

LAURENT MICLO

**Une étude des algorithmes de recuit simulé
sous-admissibles**

Annales de la faculté des sciences de Toulouse 6^e série, tome 4, n^o 4
(1995), p. 819-877

http://www.numdam.org/item?id=AFST_1995_6_4_4_819_0

© Université Paul Sabatier, 1995, tous droits réservés.

L'accès aux archives de la revue « Annales de la faculté des sciences de Toulouse » (<http://picard.ups-tlse.fr/~annales/>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

Une étude des algorithmes de recuit simulé sous-admissibles^(*)

LAURENT MICLO⁽¹⁾

RÉSUMÉ. — On considère les algorithmes du recuit simulé en temps continu, associés à un potentiel sur un graphe fini ou sur une variété compacte, dont la décroissance de la température est en $T_t = k \ln^{-1}(t+1)$, où $k > 0$. Mis à part pour un ensemble exceptionnel \mathcal{K} de valeurs de k , qui est borné et dont le seul point d'accumulation possible est 0 (et qui est fini si, par exemple, les minima locaux sont en nombre fini), on verra comment retrouver des résultats de Catoni sur le comportement p.s. et en loi de ces algorithmes en temps grand, en utilisant des techniques de semi-groupes introduites par Holley, Kusuoka et Stroock. Notamment, il existe un nombre fini $R(k)$ de cycles disjoints tels que la trajectoire finit p.s. par rester dans l'un des cycles (et est alors presque récurrente à l'intérieur), ce qui permet de voir que la loi du processus s'approche d'une combinaison linéaire des mesures de Gibbs associées à ces cycles (et à la température T_t). Par ailleurs, on vérifiera que l'espace vectoriel engendré par l'ensemble convexe des lois limites est de dimension $R(k)$, ce qui montre que pour $R(k) > 1$, la loi limite dépend de la distribution initiale.

ABSTRACT. — We consider simulated annealing algorithms in continuous time, associated to a potential on a finite graph or a compact manifold, whose evolution of the temperature is given by $T_t = k \ln^{-1}(t+1)$, with $k > 0$. Apart for an exceptional set \mathcal{K} of values of k , which is bounded and for which 0 is the only possible accumulation point (and which is finite if, for instance, the number of local minima of the potential is finite), we find again results of Catoni on the a.s. and in law behavior in large times of these algorithms, by using a semi-group approach introduced by Holley, Kusuoka and Stroock. In particular, there exists a finite number $R(k)$ of disjoint cycles, such that a.s. the trajectory eventually stay in one of the cycles and is then almost recurrent inside it. This localization property implies that the law of the process get closer and closer to some linear combination of the Gibbs measures associated to the cycles (and to the temperature T_t). We shall also verify that the dimension of the

(*) Reçu le 2 novembre 1993

(1) Institut de Recherche Mathématique Avancée, Université Louis-Pasteur et CNRS, 7 rue René Descartes, F-67084 Strasbourg Cedex (France)

convex set of the limit laws is $R(k)$, so for $R(k) > 1$, the limit depends on the initial distribution.

KEY-WORDS : Simulated Annealing, k -cycles.

1. Introduction

On va étudier le comportement en temps grand d'algorithmes du recuit simulé sur un graphe fini, dans le cas où l'on fait décroître la température trop rapidement pour qu'il converge vers les minima globaux du potentiel donné a priori.

Soit (M, μ, q) un graphe fini réversible et irréductible : M est un ensemble fini, μ est une probabilité chargeant tous les points de M et $q : M \times M \rightarrow \mathbb{R}_+$ est un noyau de probabilités de transitions irréductible (c'est-à-dire que pour tous $x, y \in M$, il existe une suite $(p(i))_{1 \leq i \leq n}$ d'éléments de M telle que $p(1) = x$, $p(n) = y$ et satisfaisant pour tout $1 \leq i < n$, $q(p(i), p(i+1)) > 0$) et réversible par rapport à μ (i.e. pour tous $x, y \in M$, $\mu(x)q(x, y) = \mu(y)q(y, x)$).

Soit U une fonction réelle sur M . Les algorithmes du recuit simulé sont une procédure stochastique de recherche des minima globaux de U , introduite par Kirkpatrick, Gelatt et Vecchi [12] et Geman et Geman [6]. Pour les décrire, considérons un paramètre $\beta \geq 0$ (l'inverse de la température) et définissons le noyau de probabilités de transitions q_β par

$$\forall x, y \in M, \quad q_\beta(x, y) = \begin{cases} \exp\left(-\beta(U(y) - U(x))_+\right) q(x, y) & \text{si } y \neq x \\ 1 - \sum_{z \neq x} q_\beta(x, z) & \text{si } y = x \end{cases}$$

(rappelons que $(\cdot)_+ = (\cdot \vee 0)$, où \vee désigne le maximum).

Il est clair que q_β est irréductible et qu'il est réversible par rapport à la mesure de Gibbs μ_β définie par

$$\forall x \in M, \quad \mu_\beta(x) = Z_\beta^{-1} \exp(-\beta U(x)) \mu(x),$$

où $Z_\beta = \sum_{y \in M} \exp(-\beta U(y)) \mu(y)$, car pour tous $x, y \in M$, on a

$$\mu_\beta(x) q_\beta(x, y) = Z_\beta^{-1} \exp(-\beta(U(y) \vee U(x))) \mu(x) q(x, y) = \mu_\beta(y) q_\beta(y, x).$$

On associe à q_β l'opérateur L_β agissant sur les fonctions réelles ϕ définies sur M , par

$$\forall x \in M, \quad L_\beta \phi(x) = \sum_{y \in M} (\phi(y) - \phi(x)) q_\beta(x, y).$$

Cet opérateur sera le générateur infinitésimal du processus du recuit simulé quand la température prendra la valeur β^{-1} . Plus précisément, soit $\Omega = D([0, +\infty[, M)$ l'ensemble des trajectoires *càdlàg* de $[0, +\infty[$ dans M . On note $(X_t)_{t \geq 0}$ le processus canonique des projections de Ω sur M , et pour $t \geq 0$, \mathcal{F}_t la tribu engendrée par $\{X_s; 0 \leq s \leq t\}$, ainsi que $\mathcal{F} = \bigvee_{t \geq 0} \mathcal{F}_t$.

Soit m une probabilité sur M (qui sera la probabilité initiale) et $\beta \in C^1([0, +\infty[, \mathbb{R}_+)$ (qui représentera l'évolution temporelle de l'inverse de la température).

Il est bien connu qu'il existe une unique probabilité $P_{m,\beta}$ sur (Ω, \mathcal{F}) telle que

$$X_0 \circ P_{m,\beta} = m$$

et telle que pour toute fonction ϕ sur M ,

$$t \mapsto \phi(X_t) - \phi(X_0) - \int_0^t L_{\beta_s} \phi(X_s) ds$$

soit une $(\mathcal{F}_t)_{t \geq 0}$ -martingale.

Le processus $(X_t)_{t \geq 0}$ sous une probabilité $P_{m,\beta}$ est appelé un algorithme du recuit simulé associé à l'espace des phases (M, μ, q) , au potentiel U , à la loi initiale m et à l'évolution β . Pour $t \geq 0$, on note m_t la loi de X_t .

En fait, on ne s'intéressera ici qu'à des évolutions β pour lesquelles il existe une constante strictement positive k telle que pour tout $t \geq 2$, on ait $\beta_t = k^{-1} \ln(t)$. Pour simplifier, on parlera alors d'un k -algorithme.

Le résultat classique en théorie du recuit simulé affirme qu'il existe une constante $\tilde{c} \geq 0$, telle que si l'on considère un k -algorithme avec $k > \tilde{c}$, alors m_t tend, en temps grand, à se concentrer au voisinage des minima globaux de U , mais que par contre, si $k < \tilde{c}$, alors m_t ne se concentre pas uniquement au voisinage de ces minima globaux, mais aussi au voisinage de certains minima locaux.

Nous allons retrouver ce résultat et plus généralement, on va se donner un $k > 0$ quelconque (mis à part un nombre fini de valeurs) et étudier le

comportement en temps grand de m_t . Ceci permettra notamment de voir exactement quels sont les minima locaux (outre les minima globaux qui sont toujours tous approchés avec une probabilité strictement positive) au voisinage desquels une “partie” non nulle de m_t tend à se concentrer en temps grand.

L’idée principale est de montrer que l’on peut “découper” l’espace des phases M en cellules disjointes de telle manière que presque sûrement, une trajectoire finisse par rester dans une même cellule et de telle sorte qu’aucune de ces cellules ne se vide complètement. Cette procédure de localisation permet alors d’étudier ce qui se passe en temps grand dans une cellule presque indépendamment de ce qui se passe dans les autres, et de voir ainsi que dans une cellule A , la loi renormalisée $m_{t,A}(\cdot) = m_t^{-1}(A)m_t(\cdot \cap A)$ s’approche de la mesure de Gibbs associée à cette cellule et à la température β_t^{-1} . De plus, comme on montrera que $m_t(A)$ converge vers une constante strictement positive, on pourra conclure à la convergence des m_t quand t tend vers l’infini et préciser la forme de la limite.

Soit $k > 0$ fixé, pour donner la définition d’une cellule de niveau k (en lesquelles on va décomposer M , si l’on considère un k -algorithme), il faut introduire les notions suivantes.

Soit A un sous-graphe connexe de (M, q) (i.e. la restriction de q à $A \times A$ est un noyau de transitions irréductible dans A). Pour $x, y \in A$, on note $\mathcal{C}_{x,y}^A$ l’ensemble des chemins allant de x à y dans A , c’est-à-dire l’ensemble des suites $p = (p(i))_{1 \leq i \leq n}$ d’éléments de A telles que $p(1) = x$, $p(n) = y$ et satisfaisant pour tout $1 \leq i < n$, $q(p(i), p(i+1)) > 0$.

L’élévation d’un tel chemin est le nombre

$$e(p) = \max_{1 \leq i \leq n} U(p(i)) .$$

On définit alors l’énergie minimale à fournir pour aller dans A de x à y ou de y à x comme étant

$$H_A(x, y) = \min_{p \in \mathcal{C}_{x,y}^A} e(p) - U(x) - U(y) + \min_{z \in A} U(z) ,$$

puis on associe à A la constante

$$c(A) = \max_{x,y \in A} H_A(x, y) .$$

Définissons également le fond de A comme étant la plus petite valeur de U sur A ,

$$f(A) = \min_{z \in A} U(z)$$

et l'ensemble des minima globaux relatifs à (A, U) ,

$$N_A = \{z \in A \mid U(z) = f(A)\}.$$

Il est facile de vérifier qu'en fait, on a pour tout $y \in N_A$ fixé,

$$c(A) = \max_{x \in A} H_A(x, y).$$

On appelle frontière de A l'ensemble ∂A des couples $(x, y) \in A \times A^c$ tels que $q(x, y) > 0$ (remarquons que par l'hypothèse de réversibilité, cette condition est équivalente à $q(y, x) > 0$). La hauteur de A est alors le nombre

$$h(A) = \min_{(x, y) \in \partial A} (U(x) \vee U(y)) - f(A).$$

On dira que le sous-graphe A est une cellule de niveau k (ou plus simplement une k -cellule) si

$$c(A) < k \quad \text{et} \quad h(A) > k.$$

On vérifiera dans la section suivante que, mis à part pour un nombre fini de valeurs de k , disons $\{k_1, \dots, k_T\}$, il existe une partition de M en k -cellules (cette partition n'est pas tout à fait unique en général).

D'autre part, définissons la mesure de Gibbs $\mu_{A, \beta}$ associée à un sous-graphe A et à l'inverse de la température $\beta \geq 0$, par

$$\mu_{A, \beta}(x) = \begin{cases} Z_{A, \beta}^{-1} \exp(-\beta U(x)) \mu(x) & \text{si } x \in A \\ 0 & \text{sinon,} \end{cases}$$

où $Z_{A, \beta} = \sum_{y \in A} \exp(-\beta U(y)) \mu(y)$ (on considérera aussi cette probabilité $\mu_{A, \beta}$ comme une probabilité sur A).

