}_\varepsilon$ et T_ε telles que

$$\forall\, t \geqq T_\varepsilon,$$

$$\int \left(\sqrt{\frac{m_t}{g_{\sigma(t)}}} - \left\langle \sqrt{\frac{m_t}{g_{\sigma(t)}}} \right\rangle_t \right)^2 g_{\sigma(t)}\, dx \geqq \frac{1}{2} \exp\left(-2\frac{\varepsilon}{\sigma(t)}\right) \left[I_t(m_t) - \tilde{A}_\varepsilon \exp\left(-\frac{\varepsilon}{\sigma(t)}\right)\right]$$

5. DÉMONSTRATION DU THÉORÈME 1

En regroupant les résultats des sections précédentes, il est clair que $I_t(m_t)$ satisfait certaines inégalités différentielles. Nous allons prouver que celles-ci impliquent, sous les hypothèses du théorème 1, que $I_t(m_t)$ tend vers 0 en l'infini.

On suppose que les hypothèses (H_1) et (H_3) sont satisfaites, que m admet des moments de tous ordres et que pour $t \geqq t_0 > 1$, on ait $\sigma(t) = \dfrac{k}{\ln(t)}$, où k est un nombre strictement plus grand que la constante c, apparaissant dans le théorème 1.

Soit $\varepsilon = \dfrac{k-c}{5}$.

D'après la définition de c et du fait que $\sigma(t)$ tend vers 0 en l'infini, il existe un temps $T'_\varepsilon > 0$ tel que pour tout $t \geqq T'_\varepsilon$, on ait

$$-\sigma(t)\ln(\lambda_t) \leqq c+\varepsilon$$

Ainsi si T_ε et $\tilde{A}_\varepsilon$ sont les constantes qui apparaissent à la fin de la section précédente, on a pour tout $t \geqq T_\varepsilon \vee T'_\varepsilon$,

$$\begin{aligned}\int \left| \nabla \sqrt{\frac{m_t}{g_{\sigma(t)}}} \right|^2 g_{\sigma(t)}\, dx &\geqq 2\,\sigma(t)^{-1} \exp\left(-\frac{c+\varepsilon}{\sigma(t)}\right) \int \left(\sqrt{\frac{m_t}{g_{\sigma(t)}}} - \left\langle \sqrt{\frac{m_t}{g_{\sigma(t)}}} \right\rangle_t \right)^2 g_{\sigma(t)}\, dx \\ &\geqq \sigma(t)^{-1} \exp\left(-\frac{c+\varepsilon}{\sigma(t)}\right) \exp\left(-\frac{2\varepsilon}{\sigma(t)}\right) \left[I_t(m_t) - \tilde{A}_\varepsilon \exp\left(-\frac{\varepsilon}{\sigma(t)}\right)\right]\end{aligned}$$

En substituant ceci dans la formule de la proposition 3, on obtient

$$\frac{d\mathrm{I}_t(m_t)}{dt} \leqq \frac{d}{dt}\left(\frac{1}{\sigma(t)}\right)\int \mathrm{U}\,(m_t - g_{\sigma(t)})\,dx$$
$$-2\exp\left(-\frac{c+3\varepsilon}{\sigma(t)}\right)\left[\mathrm{I}_t(m_t) - \tilde{\mathrm{A}}_\varepsilon \exp\left(-\frac{\varepsilon}{\sigma(t)}\right)\right]$$

Mais d'après le lemme 2, il existe une constante $\mathrm{M}_\varepsilon > 0$ telle que pour tout $t \geqq 0$,

$$\int \mathrm{U}\, m_t\, dx \leqq \mathrm{M}_\varepsilon \exp\left(\frac{\varepsilon}{\sigma(t)}\right)$$

Ainsi, pour $t \geqq \mathrm{T}_\varepsilon \vee \mathrm{T}'_\varepsilon$,

$$\frac{d\mathrm{I}_t(m_t)}{dt} \leqq \alpha_t - \beta_t\, \mathrm{I}_t(m_t)$$

où,

$$\alpha_t = -\mathrm{M}_\varepsilon \frac{d}{dt}\left(\frac{1}{\sigma(t)}\right)\exp\left(\frac{\varepsilon}{\sigma(t)}\right) + \tilde{\mathrm{A}}_\varepsilon \exp\left(-\frac{\varepsilon}{\sigma(t)}\right)\beta_t$$
$$\beta_t = 2\exp\left(-\frac{c+3\varepsilon}{\sigma(t)}\right)$$

Pour conclure, nous avons besoin du lemme classique suivant.