Il est clair que quand β tend vers l'infini, $\mu_{A, \beta}$ converge vers la probabilité $\mu_{A, \infty}$ définie par

$$\forall x \in M, \quad \mu_{A, \infty}(x) = (\mu(N_A))^{-1} \mu(x) 1_{N_A}(x).$$

Nous pouvons maintenant énoncer le principal résultat de cet article.

THÉORÈME 1. — *Soit $(m_t)_{t \geq 0}$ la famille des lois associées à un k -algorithme, avec $k > 0$, $k \notin \{k_1, \dots, k_T\}$. Notons $M = \bigsqcup_{1 \leq i \leq R} A_i$ une partition de M en k -cellules disjointes et non vides.*

Alors pour tout $1 \leq i \leq R$, $m_t(A_i)$ converge en temps grand vers une limite strictement positive λ_i et pour tout $x \in M$,

$$\lim_{t \rightarrow +\infty} m_t(x) = \sum_{i=1}^R \lambda_i \mu_{A_i, \infty}(x).$$

Ainsi les minima locaux où m_t tend à se concentrer en temps grand sont les éléments de $\bigcup_{i=1}^R N_{A_i}$. Cet ensemble ne dépend évidemment pas de la partition en k -cellules choisie et on peut voir facilement qu'un élément x de M y appartient si et seulement s'il existe une k -cellule A telle que $x \in N_A$.

Dans l'étude des algorithmes du recuit simulé, il apparaît en général deux constantes, $c(M)$ et \tilde{c} , cette dernière étant définie par

$$\tilde{c} = \max_{x \in M} \min_{y \in N_M} H_M(x, y).$$

Remarquons que $c(M) \geq \tilde{c} \geq 0$ et que ces deux constantes peuvent être égales, par exemple dans le cas où il n'y a qu'un seul minimum global. L'intérêt de ces constantes provient du fait que si l'on considère un k -algorithme avec $k > \tilde{c}$, alors m_t finit par se concentrer au voisinage des minima globaux de U (et il est connu que \tilde{c} est la meilleure constante qui vérifie cette propriété, voir par exemple [8], [1], [3] et [16]), et si de plus $k > c(M)$ alors on sait que la limite des m_t est exactement $\mu_{M, \infty}$ (outre les références précédentes voir [10], [14] et [2]).

Le théorème précédent permet de retrouver ces résultats, sous l'hypothèse $k \notin \{k_1, \dots, k_T\}$. En effet, $k > c(M)$ équivaut au fait que M est une k -cellule (et c'est alors la seule k -cellule) et $k > \tilde{c}$ implique que pour toute k -cellule A on a $f(A) = f(M)$, c'est-à-dire sous les hypothèses du théorème 1, $\bigcup_{i=1}^R N_{A_i} = N_M$, qui est l'ensemble des minima globaux de U sur M (remarquons que l'on a toujours l'inclusion $N_M \subset \bigcup_{i=1}^R N_{A_i}$). De plus, on voit que la constante \tilde{c} dans la condition $k > \tilde{c}$ est optimale pour espérer obtenir la convergence de m_t vers les minima globaux, car $k < \tilde{c}$ implique que dans une partition $M = \bigsqcup_{1 \leq i \leq R} A_i$ en k -cellules, il existe au moins une k -cellule A_i telle que $f(A_i) > f(M)$, c'est-à-dire que $N_M \neq \bigcup_{i=1}^R N_{A_i}$. D'où le corollaire suivant.

COROLLAIRE 2. — *Sous les hypothèses du théorème 1, on a*

$$k > \tilde{c} \Rightarrow \lim_{t \rightarrow +\infty} m_t(N_M) = 1$$

$$k < \tilde{c} \Rightarrow \limsup_{t \rightarrow +\infty} m_t(N_M) < 1.$$

En fait, l'hypothèse $k \notin \{k_1, \dots, k_T\}$ est inutile dans ce corollaire, comme on le verra directement dans la preuve du théorème 1, ce qui permet de retrouver entièrement le résultat classique.

Cependant, l'intérêt principal du théorème 1 est la description du comportement des algorithmes sous-admissibles (i.e. avec $k < \tilde{c}$), qui peut permettre de comprendre pourquoi dans la pratique (avec des horizons de temps finis [1], il est intéressant de considérer des algorithmes dont la décroissance de la température est de la forme $k(t) \ln^{-1}(t)$, où la fonction k est constante par palier (et décroissante).

Dans la cinquième section, on généralisera ces résultats au cas continu où M est une variété riemannienne compacte et connexe.

Dans la dernière section, on précisera le comportement p.s. des k -algorithmes : une trajectoire qui finit par rester dans une k -cellule A visitée, en des temps arbitrairement grands, tout voisinage d'un point quelconque de la composante connexe de l'ensemble $\{x \in A \mid U(x) < f(A) + k\}$ qui contient N_A (et finit par ne plus visiter les points qui n'appartiennent pas à la composante connexe de $\{x \in A \mid U(x) \leq f(A) + k\}$ contenant N_A), ce qui interdit notamment une convergence p.s. de tels algorithmes. Nous vérifierons également que pour $k < c(M)$, la loi limite dépend effectivement de la loi initiale, et ce faisant, on caractérisera le nombre R de k -cellules qui entrent dans une décomposition de M , comme la dimension de l'espace des solutions bornées d'une équation parabolique rétrograde.

Remarque. — Les propriétés des algorithmes du recuit simulé présentées dans cet article peuvent se déduire, dans le cas de processus à temps discret sur un graphe, des résultats très précis de Catoni [1], obtenus par une méthode basée sur des estimations type grandes déviations pour les temps et positions de sortie de cycles.

Mais nous utiliserons ici des techniques de semi-groupes, en poursuivant ainsi une approche inaugurée par Holley et Stroock dans [10], car elles sont mieux adaptées aux situations continues (en temps et en espace) et surtout

plus directes. Nous nous inspirerons d'ailleurs souvent des techniques et résultats de Holley, Kusuoka et Stroock, le point de départ de cette étude étant la section 3 de [9].

Revenons plus précisément sur les résultats de Catoni. On appelle cycle (notion introduite par Freidlin et Wentzell [4]) une composante connexe des ensembles de niveau de l'énergie. Soit $k > 0$ fixé. Considérons parmi les cycles $C \subset M$, ceux dont la hauteur satisfait $h(C) > k$ et qui sont minimaux pour l'inclusion ensembliste. Appelons-les des k -cycles, que l'on note désormais $C_1, \dots, C_{R'}$. Les résultats de Catoni (regrouper le théorème 2.25 et la proposition 4.5 de [1]) montrent que si l'on prend pour suite des températures $T_n = k \ln^{-1}(1 + n)$, alors pour tout $x \in M$,

$$\lim_{n \rightarrow +\infty} m_n(x) = \sum_{i=1}^{R'} \lambda_i \mu_{C_i, \infty}(x), \quad \text{où } \lambda_i > 0, \quad i = 1, \dots, R'$$

et presque sûrement X_n reste dans l'un des C_i (en fait, Catoni obtient un résultat plus fort, car il se permet des conditions un peu plus générales pour l'évolution de la température). Ainsi notamment Catoni n'a pas de restrictions sur $k > 0$ et la notion de cycle semble plus importante que celle de cellule, car les k -cycles sont uniques et non pas la partition en k -cellules. D'autant plus que, comme on vient de l'annoncer avant cette remarque, l'algorithme est transient à l'extérieur des k -cycles et finit par être presque récurrent à l'intérieur de l'un d'eux.

Cependant, nous avons trouvé commode techniquement de disposer d'une décomposition en k -cellules (notamment pour traiter le cas d'une variété compacte, car alors pour tout k -cycle C_i , on a $h(C_i) = k$ (éventuellement $+\infty$, si $R' = 1$) et il faudrait donc considérer de petites extensions des k -cycles), car outre qu'elle est naturelle pour la démonstration de la proposition 4, qui est la clé technique de l'article, elle permet aussi de se ramener directement aux cas $k > c(M)$, sans se soucier de la transience du complémentaire de $\bigcup_{i=1}^{R'} C_i$ (ceci découlant ensuite de l'étude du cas classique $k > c(M)$ à l'intérieur des cellules, sect. 6). De plus, elle a le mérite de faire apparaître l'ensemble \mathcal{K} qui a un intérêt propre, car c'est lui qui détermine le comportement asymptotique des valeurs propres des L_β quand β devient grand (voir [15]).

Mais il est vrai, a posteriori, pour se ramener à des notions plus classiques, que l'on aurait pu tout écrire en terme de k -cycles, nous décrirons dans la remarque de la fin de cet article, la démarche qu'il aurait fallu alors suivre.

2. Découpage en cellules

On va donner une démonstration dynamique de la construction d'une partition en k -cellules, en faisant varier k . Ceci permettra de donner une caractérisation de l'ensemble exceptionnel \mathcal{K} en terme de cycles.

Il est commode de disposer de différents types de cellules : on dira qu'un sous-graphe connexe A est une k -cellule large à bord relevé, en abrégé une k -clbr (respectivement une k -cellule semi-large à bord relevé du premier type, en abrégé une k -cslbr1, respectivement une k -cellule semi-large à bord relevé du second type, en abrégé une k -cslbr2) si

$$c(A) \leq k \quad \text{et} \quad h(A) \geq k$$

(respectivement $c(A) \leq k$ et $h(A) > k$, respectivement $c(A) < k$ et $h(A) \geq k$) et si pour tout $(x, y) \in \partial A$ et tout $z \in N_A$, il existe un chemin $p \in \mathcal{C}_{x',z}^{A \cup \{x'\}}$ tel que $e(p) = U(x')$, où $x' \in \{x, y\}$ est tel que $U(x') = U(x) \vee U(y)$.

Si on dispose d'une partition en k -clbr $M = \bigsqcup_{i=1}^R A_i$, on peut construire une partition en k -cslbr1 de la manière suivante :

- si $\min_{1 \leq i \leq R} h(A_i) > k$, il suffit de garder la même partition ;
- si $\min_{1 \leq i \leq R} h(A_i) = k$, soit $1 \leq i \leq R$ tel que $h(A_i) = k$, soit $(x, y) \in \partial A_i$ tel que

$$U(x) \vee U(y) = h(A_i) + f(A_i)$$

et enfin soit $j \neq i$ tel que $y \in A_j$.

On a nécessairement $f(A_i) \geq f(A_j)$ et il est alors clair que $A_i \cup A_j$ est encore une k -clbr. Considérons la partition $M = \bigsqcup_{\ell \neq i, \ell \neq j} A_\ell \sqcup (A_i \cup A_j)$. En répétant l'opération précédente un nombre fini de fois, on finit par obtenir une partition en k -cslbr1.

Notamment, en partant d'une partition en k -cslbr2 qui n'est pas une partition en k -cellules, on obtient de cette manière une partition en k -cslbr1 en diminuant strictement le nombre d'éléments de la partition.

Considérons l'ensemble L des minima locaux ; il s'agit des points $x \in M$ tels que pour tout $y \in M$ satisfaisant $U(y) < U(x)$, on ait $\min_{p \in \mathcal{C}_{x,y}^M} e(p) >$

$U(x)$. Soit \mathcal{L} l'ensemble des classes d'équivalence pour la relation \sim définie sur L par

$$x \sim y \iff U(x) = U(y) \quad \text{et} \quad \min_{p \in \mathcal{C}_{x,y}^M} e(p) = U(x).$$

Soit $\ell_1, \dots, \ell_{R(0)}$ des représentants des éléments de \mathcal{L} . On considère les domaines d'attraction $\tilde{A}_1, \dots, \tilde{A}_{R(0)}$ de ces points. Pour tout $1 \leq i \leq R(0)$, on dit que $x \in \tilde{A}_i$, s'il existe un chemin $p = (p_j)_{1 \leq j \leq n} \in \mathcal{C}_{x, \ell_i}^M$ décroissant pour U (i.e. qui satisfait $U(p_{j+1}) \leq U(p_j)$, $\forall 1 \leq j \leq n-1$). Cet ensemble ne dépend pas du choix du représentant et contient toute sa classe d'équivalence (et est d'intersection vide avec les autres classes). Puis on définit une partition $M = \bigsqcup_{i=1}^{R(0)} A_i^{(0)}$ en posant

$$\begin{aligned} A_1^{(0)} &= \tilde{A}_1 \\ A_2^{(0)} &= \tilde{A}_2 \setminus A_1^{(0)} \\ &\vdots \\ A_{R(0)}^{(0)} &= \tilde{A}_{R(0)} \setminus \left(\bigcup_{i=1}^{R(0)-1} A_i^{(0)} \right). \end{aligned}$$

Il est clair que pour tout $k > 0$ suffisamment petit, la décomposition précédente est une partition en k -cslbr1. Soit k_1 le supremum des valeurs de k pour lesquelles ceci est satisfait ($k_1 = \min_{1 \leq i \leq R(0)} h(A_i)$). Si $k_1 < +\infty$ ($k_1 = +\infty$ correspond au cas $c(M) = 0$ inintéressant en théorie du recuit simulé), les $(A_i^{(0)})_{1 \leq i \leq R(0)}$ forment une partition en k_1 -cslbr2 que l'on transforme en suivant la méthode précédente en une partition en k_1 -cslbr1, $M = \bigsqcup_{i=1}^{R(k_1)} A_i^{(k_1)}$. Puis on considère $k_2 > k_1$ le supremum des valeurs de k pour lesquelles ceci est satisfait. Ainsi de suite, on construit un ensemble fini $\mathcal{K} = \{k_1, \dots, k_T\}$ tel que pour tout $0 \leq i \leq T$ et tout $k_i \leq k < k_{i+1}$, on ait une partition en k -cslbr1, $M = \bigsqcup_{i=1}^{R(k_i)} A_i^{(k_i)}$ (où on a convenu de poser $k_0 = 0$ et $k_{T+1} = +\infty$, par ailleurs, notons que $k_T = c(M)$).

On en déduit la proposition suivante.

PROPOSITION 3. — *Pour tout $k > 0$, il existe au moins une partition de M en k -cslbr1 et une partition en k -cslbr2. De plus l'existence d'une partition en k -cellules équivaut à $k \notin \mathcal{K}$.*

Remarques

- Soit $k > 0$ et $R'(k)$ le nombre de k -cycles. Pour $k \notin \mathcal{K}$, on a $R'(k) = R(k)$, le nombre de k -cellules d'une partition de M (en reprenant les notations précédentes, $R(k) = R(k_i)$ où $k_i = \max\{k_j < k; 0 \leq j \leq T\}$), car il est facile de voir que toute k -cellule d'une décomposition de M contient un et un seul des k -cycles. On peut ainsi caractériser \mathcal{K} comme l'ensemble des points de sauts de la fonction $R'(\cdot)$.
- Quand on traitera le cas continu, on donnera une caractérisation plus explicite de \mathcal{K} (voir la proposition 10). On se servira d'ailleurs de la proposition 3 et il sera utile d'y avoir considéré des cellules à bord relevé.
- Le fait de faire varier k est en relation avec les liens qui unissent \mathcal{K} et le comportement asymptotique des valeurs propres de L_β . Mais il est possible également de donner une démonstration statique de la proposition 3. On fixe $k > 0$ et on construit "à la main" une partition en k -cslbr1 ou une partition en k -cslbr2. C'est la première démonstration que l'on avait donné de ce résultat, mais elle était extrêmement fastidieuse.

3. Première étape de la démonstration

Désormais, $k > 0$ est fixé de telle manière que $k \notin \{k_1, \dots, k_T\}$ et $M = \bigsqcup_{1 \leq i \leq R} A_i$ est une partition de M en k -cellules. On notera

$$\delta = \min \left(\min_{1 \leq i \leq R} (k - c(A_i)); \min_{1 \leq i \leq R} (h(A_i) - k) \right) > 0.$$

Sans perte de généralité, on supposera aussi que $f(M) = 0$.

On va montrer dans cette section, par une récurrence sur S , le cardinal de l'ensemble $\{f(A_i); 1 \leq i \leq R\}$, que la trajectoire d'un k -algorithme finit presque sûrement par rester dans l'une des k -cellules A_i .

On note $\{f(A_i); 1 \leq i \leq R\} = \{f_1, \dots, f_S\}$, de telle sorte que $1 \leq i < j \leq S \Rightarrow f_i > f_j$, et pour $1 \leq i \leq S$, on pose également

$$I_i = \{1 \leq j \leq R \mid f(A_j) = f_i\} \quad \text{et} \quad B_i = \bigsqcup_{j \in I_i} A_j.$$

Pour amorcer la récurrence, on va d'abord voir que la trajectoire d'un k -algorithme finit presque sûrement, soit par rester au dehors de B_1 , soit par ne plus quitter l'un des A_j , pour $j \in I_1$.

Pour cela, on étudie le comportement en temps grand de la fonctionnelle

$$F_p(t) = \sum_{x \in M} \left(\frac{m_t}{\mu_{\beta_t}}(x) \right)^p \mu_{\beta_t}(x)$$

où p est un réel fixé strictement plus grand que 1.

PROPOSITION 4. — *Il existe une constante $K(p) > 0$ telle que pour tout $t \geq 0$, on ait*

$$F_p(t) \leq K(p) \exp((p-1)f_1\beta_t).$$

Démonstration. — Pour simplifier l'écriture, on note pour $t \geq 0$ et $x \in M$, $g_t(x) = m_t/\mu_{\beta_t}(x)$. Il est bien connu que pour $t \geq 0$, $m_t(x)$ satisfait l'équation de Chapman-Kolmogorov,

$$\begin{aligned} \frac{dm_t(x)}{dt} &= L_{\beta_t}^*(m_t)(x) \\ &= \sum_{y \in M} [m_t(y)q_{\beta_t}(y, x) - m_t(x)q_{\beta_t}(x, y)] ; \end{aligned}$$

ainsi, $F_p(t)$ est dérivable sur $[0, +\infty[$ et

$$\begin{aligned} F'_p(t) &= p \sum_{x \in M} g_t^{p-1}(x) \frac{dm_t(x)}{dt} + (1-p) \sum_{x \in M} g_t^p(x) \frac{d\mu_{\beta_t}(x)}{dt} \\ &= p \sum_{x \in M} L_{\beta_t}(g_t^{p-1})(x) m_t(x) + \\ &\quad + (p-1) \frac{d\beta_t}{dt} \sum_{x \in M} g_t^p(x) (U(x) - \langle U \rangle_{\beta_t}) \mu_{\beta_t}(x) \end{aligned}$$

où $\langle U \rangle_{\beta_t} = \sum_{x \in M} U(x) \mu_{\beta_t}(x)$.

Cependant,

$$\begin{aligned} \sum_{x \in M} L_{\beta_t}(g_t^{p-1})(x) m_t(x) &= \\ &= \sum_{x \in M} \sum_{y \in M} (g_t^{p-1}(y) - g_t^{p-1}(x)) q_{\beta_t}(x, y) m_t(x) \\ &= \sum_{x \in M} \sum_{y \in M} (g_t^{p-1}(y) - g_t^{p-1}(x)) g_t(x) q_{\beta_t}(x, y) \mu_{\beta_t}(x) \\ &= -\frac{1}{2} \sum_{x, y \in M} (g_t^{p-1}(x) - g_t^{p-1}(y)) (g_t(x) - g_t(y)) q_{\beta_t}(x, y) \mu_{\beta_t}(x) \end{aligned}$$

(par symétrie en x, y de $q_{\beta_t}(x, y)\mu_{\beta_t}(x)$)

$$\begin{aligned} &\leq -\frac{1}{2} \sum_{i=1}^R \sum_{x, y \in A_i} (g_t^{p-1}(x) - g_t^{p-1}(y)) (g_t(x) - g_t(y)) q_{\beta_t}(x, y) \mu_{\beta_t}(x) \\ &\leq -2(p-1)/p^2 \sum_{i=1}^R \sum_{x, y \in A_i} (g_t^{p/2}(x) - g_t^{p/2}(y))^2 q_{\beta_t}(x, y) \mu_{\beta_t}(x) \end{aligned}$$

car il est facile de vérifier (voir, par exemple, [2]) que pour tous réels X et Y , on a

$$(X^{p-1} - Y^{p-1})(X - Y) \geq 4 \frac{(p-1)}{p^2} (X^{p/2} - Y^{p/2})^2.$$

Pour évaluer le terme

$$\sum_{x, y \in A_i} (g_t^{p/2}(x) - g_t^{p/2}(y))^2 q_{\beta_t}(x, y) \mu_{\beta_t}(x)$$

pour un $1 \leq i \leq R$ fixé, on va s'inspirer d'un calcul de Holley et Stroock [10, p. 562], pour ce type d'inégalité, dite inégalité géométrique de Poincaré, on renvoie aussi à Götze [7] et aux références qu'il donne).

Soit $x_i \in N_{A_i}$ et pour $x \in A_i$, notons $p_{x_i, x} = (p_{x_i, x}(j))_{1 \leq j \leq n_x}$ un chemin allant de x_i à x dans A_i tel que $e(p_{x_i, x}) \leq U(x) + k - \delta$.

Pour $z, w \in A_i$, on pose

$$\chi_{x_i, x}(z, w) = \begin{cases} 1 & \text{s'il existe } 1 \leq j < n_x \text{ tel que } z = p_{x_i, x}(j) \\ & \text{et } w = p_{x_i, x}(j+1) \\ 0 & \text{sinon.} \end{cases}$$

Alors,

$$\begin{aligned} \sum_{x \in A_i} g_t^p \mu_{\beta_t}(x) &= \\ &= \sum_{x \in A_i} (g_t^{p/2}(x) - g_t^{p/2}(x_i) + g_t^{p/2}(x_i))^2 \mu_{\beta_t}(x) \\ &\leq 2 \sum_{x \in A_i} (g_t^{p/2}(x) - g_t^{p/2}(x_i))^2 + 2g_t^p(x_i) \mu_{\beta_t}(A_i) \end{aligned}$$

$$\begin{aligned}
&\leq 2 \sum_{x \in A_i} \left[\sum_{j=1}^{n_x-1} g_t^{p/2}(p_{x_i,x}(j)) - g_t^{p/2}(p_{x_i,x}(j+1)) \right]^2 \mu_{\beta_t}(x) + \\
&\quad + 2g_t^p(x_i) \mu_{\beta_t}(A_i) \\
&\leq 2 \sum_{x \in A_i} (n_x - 1) \sum_{j=1}^{n_x-1} \left[g_t^{p/2}(p_{x_i,x}(j)) - g_t^{p/2}(p_{x_i,x}(j+1)) \right]^2 \mu_{\beta_t}(x) + \\
&\quad + 2g_t^p(x_i) \mu_{\beta_t}(A_i) \\
&\leq 2N \sum_{x \in A_i} \sum_{j=1}^{n_x-1} \left[g_t^{p/2}(p_{x_i,x}(j)) - g_t^{p/2}(p_{x_i,x}(j+1)) \right]^2 \mu_{\beta_t}(x) + \\
&\quad + 2g_t^p(x_i) \mu_{\beta_t}(A_i) \\
&= 2N \sum_{x \in A_i} \sum_{z, w \in A_i} (g_t^{p/2}(z) - g_t^{p/2}(w))^2 \chi_{x_i,x}(z, w) \mu_{\beta_t}(x) + \\
&\quad + 2g_t^p(x_i) \mu_{\beta_t}(A_i) \\
&\leq 2N \sum_{z, w \in A_i} (g_t^{p/2}(z) - g_t^{p/2}(w))^2 \left(\sum_{x \in A_i} \frac{\chi_{x_i,x}(z, w) \mu_{\beta_t}(x)}{q_{\beta_t}(z, w) \mu_{\beta_t}(z)} \right) \times \\
&\quad \times q_{\beta_t}(z, w) \mu_{\beta_t}(z) + 2g_t^p(x_i) \mu_{\beta_t}(A_i)
\end{aligned}$$

où on a fait la convention que

$$\frac{\chi_{x_i,x}(z, w) \mu_{\beta_t}(x)}{q_{\beta_t}(z, w) \mu_{\beta_t}(z)} = 0 \quad \text{si } \chi_{x_i,x}(z, w) = 0.$$