LEMME 6. — *Soit $f: [0, +\infty[\to \mathbb{R}_+$ une fonction absolument continue satisfaisant, p.s. en t, l'inégalité différentielle*

$$f'(t) \leqq \alpha_t - \beta_t f(t)$$

où α et β sont deux fonctions positives et continues telles que

$$\int^\infty \beta_t\, dt = \infty$$

$$\lim_{t\to\infty} \frac{\alpha_t}{\beta_t} = 0$$

Alors,

$$\lim_{t\to\infty} f(t) = 0$$

Démonstration. — On pose

$$g(t) = \exp\left(\int_0^t \beta_s\, ds\right) f(t)$$

g est une fonction absolument continue, et p. s. en t,

$$g'(t)=(f'(t)+\beta_t f(t))\exp\left(\int_0^t \beta_s\,ds\right)$$
$$\leqq \alpha_t \exp\left(\int_0^t \beta_s\,ds\right)$$

d'où,

$$g(t)\leqq g(0)+\int_0^t \alpha_s \exp\left(\int_0^s \beta_u\,du\right)ds$$

Ainsi, pour tout $t_0 \geqq 0$,

$$f(t)\leqq f(0)\exp\left(-\int_0^t \beta_s\,ds\right)$$
$$+\exp\left(-\int_0^t \beta_s\,ds\right)\int_0^{t_0} \alpha_s \exp\left(\int_0^s \beta_u\,du\right)ds$$
$$+\exp\left(-\int_0^t \beta_s\,ds\right)\int_{t_0}^t \alpha_s \exp\left(\int_0^s \beta_u\,du\right)ds$$

Soit $\varepsilon>0$. Choisissons $t_0\geqq 0$ tel que pour $t\geqq t_0$, on ait $\frac{\alpha_s}{\beta_s}\leqq \varepsilon$. Le dernier terme de l'inégalité ci-dessus est alors majoré, pour $t\geqq t_0$, par

$$\exp\left(-\int_0^t \beta_s\,ds\right)\int_{t_0}^t \varepsilon\beta_s \exp\left(\int_0^s \beta_u\,du\right)ds$$
$$\leqq \varepsilon \exp\left(-\int_0^t \beta_s\,ds\right)\left[\exp\left(\int_0^s \beta_u\,du\right)\right]_{t_0}^t \leqq \varepsilon$$

D'autre part, les deux premiers termes tendent vers 0, quand t tend vers l'infini, on a donc

$$\limsup_{t\to\infty} f(t)\leqq \varepsilon$$

et le résultat annoncé en découle, $\varepsilon>0$ pouvant être choisi arbitrairement petit. □

En appliquant ce lemme avec $f(t)=\mathrm{I}_{\mathrm{T}_\varepsilon \vee \mathrm{T}'_\varepsilon+t}(m_{\mathrm{T}_\varepsilon \vee \mathrm{T}'_\varepsilon+t})$ (remarquer que l'inégalité de Jensen implique la positivité de f), il suffit, pour finir la démonstration du théorème, de vérifier que

$$\int^{\infty} t^{-(c+3\varepsilon)/k}\,dt=+\infty$$
$$\lim_{t\to\infty}\frac{1}{t}\,t^{(c+4\varepsilon)/k}=0$$

ce qui est bien satisfait, puisque $k=c+5\varepsilon$.

APPENDICE

On va présenter des majorations sur les moments des m_t plus précises que celles du lemme 2, mais nous ne les démontrerons pas ici. On s'intéressera ensuite à certains potentiels, pour lesquels la démonstration des estimées précédentes est facile, et pour lesquels on prouvera que la convergence de l'algorithme du recuit simulé a lieu avec une décroissance de la température en $\frac{c}{\ln(t)-\mathrm{A}\ln(\ln t)}$, pour une certaine constante $\mathrm{A}>0$.

Nous supposons que la convention $\min_{x\in\mathbb{R}^n} \mathrm{U}(x)=0$ est satisfaite. On a le raffinement suivant du lemme 2.

LEMME 7. — *Supposons que l'hypothèse* (H_1) *est vérifiée, qu'il existe un nombre* $q>0$ *tel que pour* $|x|$ *suffisamment grand on ait* $|\nabla \mathrm{U}|(x)^2 \leqq \mathrm{U}^q(x)$, *que pour un* $p\geqq 1$, *la probabilité initiale m satisfait* $\int \mathrm{U}^p(x)\,m(dx)<\infty$, *et que pour t assez grand,* $\sigma(t)$ *soit décroissante et asymptotiquement équivalente à* $\frac{l}{\ln(t)}$, *où l est un nombre strictement positif.*

Alors, pour tout nombre $r>0$, *il existe une constante* $\mathrm{M}_r>0$ *telle que*

$$\forall t\geqq 0,$$
$$\int \mathrm{U}^p(x)\,m_t(dx)\leqq \mathrm{M}_r\left(\frac{1}{\sigma(t)}\right)^r$$

Pour la démonstration de ce lemme, on renvoie à [10]. Notons qu'il est alors facile d'adapter la démonstration de la proposition 5 pour obtenir le résultat suivant.

PROPOSITION 8. — *Supposons que les hypothèses* (H_1) *et* (H_3) *sont satisfaites, qu'il existe un nombre* $q>0$ *tel que pour* $|x|$ *suffisamment grand on ait* $|\nabla \mathrm{U}|(x)^2\leqq \mathrm{U}^q(x)$, *que la probabilité initiale m vérifie* $\int \mathrm{U}^{p+1}(x)\,m(dx)<\infty$, *pour un* $p\geqq 1$, *et que* $\sigma(t)$ *finisse par être décroissante et soit asymptotiquement équivalente à* $\frac{l}{\ln(t)}$, *où l est un nombre strictement positif.*

Alors, pour tout nombre $r>0$, il existe une constante $A_r>0$ telle que pour tout t assez grand, on ait

$$\int \ln^p\left(\frac{m_t}{g_{\sigma(t)}}+e\right) m_t\,dx \leqq A_r \frac{1}{\sigma(t)^{p+r}}$$

$\left[\right.$Remarquons que sous l'hypothèse (H_1), si ∇U est globalement lipschitzien, alors, on a pour tout $q>1$, $\lim_{|x|\to\infty} \frac{|\nabla U|}{U^q}(x)=+\infty.\left.\right]$

Cette majoration nous permet d'obtenir des inégalités différentielles satisfaites par $I_t(m_t)$, qui sont plus précises que celles qui découlent des estimées de la proposition 5, et qui sont intéressantes si on veut étudier finement la vitesse de convergence de l'algorithme du recuit simulé, ou trouver des taux de décroissance de la température qui impliquent $\lim_{t\to\infty} I_t(m_t)=0$ et qui soient un peu meilleurs que ceux du théorème 1.

Nous allons maintenant étudier un cas particulier. On supposera, dans la suite de cette section, qu'il existe des constantes $a>0$, $b\in\mathbb{R}$ et $R>0$, telles que le potentiel satisfasse,

$$|x|\geqq R \quad \Rightarrow \quad U(x)=a|x|^2+b$$

Sous cette hypothèse, la démonstration du lemme 7 est immédiate, et en fait, on peut même prendre $r=0$.

En effet, il suffit de voir que pour tout $p\geqq 1$,

$$t\mapsto \int_{\mathbb{R}^n} |x|^{2p} m_t(dx)$$

est une application bornée sur $\mathbb{R}_+$, si on suppose que la probabilité initiale satisfait $\int U^p(x)\,m(dx)<\infty$, et si σ est borné sur $\mathbb{R}_+$.

Mais ceci découle du fait qu'il existe des constantes M_p, $N_p>0$ telles que

$$\forall t\geqq 0,\quad \forall x\in\mathbb{R}^n,$$
$$\mathscr{L}_t(|\,.\,|^{2p})(x)\leqq M_p - N_p|x|^{2p}$$

Ainsi, une application de la formule d'Itô (avec les précautions habituelles) à la fonction $|\,.\,|^{2p}$ et à la diffusion décrite par (1), et un résultat classique sur les inégalités différentielles démontrent l'affirmation précédente de bornitude.