Cependant, pour un couple (z, w) de A_i ne satisfaisant pas cette condition, on a

$$\begin{aligned}
\frac{\chi_{x_i,x}(z, w) \mu_{\beta_t}(x)}{q_{\beta_t}(z, w) \mu_{\beta_t}(z)} &= \frac{1}{q(z, w) \mu(z)} \exp(-\beta_t [U(x) - U(z) \vee U(w)]) \\
&\leq K_1^{-1} \exp((k - \delta)\beta_t)
\end{aligned}$$

où

$$K_1 = \min_{z, w \in M | q(z, w) > 0} \mu(z) q(z, w).$$

Ainsi,

$$\begin{aligned}
&\sum_{x, y \in A_i} (g_t^{p/2}(x) - g_t^{p/2}(y))^2 q_{\beta_t}(x, y) \mu_{\beta_t}(x) \geq \\
&\geq \frac{K_1}{2N^2} \exp(-(k - \delta)\beta_t) \left[\sum_{x \in A_i} g_t^p(x) \mu_{\beta_t}(x) - 2g_t^p(x_i) \mu_{\beta_t}(A_i) \right].
\end{aligned}$$

Cependant, notons que pour tout $t \geq 0$ et tout $x \in M$, on a $m_t(x) \leq 1$ et qu'il existe une constante $K_2 \geq 0$, telle que pour tout $1 \leq i \leq R$ et tout $t \geq 0$, on ait

$$\begin{aligned} \frac{1}{\mu_{\beta_t}^p(x_i)} \mu_{\beta_t}(A_i) &\leq K_2 \exp((p-1)f(A_i)\beta_t) \\ &\leq K_2 \exp((p-1)f_1\beta_t). \end{aligned}$$

En utilisant le fait que pour $t \geq 2$, l'inverse de la température est donné par $\beta_t = k^{-1} \ln(t)$, on a donc obtenu que pour $t \geq 2$

$$\begin{aligned} p \sum_{x \in M} L_{\beta_t}(g_t^{p-1})(x) m_t(x) &\leq \\ &\leq -\frac{(p-1)}{pN^2} K_1 t^{-(k-\delta)/k} \left[\sum_{x \in M} g_t^p(x) \mu_{\beta_t}(x) - 2RK_2 t^{k^{-1}(p-1)f_1} \right]. \end{aligned}$$

D'autre part, il est clair qu'en prenant $K_3(p) = k^{-1}(p-1)(\max_{x \in M} U(x) - \min_{x \in M} U(x))$, on a pour tout $t \geq 2$

$$\left| (1-p) \frac{d\beta_t}{dt} \sum_{x \in M} g_t^p(x) (U(x) - \langle U \rangle_{\beta_t}) \mu_{\beta_t}(x) \right| \leq K_3(p) \frac{1}{t} \sum_{x \in M} g_t^p(x) \mu_{\beta_t}(x).$$

La fonctionnelle F_p satisfait donc pour $t \geq 2$ l'inégalité différentielle

$$F_p'(t) \leq K_5(p) t^{-(k-\delta)/k} t^{k^{-1}(p-1)f_1} - \left[2K_4(p) t^{-(k-\delta)/k} - K_3(p) t^{-1} \right] F_p(t)$$

où $K_4(p) = 2^{-1}(p-1)p^{-1}N^{-2}K_1$ et $K_5(p) = 4RK_2K_4(p)$.

Or, il est clair qu'il existe $t_1 \geq 2$ tel que pour tout $t \geq t_1$ on ait

$$2K_4(p) t^{-(k-\delta)/k} - K_3(p) t^{-1} \geq K_4(p) t^{-(k-\delta)/k};$$

ainsi, pour $t \geq t_1$, la fonctionnelle F_p vérifie l'inégalité différentielle

$$F_p'(t) \leq K_5(p) t^{\alpha_1} - K_4(p) t^{\alpha_2} F_p(t)$$

où

$$\alpha_1 = \alpha_2 + k^{-1}(p-1)f_1, \quad \alpha_2 = -\frac{k-\delta}{k}.$$

L'estimée annoncée dans la proposition (qu'il suffit de prouver pour t assez grand) découle alors immédiatement du résultat classique suivant.

LEMME 5. — Soit F une fonction positive et dérivable sur $[t_1, +\infty[$ (où $t_1 > 0$) satisfaisant l'inégalité différentielle

$$F'(t) \leq At^a - Bt^b F(t)$$

où $A, B > 0$ et $a > b > -1$.

Il existe alors une constante $K > 0$ (dépendant de A, B, a, b et $F(t_1)$) telle que pour tout $t \geq t_1$ on ait

$$F(t) \leq Kt^{a-b}$$

Démonstration. — On considère la fonction G définie sur $[t_1, +\infty[$ par $G(t) = F(t) \exp(B(b+1)^{-1}t^{b+1})$. On a

$$\begin{aligned} G'(t) &= [F'(t) + Bt^b F(t)] \exp(B(b+1)^{-1}t^{b+1}) \\ &\leq At^a \exp(B(b+1)^{-1}t^{b+1}) \end{aligned}$$

d'où

$$\begin{aligned} G(t) &\leq G(t_1) + A \int_{t_1}^t s^a \exp(B(b+1)^{-1}s^{b+1}) ds \\ &\leq G(t_1) + At^{a-b} \int_{t_1}^t s^b \exp(B(b+1)^{-1}s^{b+1}) ds \\ &\leq G(t_1) + AB^{-1}t^{a-b} \exp(B(b+1)^{-1}t^{b+1}) \end{aligned}$$

puis

$$F(t) \leq G(t_1) \exp(-B(b+1)^{-1}t^{b+1}) + AB^{-1}t^{a-b}$$

ce qui peut encore se récrire sous la forme énoncée dans le lemme. \square

Remarque. — Notons que dans la démonstration précédente, on n'a eu besoin que de l'hypothèse $\min_{1 \leq i \leq R} (k - c(A_i)) > 0$, c'est-à-dire que la proposition est encore satisfaite si on ne considère qu'une partition en k -cellules semi-larges du second type. Soit $k > \tilde{c}$ et $M = \bigsqcup_{1 \leq i \leq R'} A'_i$ une telle partition (qui existe toujours, sans l'hypothèse $k \notin \{k_1, \dots, k_T\}$). On a alors nécessairement pour tout $1 \leq i \leq R'$, $f(A'_i) = 0$ (i.e. $S = 1$ et $f_1 = 0$) et la proposition précédente permet alors de voir que la loi du processus tend à se concentrer sur l'ensemble des minima globaux de U sur M , car si

$p > 1$ est fixé et si on note q son exposant conjugué, on a

$$\begin{aligned} m_t(N_M^c) &= \sum_{x \notin N_M} f_t(x) \mu_{\beta_t}(x) \\ &\leq \left(\sum_{x \in M} f_t^p(x) \mu_{\beta_t}(x) \right)^{\frac{1}{p}} \left(\sum_{x \notin N_M} \mu_{\beta_t}(x) \right)^{\frac{1}{q}} \\ &\leq K(p)^{1/p} \mu_{\beta_t}^{1/q}(N_M^c) \xrightarrow{t \rightarrow +\infty} 0 ; \end{aligned}$$

d'où

$$\lim_{t \rightarrow +\infty} m_t(N_M) = 1 .$$

On va maintenant déduire de la proposition 4 le second résultat annoncé au début de cette section sur le comportement des trajectoires. Rappelons que le processus $(X_t)_{t \geq 0}$ est la réalisation canonique d'un k -algorithme.

PROPOSITION 6. — *Soit $j \in I_1$. On a alors l'existence p.s. de*

$$\lim_{t \rightarrow +\infty} 1_{A_j}(X_t) ,$$

c'est-à-dire que X_t finit p.s., soit par rester dans A_j , soit par ne plus y entrer.

Démonstration. — On s'inspire de la démonstration du lemme 3.2 de [9].

Fixons $j \in I_1$ et notons $\phi(\cdot) = 1_{A_j}(\cdot)$. Le processus $(\phi(X_t))_{t \geq 0}$ admet pour décomposition canonique

$$\phi(X_t) = \phi(X_0) + M_t + \int_0^t L_{\beta_s} \phi(X_s) ds$$

où $(M_t)_{t \geq 0}$ est une martingale (sous-entendu une $(\mathcal{F}_t)_{t \geq 0}$ -martingale).

Rappelons que le crochet oblique de cette martingale est

$$\langle M \rangle_t = \int_0^t \Gamma_{\beta_s}(\phi, \phi)(X_s) ds$$

où $\Gamma_{\beta_s}(\phi, \phi) = L_{\beta_s} \phi^2 - 2\phi L_{\beta_s} \phi$ est le carré du champ associé à L_{β_s} .

En effet,

$$\begin{aligned} (M_t - \phi(X_0))^2 &= \left(\phi(X_t) - \int_0^t L_{\beta_s} \phi(X_s) ds \right)^2 \\ &= \phi(X_t)^2 - 2\phi(X_t) \left(\int_0^t L_{\beta_s} \phi(X_s) ds \right) + \\ &\quad + \left(\int_0^t L_{\beta_s} \phi(X_s) ds \right)^2. \end{aligned}$$

Cependant, on a

$$\phi^2(X_t) = \phi^2(X_0) + \widetilde{M}_t + \int_0^t L_{\beta_s} \phi^2(X_s) ds$$

où $(\widetilde{M}_t)_{t \geq 0}$ est une martingale. De plus,

$$\left(\int_0^t L_{\beta_s} \phi(X_s) ds \right)^2 = 2 \int_0^t \left(\int_0^s L_{\beta_u} \phi(X_u) du \right) L_{\beta_s} \phi(X_s) ds$$

et

$$\begin{aligned} \phi(X_t) \left(\int_0^t L_{\beta_s} \phi(X_s) ds \right) &= \int_0^t \left(\int_0^s L_{\beta_u} \phi(X_u) du \right) dM_s + \\ &\quad + \int_0^t \left(\int_0^s L_{\beta_u} \phi(X_u) du \right) L_{\beta_s} \phi(X_s) ds + \\ &\quad + \int_0^t \phi(X_s) L_{\beta_s} \phi(X_s) ds. \end{aligned}$$

Remarquons que

$$t \mapsto \int_0^t \left(\int_0^s L_{\beta_u} \phi(X_u) du \right) dM_s \quad \text{et} \quad t \mapsto \phi(X_0) M_t$$

sont des martingales, ce qui permet de voir qu'il existe une martingale $(\widehat{M}_t)_{t \geq 0}$ telle que

$$M_t^2 = \widehat{M}_t + \int_0^t [L_{\beta_s} \phi^2(X_s) - 2\phi L_{\beta_s} \phi(X_s)] ds ;$$

d'où le résultat annoncé.