La proposition 8 s'en déduit aisément (également avec $r=0$), car rappelons qu'en notant J l'application

$$t\mapsto \int \ln^2\left(\frac{m_t}{g_{\sigma(t)}}+e\right) m_t(dx)$$

J satisfait l'inégalité différentielle

$$\frac{dJ_t}{dt} \leqq 2 \left| \frac{d}{dt} \frac{1}{\sigma(t)} \right| \left(\int (U - \langle U \rangle_t)^2 \, dm_t \right)^{1/2} J_t^{1/2}$$

(*cf.* la démonstration de la proposition 5).

Ainsi, si on suppose que $\sigma(t)$ finit par décroître et tend vers 0 en l'infini, il existe une certaine constante $A > 0$ telle que pour tout t assez grand, on ait

$$J_t \leqq A \frac{1}{\sigma(t)^2}$$

Il apparaît alors, en reprenant la démonstration de la fin de la section 4 et de la section 5, et en choisissant $\delta = \sigma(t)^3$, pour tout t suffisamment grand, que $I_t(m_t)$ finit par satisfaire l'inégalité différentielle

$$\frac{dI_t(m_t)}{dt} \leqq \alpha_t - \beta_t I_t(m_t)$$

avec,

$$\alpha_t = K_1 \left| \frac{d}{dt} \frac{1}{\sigma(t)} \right| + K_2 \sigma(t) \beta_t$$

$$\beta_t = 2 \sigma(t)^3 \lambda_t$$

où K_1, K_2 sont deux constantes strictement positives.

Cependant, d'après [8], l'hypothèse faite ici sur la forme de U implique qu'il existe un entier positif n_0 tel que pour $\sigma(t)$ suffisamment petit, on ait

$$\lambda_t \geqq (\sigma(t))^{n_0} \exp\left(- \frac{c}{\sigma(t)} \right)$$

On en déduit, si on prend, pour tout t assez grand,

$$\sigma(t) = \frac{c}{\ln(t) - A \ln(\ln(t))}$$

avec $A > n_0 + 3$, que les conditions

$$\int^{\infty} \beta_t \, dt = \infty$$

$$\lim_{t \to \infty} \frac{\alpha_t}{\beta_t} = 0$$

sont satisfaites, et donc que

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

RÉFÉRENCES

[1] O. CATONI, *Thèse de doctorat*, Université Paris-XI, Laboratoire de Statistiques Appliquées, 1990.
[2] T. S. CHIANG, C. R. HWANG et S. H. SHEU, Diffusion for Global Optimization in $\mathbb{R}^n$, *S.I.A.M. J. Control Optimization*, vol. **25**, n° 3, 1987, p.737-753.
[3] L. GROSS, Logarithmic Sobolev Inequalities, *Am. J. Math.*, vol. 97, 1976, p. 1061-1083.
[4] R. HOLLEY et D. STROOCK, Diffusion on an Infinite Dimensional Torus, *J.F.A.*, vol. **42**, 1981, p. 29-63.
[5] R. HOLLEY et D. STROOCK, Annealing via Sobolev Inequalities, *C.M.P.*, vol. **115**, 1988, p. 553-569.
[6] R. HOLLEY, S. KUSUOKA et D. STROOCK, Asymptotics of the Spectral Gap with Applications to the Theory of Simulated Annealing, *J.F.A.*, vol. **83**, 1989, p. 333-347.
[7] C. R. HWANG, Laplace's Method Revisited: Weak Convergence of Probability Measures, *Ann. Proba.*, vol. **8**, 1980, p. 1177-1182.
[8] S. JACQUOT, Comportement asymptotique de la seconde valeur propre des processus de Kolmogorov, *Preprint*, Département de Mathématiques et d'Informatique, Université d'Orléans, 1990.
[9] S. KUSUOKA et D. STROOCK, Applications of the Malliavin Calculus, Part I, *Taniguchi Symp. S.A. Katata*, 1982, p. 271-306.
[10] L. MICLO, *Thèse de doctorat*, Université Paris-VI (dont est extrait cet article), Laboratoire de Probabilités, 1991.
[11] G. ROYER, A Remark on Simulated Annealing of Diffusion Processes, *Preprint*, Département de Mathématiques et d'Informatique, Université d'Orléans, 1988.
[12] D. STROOCK et S. R. S. VARADHAN, *Multidimensional Diffusion Processes*, Springer-Verlag, 1979.

(*Manuscrit reçu le* 29 *janvier* 1991,
corrigé le 9 *juillet* 1991.)