On calcule que pour $x \in M$,

$$\begin{aligned}\Gamma_{\beta_s}(\phi, \phi)(x) &= \sum_{y \in M} (\phi(y) - \phi(x))^2 q_{\beta_s}(x, y) \\ &= \sum_{y \mid (x, y) \in \tilde{\partial} A_j} q_{\beta_s}(x, y)\end{aligned}$$

où $\tilde{\partial} A_j = \{(x, y) \in M \times M \mid (y, x) \in \partial A_j \text{ ou } (x, y) \in \partial A_j\}$. Ainsi, pour $s \geq 0$,

$$\begin{aligned}\mathbb{E}[\Gamma_{\beta_s}(\phi, \phi)(X_s)] &= \\ &= \sum_{(x, y) \in \tilde{\partial} A_j} q_{\beta_s}(x, y) m_s(x) = \sum_{(x, y) \in \tilde{\partial} A_j} g_s(x) q_{\beta_s}(x, y) \mu_{\beta_s}(x) \\ &\leq \sum_{y \in M} \left(\sum_{x \mid (x, y) \in \tilde{\partial} A_j} g_s^p(x) \mu_{\beta_s}(x) \right)^{\frac{1}{p}} \left(\sum_{x \mid (x, y) \in \tilde{\partial} A_j} q_{\beta_s}^q(x, y) \mu_{\beta_s}(x) \right)^{\frac{1}{q}}\end{aligned}$$

où $p > 1$, $q > 1$ satisfont

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Cependant, il existe une constante $K_1 > 0$ telle que pour tout $(x, y) \in \tilde{\partial} A_j$ on ait

$$\begin{aligned}\mu_{\beta_s}(x) q_{\beta_s}(x, y) &= Z_{\beta_s}^{-1} \exp(-\beta_s [U(x) \vee U(y)]) \\ &\leq K_1 \exp(-\beta_s (f_1 + k + \delta))\end{aligned}$$

et

$$\mu_{\beta_s}(x) = Z_{\beta_s}^{-1} \exp(-\beta_s U(x)) \leq K_1 \exp(\beta_s \|U\|_\infty)$$

d'où

$$\begin{aligned}\sum_{x \mid (x, y) \in \tilde{\partial} A_j} q_{\beta_s}^q(x, y) \mu_{\beta_s}(x) &= \\ &= \sum_{x \mid (x, y) \in \tilde{\partial} A_j} [\mu_{\beta_s}(x) q_{\beta_s}(x, y)]^q \mu_{\beta_s}^{1-q}(x) \\ &\leq K_1^2 N \exp(-q\beta_s (f_1 + k + \delta)) \exp((1-q)\beta_s \|U\|_\infty).\end{aligned}$$

Choisissons $q > 1$ suffisamment proche de 1 pour que

$$q(f_1 + k + \delta) - (q-1)\|U\|_\infty \geq f_1 + k + \frac{\delta}{2}$$

et soit p l'exposant conjugué. La proposition 4 permet alors de voir que pour $s \geq 2$,

$$\begin{aligned} \mathbb{E}[\Gamma_{\beta_s}(\phi, \phi)(X_s)] &\leq \\ &\leq \sum_{y \in M} K_1^2 N(K(p))^{1/p} \exp\left(\beta_s f_1 \frac{p-1}{p}\right) \exp\left(-\beta_s \left(f_1 + k + \frac{\delta}{2}\right)\right) \\ &\leq K_1^2 N^2(K(p))^{1/p} \exp\left(-\beta_s \left(k + \frac{\delta}{2}\right)\right) \\ &= K_1^2 N^2(K(p))^{1/p} s^{-\frac{2k+\delta}{2k}}. \end{aligned}$$

On a donc pour $t \geq 2$

$$\begin{aligned} \mathbb{E}[\langle M \rangle_t] &= \int_0^t \mathbb{E}[\Gamma_{\beta_s}(\phi, \phi)(X_s)] ds \\ &\leq K_1^2 N^2(K(p))^{1/p} \left[\int_0^2 \exp\left(-\beta_s \left(k + \frac{\delta}{2}\right)\right) ds + \int_2^t s^{-\frac{2k+\delta}{2k}} ds \right] \\ &\leq K_1^2 N^2(K(p))^{1/p} \left[\int_0^2 \exp\left(-\beta_s \left(k + \frac{\delta}{2}\right)\right) ds + \frac{2k}{\delta} \right]. \end{aligned}$$

Ceci nous permet de constater que

$$\sup_{t \geq 0} \mathbb{E}[|M_t|] \leq \sup_{t \geq 0} \sqrt{\mathbb{E}[M_t^2]} \leq \sup_{t \geq 0} \sqrt{\mathbb{E}[\langle M \rangle_t]} < +\infty ;$$

ce qui d'après un théorème classique, implique que la martingale M_t converge en temps grand.

D'autre part, pour $s \geq 0$,

$$\begin{aligned} \mathbb{E}[|L_{\beta_s} \phi(X_s)|] &= \sum_{x \in M} \left| \sum_{y \in M} (\phi(y) - \phi(x)) q_{\beta_s}(x, y) \right| m_s(x) \\ &\leq \sum_{x, y \in M} |\phi(y) - \phi(x)| q_{\beta_s}(x, y) m_s(x) \\ &= \mathbb{E}[\Gamma_{\beta_s}(\phi, \phi)(X_s)] \end{aligned}$$

car ϕ ne prend que les valeurs zéro et un. Les calculs précédents montrent alors que

$$\mathbb{E} \left[\int_0^\infty |L_{\beta_s} \phi(X_s)| ds \right] < +\infty ,$$

c'est-à-dire que p.s.,

$$\int_0^\infty L_{\beta_s} \phi(X_s) ds$$

est absolument convergente, d'où finalement la convergence p.s. de $\phi(X_t)$ en temps grand. \square

Notons que les calculs précédents montrent également que pour $j \in I_1$, la quantité $m_t(A_j)$ converge en temps grand. En effet, en reprenant les notations de la démonstration, on a

$$\frac{d}{dt} m_t(A_j) = \mathbb{E}[L_{\beta_t} \phi(X_t)]$$

d'où

$$\int_0^\infty \left| \frac{d}{dt} m_t(A_j) \right| dt \leq \int_0^\infty \mathbb{E}[|L_{\beta_t} \phi(X_t)|] dt < +\infty$$

puis la convergence annoncée. Mais ce résultat peut aussi se déduire directement de la proposition, en remarquant que pour $j \in I_1$,

$$\lim_{t \rightarrow +\infty} m_t(A_j) = \sup_{t \geq 0} P_{m,\beta}[\forall s \geq t, X_s \in A_j].$$

Prouvons maintenant que la proposition 6 est en fait satisfaite par toutes les k -cellules de la partition.

PROPOSITION 7. — *La trajectoire d'un k -algorithme finit presque sûrement par rester dans l'une des k -cellules de la partition.*

Démonstration. — Si $S = 1$, il s'agit là d'une réécriture de la proposition 6. Remarquons que cette proposition est aussi satisfaite si au lieu de s'intéresser à un k -algorithme, on considère un algorithme du recuit simulé tel que pour tout $t \geq 2$, on ait $\beta_t = k \ln(A + t)$, où $A \geq 0$. Dans la suite de cette démonstration, on appellera encore k -algorithme un algorithme du recuit simulé satisfaisant cette propriété.

Supposons que $S > 1$ et que la proposition 7 est satisfaite par tout k -algorithme sur un graphe (irréductible et réversible) M' qui admet une partition en k -cellules dont le cardinal de l'ensemble des fonds est $S' < S$. On va montrer que la proposition 7 est satisfaite par tout k -algorithme sur un graphe (irréductible et réversible) M qui admet une partition en k -cellules $M = \bigsqcup_{1 \leq i \leq R} A_i$ dont le cardinal de l'ensemble des fonds est S .

Soit $\varepsilon > 0$. D'après la proposition 6, il existe $T > 0$ tel que la probabilité (sous $P_{m,\beta}$) qu'une trajectoire sorte ou entre après l'instant T d'une k -cellule A_j , pour $j \in I_1$, soit inférieure à ε .

Pour $x \neq y \in M$, soit

$$\tilde{q}(x, y) = \begin{cases} q(x, y) & \text{s'il n'existe pas de } j \in I_1 \text{ tel que } (x, y) \in \tilde{\partial}A_j, \\ 0 & \text{sinon} \end{cases}$$

et pour $x \in M$, notons

$$\tilde{q}(x, x) = 1 - \sum_{y \neq x} \tilde{q}(x, y).$$

La matrice \tilde{q} est un noyau de probabilités de transitions (qui n'est pas irréductible) et comme dans l'introduction, on considère pour $\beta \geq 0$, le noyau \tilde{q}_β défini par

$$\forall x, y \in M, \quad \tilde{q}_\beta(x, y) = \begin{cases} \exp\left(-\beta(U(y) - U(x))_+\right) \tilde{q}(x, y) & \text{si } y \neq x \\ 1 - \sum_{z \neq x} \tilde{q}_\beta(x, z) & \text{si } y = x \end{cases}$$

et l'opérateur \tilde{L}_β agissant sur les fonctions réelles ϕ définies sur M , par

$$\forall x \in M, \quad \tilde{L}_\beta \phi(x) = \sum_{y \in M} (\phi(y) - \phi(x)) \tilde{q}_\beta(x, y).$$

Soit Q^T la probabilité sur $\Omega^T = D([T, +\infty[; M)$ (muni de la tribu \mathcal{F}^T engendrée par les applications coordonnées) qui admet m_T pour probabilité "initiale" à l'instant T et qui est solution du problème de martingales associé à la famille de générateurs $(\tilde{L}_{\beta_t})_{t \geq T}$. Notons également P^T la projection de la probabilité $P_{m,\beta}$ sur $(\Omega^T, \mathcal{F}^T)$ (par l'application canonique $(X_t)_{t \geq 0} \mapsto (X_t)_{t \geq T}$).

Soit \mathcal{D}^T l'ensemble des trajectoires de Ω^T qui, si $X_T \in A_j$ pour un $j \in I_1$, restent ensuite dans A_j et qui si $X_T \notin B_1$ restent ensuite en dehors de B_1 . Par choix de T , on a

$$P^T(\mathcal{D}^T) \geq 1 - \varepsilon. \quad (*)$$

D'autre part, on voit facilement (en considérant par exemple la loi de la suite des temps de sauts successifs des trajectoires sous P^T et sous Q^T) que pour tout borélien $A \in \mathcal{F}^T$, on a

$$P^T(A \cap \mathcal{D}^T) \leq Q^T(A).$$

Remarquons que $Q^T(\mathcal{D}^T) = 1$ et plus précisément que l'on peut construire la probabilité Q^T de la manière suivante. Notons M'_1, \dots, M'_r les composantes connexes de $M \setminus B_1$ pour \tilde{q} (ou q). $\mathcal{Y} = \{M'_1, \dots, M'_r, A_j; j \in I_1\}$ est alors l'ensemble des composantes connexes de M pour \tilde{q} . Pour $Y \in \mathcal{Y}$, soient $\Omega^T(Y) = D([T, +\infty[; Y)$, $m_{T,Y}$ la probabilité sur Y définie par

$$\forall y \in Y, \quad m_{T,Y}(y) = \frac{m_T(y)}{m_T(Y)}$$

et pour $\beta \geq 0$, $\tilde{L}_{\beta,Y}$ la restriction de l'opérateur \tilde{L}_β aux fonctions définies sur Y (ce qui a bien un sens, car pour $y \in Y$, $\tilde{L}_\beta(\phi)(y)$ ne dépend que des valeurs de ϕ sur Y). Notons que $\tilde{L}_{\beta,Y}$ est symétrique pour la mesure $\mu_{Y,\beta}$ définie dans l'introduction (c'est-à-dire que la restriction de \tilde{q}_β à $Y \times Y$ est réversible par rapport à $\mu_{Y,\beta}$). Appelons $Q^{T,Y}$ la probabilité sur $\Omega^T(Y)$ qui admet pour probabilité initiale (à l'instant T) $m_{T,Y}$ et qui est solution du problème de martingales associé à la famille de générateurs $(\tilde{L}_{\beta_t,Y})_{t \geq T}$. On note encore $Q^{T,Y}$ la probabilité obtenue sur Ω^T à partir de $Q^{T,Y}$ par identification de $\Omega^T(Y)$ comme un sous-ensemble mesurable de Ω^T .

Il est alors immédiat de vérifier que

$$Q^T(\cdot) = \sum_{Y \in \mathcal{Y}} m_T(Y) Q^{T,Y}(\cdot).$$

Soit $1 \leq i \leq r$ fixé. On note $A'_1, \dots, A'_{R'}$ les k -cellules (de $\{A_j; j \notin I_1\}$) qui sont incluses dans M'_i . Puisque les k -cellules A_1, \dots, A_R sont aussi connexes pour \tilde{q} , $M'_i = \bigsqcup_{1 \leq j \leq R'} A'_j$ est une décomposition de M'_i en k -cellules, et on peut appliquer l'hypothèse de récurrence pour obtenir que sous $Q^{T,Y}$, une trajectoire de $\Omega^T(Y)$ finit par rester dans l'une des k -cellules A'_j pour $1 \leq j \leq R'$. On en déduit que sous Q^T , une trajectoire finit par rester dans l'une des k -cellules A_i , pour $1 \leq i \leq R$, puis qu'il en est de même sous P^T (et sous $P_{m,\beta}$), pour une trajectoire de \mathcal{D}^T . Le résultat annoncé découle alors de (*) et du fait que l'on peut choisir ε arbitrairement petit. \square

4. Seconde étape de la démonstration

Le résultat de la section précédente permet d'étudier ce qui se passe en temps grand dans une k -cellule de la partition presque indépendamment de ce qui se passe dans les autres, et le théorème 1 va s'en déduire facilement.

Commençons par nous intéresser aux masses des k -cellules sous m_t . Pour un $1 \leq j \leq R$ fixé, la proposition 7 permet de voir que la quantité $m_t(A_j)$ converge en temps grand et que la limite vaut

$$\lim_{t \rightarrow +\infty} m_t(A_j) = \sup_{t \geq 0} P_{m, \beta}[\forall s \geq t, X_s \in A_j].$$

Pour prouver que cette limite est strictement positive, on va reprendre une démonstration due à Holley, Kusuoka et Stroock (voir la troisième section de [9]).

PROPOSITION 8. — Pour $1 \leq j \leq R$, notons

$$\lambda_j = \lim_{t \rightarrow +\infty} m_t(A_j).$$

On a $\lambda_j > 0$.

Démonstration. — Soit $1 \leq j \leq R$ fixé. On considère le potentiel V sur M défini par :

$$V(x) = \begin{cases} U(x) & \text{si } x \in A_j \\ \|U\|_\infty & \text{sinon.} \end{cases}$$

Il est clair que M est une k -cellule pour ce potentiel.

Appelons Q la probabilité sur (Ω, \mathcal{F}) associée, comme dans l'introduction, à ce potentiel, à la mesure initiale m et à l'évolution β , et pour $t \geq 0$, notons n_t la loi de X_t sous Q .

On applique la proposition 4 (et la remarque qui la suit, après avoir noté que l'ensemble des minima globaux de V est N_{A_j}) à la famille des $(n_t)_{t \geq 0}$ pour obtenir que

$$\lim_{t \rightarrow +\infty} n_t(N_{A_j}) = 1.$$

Cependant, du fait que pour tout couple $(x, y) \in \partial A_j$, on a

$$V(x) \vee V(y) \geq \min_{z \in M} V(z) + k + \delta$$

il est encore vrai que sous Q , $1_{A_j}(X_t)$ converge p.s. en temps grand (voir la démonstration de la proposition 6) et l'égalité ci-dessus montre alors que la limite est 1 Q -p.s.

Pour $x \in M$ et $s > 0$, notons $u(x, s)$ la probabilité conditionnelle

$$u(x, s) = Q[\forall t \geq s, X_t \in A_j \mid X_s = x]$$

(remarquons que cette quantité est définie partout, car pour tout $x \in M$ et tout $s > 0$, on a $Q(X_s = x) = n_s(x) > 0$ (par exemple, [14])).

Pour $x \notin A_j$ et $s > 0$, on a $u(x, s) = 0$, et il est bien connu que cette fonction satisfait sur $A_j \times]0, +\infty[$ l'équation

$$\frac{\partial}{\partial s} u(x, s) = -L_{\beta_s}^V(u(\cdot, s))(x) \quad (**)$$

où $L_{\beta_s}^V$ est le générateur du processus $(X_t)_{t \geq 0}$ sous Q .

En effet, en notant $(\tau_s(n))_{n \geq 1}$ la suite des temps de sauts après l'instant $s \geq 0$, il existe, du fait que les taux de sauts sont bornés uniformément, une constante $K > 0$ telle que pour tout $s \geq 0$, tout $x \in M$ et tout $h > 0$ on ait

$$Q[\tau_s(1) \leq h \mid X_s = x] \leq Kh;$$

il en découle que

$$\lim_{h \rightarrow 0^+} \sup_{\substack{s \geq 0 \\ x \in M}} h^{-1} Q[\tau_s(2) \leq s + h \mid X_s = x] = 0;$$

d'où pour $x \in A_j$,

$$\lim_{h \rightarrow 0^+} \sup_{\substack{s \geq 0 \\ x \in A_j}} h^{-1} (Q[\forall t \geq s + h, X_t \in A_j \mid X_s = x] - u(x, s)) = 0.$$

Or, pour $x \in A_j$, $s > 0$ et $h > 0$ suffisamment petit, on a

$$\begin{aligned} u(x, s) - u(x, s - h) &= \\ &= u(x, s) - Q[\forall t \geq s, X_t \in A_j \mid X_{s-h} = x] + \\ &\quad + Q[\forall t \geq s, X_t \in A_j \mid X_{s-h} = x] - u(x, s - h) \\ &= \mathbb{E}_Q[u(x, s) - u(X_s, s) \mid X_{s-h} = x] + \\ &\quad + Q[\forall t \geq s, X_t \in A_j \mid X_{s-h} = x] - u(x, s - h) \\ &= \int_{s-h}^s \mathbb{E}_Q[L_{\beta_r}^V u(X_r, s) \mid X_{s-h} = x] dr + \\ &\quad + Q[\forall t \geq s, X_t \in A_j \mid X_{s-h} = x] - u(x, s - h); \end{aligned}$$

d'où

$$\lim_{h \rightarrow 0^+} h^{-1} [u(x, s) - u(x, s - h)] = L_{\beta_s}^V u(\cdot, s)(x)$$

et on déduit aisément le résultat annoncé.

Cependant, l'opérateur spatio-temporel $\partial/\partial s + L_{\beta}^V$ satisfait un principe du maximum sur $A_j \times]0, +\infty[$: si une fonction u satisfait (**) et admet un minima global sur cet ensemble, alors elle y est constante. En effet, en un tel minima global (x_0, s_0) $\partial/\partial s u$ s'annule, d'où $L_{\beta_{s_0}}^V(u(\cdot, s_0))(x_0) = 0$. Mais ceci et le fait que $A_j \ni x \mapsto u(x, s_0)$ admet un minimum global en x_0 permet de voir que tous les plus proches voisins de x_0 sont également des minima globaux pour cette fonction, et par suite pour l'application u sur $A_j \times]0, +\infty[$. En répétant ce raisonnement un nombre fini de fois, on obtient, grâce à l'irréductibilité du graphe, que pour tout $x \in A_j$, $u(x, s_0) = u(x_0, s_0)$. L'unicité de la solution de l'équation différentielle ordinaire (**), étant donné $u(\cdot, s_0)$, implique alors que $u \equiv u(x_0, s_0)$ sur $A_j \times]0, +\infty[$.

Notamment, s'il existe un point $(x_0, s_0) \in A_j \times]0, +\infty[$ tel que $u(x_0, s_0) = 0$, il apparaît, par positivité de u (qui est à nouveau la fonction définie précédemment), que $u \equiv 0$ sur $A_j \times]0, +\infty[$ (et donc en fait sur $M \times]0, +\infty[$). Supposons que ceci soit vérifié. Du fait que X_t finit par rester dans A_j , on devrait avoir

$$1 = \lim_{s \rightarrow +\infty} Q[\forall t \geq s, X_t \in A_j] = \lim_{s \rightarrow +\infty} \mathbb{E}_Q[u(X_s, s)] = 0 ;$$

ce qui nous fournit une contradiction. On a donc pour tout $x \in A_j$ et tout $s > 0$, $u(x, s) > 0$. Or le fait qu'une particule issue de $x \in A_j$ au temps $s = 1$ puisse rester ensuite tout le temps dans A_j avec une probabilité strictement positive sous $Q[\cdot \mid X_1 = x]$ montre qu'il en est de même sous $P_{m, \beta}[\cdot \mid X_1 = x]$ car tant qu'une particule reste dans A_j , elle a les mêmes probabilités de transitions sous $Q[\cdot \mid X_1 = x]$ et sous $P_{m, \beta}[\cdot \mid X_1 = x]$. On a donc pour $x \in A_j$,

$$P_{m, \beta}[\forall t \geq 1, X_t \in A_j \mid X_1 = x] > 0$$

d'où, puisque $m_1(A_j) > 0$

$$\alpha_j \stackrel{\text{déf}}{=} P_{m, \beta}[\forall t \geq 1, X_t \in A_j] > 0.$$

Or pour $t \geq 1$, $m_t(A_j) \geq \alpha_j$, et on en déduit en passant à la limite, que $\lambda_j \geq \alpha_j > 0$. \square

Remarque. — Soit $0 < k < \tilde{c}$ quelconque (i.e. sans l'hypothèse $k \notin \{k_1, \dots, k_T\}$). On peut adapter la démonstration précédente pour voir qu'un k -algorithme ne convergera pas totalement vers les minima globaux du potentiel. En effet, soit $M = \bigsqcup_{1 \leq i \leq R'} A'_i$ une partition en k -cellules semi-larges du premier type. Nécessairement, l'une d'elle satisfait $f(A'_i) > 0$. Décomposons $A'_i = \bigsqcup_{1 \leq j \leq R''} A''_j$ en k -cellules semi-larges du second type. On note A'''_1, \dots, A'''_r les A''_j qui sont telles que $f(A''_j) = f(A'_i)$ et on considère l'ensemble $B = \bigcup_{1 \leq j \leq r} A'''_j$. Il est clair que $h(B) > k$. Soit V le potentiel sur M défini par

$$V(x) = \begin{cases} U(x) & \text{si } x \in B \\ \|U\|_\infty & \text{sinon.} \end{cases}$$

On peut rajouter les composantes connexes de $M \setminus B$ à des A'''_j qui leur sont voisins pour obtenir une partition de M en k -cellules semi-larges du second type pour V . Pour cette décomposition, $S = 1$, ainsi la remarque suivant la proposition 4 et la démonstration de la proposition 8 permettent de voir que la trajectoire d'un k -algorithme associé à V qui se trouve au temps 1 en B peut ensuite, avec une probabilité strictement positive, toujours y rester. On en déduit comme dans la démonstration précédente que pour un k -algorithme associé à U on a

$$P_{m,\beta}[\forall t \geq 1, X_t \in B] > 0;$$

ce qui termine de montrer que dans le corollaire 2, l'hypothèse $k \notin \{k_1, \dots, k_T\}$ est inutile.

Reste à montrer la convergence des m_t vers $\sum_{j=1}^R \lambda_j \mu_{A_j, \infty}$. On note $\|\cdot\|$ la variation totale des mesures.

Soit $\varepsilon > 0$. D'après la proposition 7, il existe $T > 0$ tel que la probabilité (sous $P_{m,\beta}$) qu'une trajectoire sorte après l'instant T d'une k -cellule A_j , pour $1 \leq j \leq R$, soit inférieure à ε .

On considère la probabilité Q^T sur $(\Omega^T, \mathcal{F}^T)$ associée, comme dans la proposition 7, au noyau de probabilités de transitions \tilde{q} défini par

$$\tilde{q}(x, y) = \begin{cases} q(x, y) & \text{si } x \neq y \text{ et s'il n'existe pas de} \\ & 1 \leq j \leq R \text{ tel que } (x, y) \in \tilde{\partial} A_j, \\ 0 & \text{s'il existe } 1 \leq j \leq R \text{ tel que} \\ & (x, y) \in \partial A_j, \\ 1 - \sum_{y \neq x} \tilde{q}(x, y) & \text{si } x = y. \end{cases}$$

Notons également P^T la projection de $P_{m,\beta}$ sur $(\Omega^T, \mathcal{F}^T)$. Par choix de T , on a sur $(\Omega^T, \mathcal{F}^T)$

$$\|P^T - Q^T\| \leq \varepsilon. \quad (\dagger)$$

Cependant, on peut écrire

$$Q^T(\cdot) = \sum_{1 \leq j \leq R} m_T(A_j) Q_j^T(\cdot) \quad (\ddagger)$$

où Q_j^T est la loi d'un processus de recuit simulé sur $(A_j, q_{|A_j})$, associé au potentiel $U_{|A_j}$, à l'évolution de température $(\beta_t^{-1})_{t \geq T}$ et admettant $m_T(A_j)^{-1}(m_T)_{|A_j}$ pour probabilité initiale à l'instant T . Notons $(n_{j,t})_{t \geq T}$ la famille des lois (a priori sur A_j , mais que l'on considère aussi comme des probabilités sur M) associée à Q_j^T . Il est bien connu (voir par exemple [14]), du fait que $c(A_i) < k$, que sur M

$$\lim_{t \rightarrow +\infty} \|n_{j,t} - \mu_{A_j, \infty}\| = 0 ;$$

ainsi, d'après (\dagger) et (\ddagger) , on a sur M

$$\overline{\lim}_{t \rightarrow +\infty} \left\| m_t - \sum_{j=1}^R m_T(A_j) \mu_{A_j, \infty} \right\| \leq \varepsilon.$$

Cependant, remarquons cette inégalité implique que

$$\sum_{j=1}^R |\lambda_j - m_T(A_j)| \leq \varepsilon ;$$

d'où

$$\overline{\lim}_{t \rightarrow +\infty} \left\| m_t - \sum_{j=1}^R \lambda_j \mu_{A_j, \infty} \right\| \leq 2\varepsilon$$

puis, ε pouvant être choisi arbitrairement petit,

$$\lim_{t \rightarrow +\infty} \left\| m_t - \sum_{j=1}^R \lambda_j \mu_{A_j, \infty} \right\| = 0.$$

5. Le cas continu

Il est clair que l'idée générale de la méthode présentée ci-dessus est également valable dans le cas continu où M est une variété riemannienne compacte et connexe. Le but de cette section est d'énoncer un résultat similaire au théorème 1 dans cette situation et de présenter les modifications techniques qu'il faut apporter à la démonstration précédente pour le prouver.

Soit M une variété riemannienne compacte et connexe, comme d'habitude, on note $\langle \cdot, \cdot \rangle$, $|\cdot|$, ∇ , div , Δ , λ et d , le produit scalaire, la norme, le gradient, la divergence, le laplacien, la probabilité et la distance associés à la structure riemannienne de M . Soit $U \in C^\infty(M)$ un potentiel régulier. Sans perte de généralité, on supposera que U a été normalisée de telle manière que $\min_{z \in M} U(z) = 0$ et $\max_{z \in M} |\nabla U|(z) \leq 1$. Comme dans le cas discret, les algorithmes du recuit simulé sont décrits à l'aide d'une famille $(L_\beta)_{\beta \geq 0}$ d'opérateurs agissant sur $C^\infty(M)$ par

$$\forall f \in C^\infty(M), \quad L_\beta f = \text{tr } f - \beta \langle \nabla U, \nabla f \rangle.$$

Soient $\Omega = C([0, +\infty[; M)$ l'ensemble des trajectoires continues de $[0, +\infty[$ dans M et $(X_t)_{t \geq 0}$ le processus canonique des coordonnées sur Ω . Pour $t \geq 0$, on note \mathcal{F}_t la tribu sur Ω engendrée par $(X_s)_{0 \leq s \leq t}$ et $\mathcal{F} = \bigvee_{t \geq 0} \mathcal{F}_t$.

Si on se donne une probabilité initiale m sur M et une l'évolution temporelle de l'inverse de la température $\beta \in C^1([0, +\infty[, \mathbb{R}_+)$, il est bien connu qu'il existe une unique probabilité $P_{m,\beta}$ sur (Ω, \mathcal{F}) telle que

$$X_0 \circ P_{m,\beta} = m$$

et telle que pour toute fonction $\phi \in C^\infty(M)$,

$$t \mapsto \phi(X_t) - \phi(X_0) - \int_0^t L_{\beta_s} \phi(X_s) ds$$

soit une $(\mathcal{F}_t)_{t \geq 0}$ -martingale.

Le processus $(X_t)_{t \geq 0}$ sous une telle probabilité $P_{m,\beta}$ est appelé un algorithme du recuit simulé associé au potentiel U , à la loi initiale m et à l'évolution β . Pour $t \geq 0$, on note m_t la loi de X_t .

En fait comme précédemment, on ne s'intéressera qu'à des évolutions β pour lesquelles il existe une constante strictement positive k telle que, pour tout $t \geq 2$, on ait $\beta_t = k^{-1} \ln(t)$. On parlera alors encore d'un k -algorithme.

Soit $k > 0$ fixé, pour donner la définition d'une k -cellule (en lesquelles on va essayer de décomposer M si l'on considère un k -algorithme), il faut adapter à la situation continue les notions introduites dans la section 1.

Soit A un domaine de M (i.e. un ouvert non vide et connexe de M). Pour $x, y \in A$, on note $\mathcal{C}_{x,y}^A$ l'ensemble des chemins continus allant de x à y dans A , c'est-à-dire l'ensemble des applications continues $\varphi : [0, 1] \ni t \mapsto \varphi(t) \in A$ telles que $\varphi(0) = x$ et $\varphi(1) = y$.

L'élévation d'un tel chemin est le nombre

$$e(\varphi) = \max_{0 \leq t \leq 1} U(\varphi(t))$$

et on définit l'énergie minimale à fournir pour aller dans A de x à y ou de y à x comme étant

$$H_A(x, y) = \inf_{\varphi \in \mathcal{C}_{x,y}^A} e(\varphi) - U(x) - U(y) + \min_{z \in A} U(z).$$

Il n'est pas difficile de voir que H_A est continue sur $A \times A$, et on associe à A la constante

$$c(A) = \sup_{x,y \in A} H_A(x, y)$$

(pour faire apparaître la dépendance en U , on notera aussi $c(A, U)$ ce nombre).

Définissons également le fond de A comme étant la plus petite valeur de U sur A

$$f(A) = \inf_{z \in A} U(z)$$

et la hauteur de A comme étant le nombre

$$h(A) = \min_{x \in \partial A} U(x) - f(A)$$

où $\partial A = \overline{A} \setminus A$ est la frontière topologique de A .

Le domaine A sera une k -cellule si et seulement si

$$c(A) < k \quad \text{et} \quad h(A) > k.$$

Notons que si A est une k -cellule, l'infimum qui définit $f(A)$ est en fait un minimum, et si on appelle N_A l'ensemble des points où il est atteint,

$$N_A = \{z \in A \mid U(z) = f(A)\} ;$$

il est facile de vérifier que pour tout $y \in N_A$ fixé, on a

$$c(A) = \max_{x \in A} H_A(x, y) .$$

On dira qu'une suite finie $(A_i)_{1 \leq i \leq R}$ de k -cellules est une décomposition de M , si $M = \bigcup_{1 \leq i \leq R} A_i$ et si pour $1 \leq i \neq j \leq R$, on a $\inf_{z \in A_i \cap A_j} U(z) > f(A_i) \vee f(A_j) + \bar{k}$ (on fait la convention habituelle $\inf \emptyset = +\infty$).

Appelons \mathcal{K} l'ensemble des $k > 0$ pour lesquels il n'existe pas de décomposition de M en k -cellules. La première différence qui apparaît avec le cas fini, c'est qu'il se peut que \mathcal{K} soit infini, comme le montre l'exemple où M est le cercle et où U est le potentiel défini sur $[0, 2\pi[$ par

$$U(x) = \begin{cases} \exp\left(-\frac{1}{\sin(x/2)}\right) \sin^2\left(\frac{1}{x}\right) & \text{si } x \in]0, 2\pi[\\ 0 & \text{si } x = 0 \end{cases}$$

(voir aussi l'appendice).

Cependant, on verra que \mathcal{K} est toujours borné et dénombrable, et que si \mathcal{K} n'est pas fini, alors son seul point d'accumulation dans \mathbb{R}_+ est 0.

D'autre part, définissons la densité de Gibbs $\mu_{A,\beta}$ associée à un domaine A et à l'inverse de la température $\beta \geq 0$, par

$$\mu_{A,\beta}(x) = \begin{cases} Z_{A,\beta}^{-1} \exp(-\beta U(x)) & \text{si } x \in A \\ 0 & \text{sinon} \end{cases}$$

où $Z_{A,\beta} = \int_A \exp(-\beta U(y)) \lambda(dy)$. On notera aussi $\mu_{A,\beta}$ la probabilité sur M qui admet cette fonction pour densité par rapport à λ (et plus généralement, on désignera par un même symbole une mesure sur M absolument continue par rapport à λ et sa densité).

Il est clair que quand β tend vers l'infini, la probabilité $\mu_{A,\beta}$ tend à se concentrer au voisinage de N_A . De plus, sous certaines conditions de régularité de U au voisinage de N_A , on peut obtenir des résultats plus précis, ainsi Hwang a montré la convergence étroite des $\mu_{A,\beta}$ vers une certaine probabilité $\mu_{A,\infty}$ (cf. [11]).

Nous pouvons maintenant énoncer l'équivalent du théorème 1.

THÉOREME 9. — Soit $(m_t)_{t \geq 0}$ la famille des lois associées à un k -algorithme, avec $k > 0$, $k \notin \mathcal{K}$. Soit $(A_i)_{1 \leq i \leq R}$ une décomposition de M en k -cellules.

Alors pour tout $1 \leq i \leq R$, $m_t(A_i)$ converge en temps grand vers une limite strictement positive λ_i et si on note $\|\cdot\|$ la variation totale des mesures, on a

$$\lim_{t \rightarrow +\infty} \left\| m_t - \sum_{i=1}^R \lambda_i \mu_{A_i, \beta_t} \right\| = 0.$$

Ainsi les minima locaux au voisinage desquels m_t tend à se concentrer en temps grand sont les éléments de $\bigcup_{i=1}^R N_{A_i}$. Comme dans le cas discret, cet ensemble ne dépend évidemment pas de la partition en k -cellules choisie et on peut voir facilement qu'un élément x de M y appartient si et seulement s'il existe une k -cellule A telle que $x \in N_A$.

Plus généralement, les remarques et le corollaire qui suivent le théorème 1 demeurent valables, mis à part que le cas où $k < c(M)$ n'a pas été très étudié dans cette situation.

Avant d'esquisser la démonstration du théorème, prouvons les résultats annoncés sur \mathcal{K} . On va pour cela donner une caractérisation de cet ensemble.

Pour $x, y \in M$, notons

$$\tilde{H}(x, y) \longrightarrow \inf_{\phi \in \mathcal{C}_{x,y}^M} e(\phi) - U(x) - U(y) + U(x) \wedge U(y)$$

et si $A \subset M$, appelons $D(A)$ l'ensemble $\{x \notin A \mid U(x) \leq f(A)\}$.

Notons $\hat{\mathcal{N}}$ l'ensemble des fermés non vides $N \subset M$ sur lesquels U est constant. Si $N \in \hat{\mathcal{N}}$, on définit

$$K(N) = \inf_{\substack{x \in N \\ y \in D(N)}} \tilde{H}(x, y) \geq 0$$

et on appelle

$$\mathcal{N} = \left\{ N \in \hat{\mathcal{N}} \mid \max_{x, y \in N} \tilde{H}(x, y) \leq K(N) \right\}.$$

Posons aussi

$$\tilde{\mathcal{K}} = \{K(N) \mid N \in \hat{\mathcal{N}}\} \cap]0, +\infty[$$

et

$$\tilde{\mathcal{K}} = \{K(N) \mid N \in \mathcal{N}\} \cap]0, +\infty[.$$

On a clairement $\tilde{\mathcal{K}} \subset \hat{\mathcal{K}}$. Vérifions qu'en fait $\tilde{\mathcal{K}} = \hat{\mathcal{K}}$.

Soit donc $k \in \hat{\mathcal{K}}$ et $N \in \hat{\mathcal{N}}$ tel que $K(N) = k$. Du fait que $k > 0$ et de la compacité de M , remarquons qu'il existe $x_0 \in N$ et $y_0 \in D(N)$ tels que $\tilde{H}(x_0, y_0) = K(N)$. Considérons

$$N' = \{x \in N \mid \tilde{H}(x_0, x) \leq K(N)\}.$$

Cet ensemble est un fermé non vide et il est immédiat de vérifier que $N' \in \mathcal{N}$. De plus, on montre facilement que $K(N') = K(N)$, d'où finalement $\tilde{\mathcal{K}} = \hat{\mathcal{K}}$.

L'intérêt que l'on a à considérer \mathcal{N} (par rapport à $\hat{\mathcal{N}}$) est que l'on peut montrer facilement que pour tout $\varepsilon > 0$, $\tilde{\mathcal{K}} \cap [\varepsilon, +\infty[$ est un ensemble fini. En effet, remarquons que si $N, N' \in \mathcal{N}$ sont tels que $K(N) > K(N') > 0$, alors $N \cap N' = \emptyset$ ou $N' \subset N$, auquel cas on a $K(N') = \inf_{x \in N', y \in N \setminus N'} \tilde{H}(x, y)$, ce qui permet de voir que $N \setminus N'$ est fermé (et appartient donc à $\hat{\mathcal{N}}$) et que $K(N \setminus N') = K(N')$.

Ainsi s'il existait une suite strictement décroissante $(k_n)_{n \geq 1}$ d'éléments de $\tilde{\mathcal{K}} \cap [\varepsilon, +\infty[$, quitte à en extraire une sous-suite, soit il existe une suite $(N_n)_{n \geq 1}$ d'éléments de \mathcal{N} deux à deux disjoints tels que $K(N_n) = k_n$, soit il existe une suite décroissante $(N_n)_{n \geq 1}$ d'éléments de \mathcal{N} tels que $K(N_n) = k_n$. Dans ce dernier cas, posons pour $n \geq 1$, $R_n = N_n \setminus N_{n+1}$, on a $R_n \in \hat{\mathcal{N}}$ et $K(R_n) = K(N_{n+1})$. Dans tous les cas, on a donc mis en évidence une suite infinie d'éléments deux à deux disjoints de $\{N \in \hat{\mathcal{N}} \mid K(N) \geq \varepsilon\}$. Cependant, si N et N' sont deux éléments disjoints de cet ensemble, alors $N_\varepsilon = \{x \in M \mid \min_{y \in N} d(x, y) < \varepsilon\}$ et $N'_\varepsilon = \{x \in M \mid \min_{y \in N'} d(x, y) < \varepsilon\}$ sont encore disjoints, du fait que $\min_{z \in M} |\nabla U|(z) \leq 1$. La compacité de M entraîne alors que $\{N \in \hat{\mathcal{N}} \mid K(N) \geq \varepsilon\}$ ne peut pas contenir une suite infinie d'éléments deux à deux disjoints, d'où la contradiction recherchée et la finitude de $\tilde{\mathcal{K}} \cap [\varepsilon, +\infty[$.

Les résultats annoncés sur \mathcal{K} découlent alors immédiatement de la caractérisation suivante.

PROPOSITION 10. — *On a l'égalité*

$$\mathcal{K} = \tilde{\mathcal{K}}$$

Démonstration. — L'inclusion $\tilde{\mathcal{K}} \subset \mathcal{K}$ est une conséquence immédiate du lemme présenté dans l'appendice, on ne prouvera donc ici que l'inclusion intéressante $\mathcal{K} \subset \tilde{\mathcal{K}}$.

Soit $k \notin \tilde{\mathcal{K}}$, $k > 0$, on va montrer qu'il existe une décomposition de M en k -cellules. D'après les résultats précédents

$$\delta = \left(\min_{r \in \tilde{\mathcal{K}}} |k - r| \right) \wedge \frac{k}{2}$$

est une constante strictement positive. Pour se ramener au cas des graphes, recouvrons M par un nombre fini N de boules ouvertes $(B_\ell)_{1 \leq \ell \leq N}$ de rayon $\delta/6$. Notons E l'ensemble $\{1, \dots, N\}$ que l'on munit de la probabilité uniforme μ et du noyau de probabilités de transition q défini par

$$\forall \ell \neq \ell' \in E, \quad q(\ell, \ell') = \begin{cases} N^{-1} & \text{si } B_\ell \cap B_{\ell'} \neq \emptyset \\ 0 & \text{sinon} \end{cases}$$

et

$$\forall \ell \in E, \quad q(\ell, \ell) = 1 - \sum_{\ell' \neq \ell} q(\ell, \ell').$$

Le noyau q est clairement réversible par rapport à μ et irréductible. Considérons le potentiel \tilde{U} défini sur E par

$$\forall \ell \in E, \quad \tilde{U}(\ell) = \min_{x \in \overline{B}_\ell} U(x).$$

On peut maintenant appliquer la proposition 3 pour obtenir une partition de E en $(k - (2\delta/3))$ -cellules larges à bord relevé, relativement à (q, \tilde{U}) , $E = \bigsqcup_{1 \leq i \leq R} V_i$. Pour $1 \leq i \leq R$, notons

$$A_i = \bigcup_{\ell \in V_i} B_\ell.$$

Il est clair que les A_i sont des domaines qui satisfont

$$c(A_i) \leq k - \frac{2\delta}{3} + \frac{\delta}{3} = k - \frac{\delta}{3}.$$

Pour vérifier que la famille $(A_i)_{1 \leq i \leq R}$ est une décomposition de M en k -cellules, il suffit de voir que pour tout $1 \leq i \neq j \leq R$

$$\inf_{z \in A_i \cap A_j} U(z) > f(A_i) \vee f(A_j) + k. \quad (\star)$$

Soit $1 \leq i_0 \leq R$ fixé, remarquons que

$$\inf_{z \in A_{i_0} \cap (\cup_{j \neq i_0} A_j)} U(z) \geq f(A_{i_0}) + k - \frac{2\delta}{3}.$$

En effet, soient $1 \leq j_0 \leq R$, $j_0 \neq i_0$, puis $\ell \in V_{i_0}$ et $\ell' \in V_{j_0}$, et enfin $z_0 \in \overline{B_\ell} \cap \overline{B_{\ell'}}$ tels que

$$\inf_{z \in A_{i_0} \cap (\cup_{j \neq i_0} A_j)} U(z) = \inf_{z \in A_{i_0} \cap A_{j_0}} U(z) = \min_{z \in \overline{B_\ell} \cap \overline{B_{\ell'}}} U(z) = U(z_0).$$

Puisque les V_i forment une partition de E en $(k - 2\delta/3)$ -cellules larges, on a

$$\begin{aligned} U(z_0) &\geq \tilde{U}(\ell) \vee \tilde{U}(\ell') \geq f(V_{i_0}) \vee f(V_{j_0}) + k - \frac{2\delta}{3} \\ &= f(A_{i_0}) \vee f(A_{j_0}) + k - \frac{2\delta}{3}; \end{aligned}$$

d'où le résultat annoncé. Celui-ci montre que pour tout $1 \leq i \leq R$, on a

$$K(N_{A_i}) \geq k - \frac{2\delta}{3}$$

et du fait que $\tilde{\mathcal{K}} \cap]k - \delta, k + \delta[= \emptyset$, cette inégalité implique que

$$K(N_{A_i}) \geq k + \delta.$$

Fixons maintenant $1 \leq i \neq j \leq R$ tels que $A_i \cap A_j \neq \emptyset$, et montrons (★). Par symétrie du problème, on peut supposer que $f(A_i) \geq f(A_j)$. Soient $x_i \in N_{A_i}$ et $x_j \in N_{A_j}$. L'inégalité précédente montre que

$$\tilde{H}(x_i, x_j) \geq k + \delta.$$

Cependant, soient $\ell \in V_i$ et $\ell' \in V_j$, puis $z_0 \in \overline{B_\ell} \cap \overline{B_{\ell'}}$ tels que

$$U(z_0) = \inf_{z \in A_i \cap A_j} U(z).$$

Puisque V_i et V_j sont à bord relevé, il est facile de construire un chemin $\phi \in C_{x_i, x_j}^{\overline{A_i \cup A_j}}$ tel que $e(\phi) \leq U(z_0) + \delta/3$, ce qui permet d'obtenir que

$$\begin{aligned} U(z_0) &\geq \min_{\phi \in C_{x_i, x_j}^M} e(\phi) - \frac{\delta}{3} \\ &\geq \tilde{H}(x_i, x_j) + f(A_i) - \frac{\delta}{3} \\ &\geq f(A_i) + k + \frac{2\delta}{3} \end{aligned}$$

puis finalement le résultat annoncé. \square

Remarques

La proposition précédente permet de donner un critère simple de finitude de \mathcal{K} . Soit L l'ensemble des minima locaux, c'est-à-dire l'ensemble des $x \in M$ pour lesquels il existe $\varepsilon > 0$ tel que pour tout $y \in M$, on ait

$$U(y) < U(x) \implies \inf_{\varphi \in \mathcal{C}_{x,y}^M} e(\varphi) \geq U(x) + \varepsilon.$$

On introduit sur L la relation d'équivalence \sim définie par

$$x \sim y \iff \left[U(x) = U(y) \text{ et } \inf_{\varphi \in \mathcal{C}_{x,y}^M} e(\varphi) = U(x) \right].$$

Il est clair qu'une condition suffisante pour que \mathcal{K} soit fini est que l'ensemble \mathcal{L} des classes d'équivalence de \sim sur L soit fini (car on vérifie facilement que l'on a également $\tilde{\mathcal{K}} = \{K(\ell) \mid \ell \in \mathcal{L}\} \cap]0, +\infty[$).

On montrera dans l'appendice que cette condition est aussi nécessaire.

D'autre part, notons que $c(M)$ est l'élément maximum de \mathcal{K} et que la caractérisation de \mathcal{K} donnée dans la proposition précédente est également valable pour le cas discret.

Prouvons maintenant le théorème 9 en suivant la méthode donnée dans le cas discret. Le point technique le plus délicat est la preuve de la proposition 11 suivante.

On fixe désormais $k > 0$, $k \notin \mathcal{K}$, et $(A_i)_{1 \leq i \leq R}$ désigne une décomposition de M en k -cellules. On note

$$\delta = \min \left[\min_{1 \leq i \leq R} (k - c(A_i)); \min_{1 \leq i, j \leq R} (f(A_i \cap A_j) - f(A_i) \vee f(A_j) - k) \right] > 0$$

et $\{f_1, \dots, f_S\} = \{f(A_i) \mid 1 \leq i \leq R\}$, de telle sorte que $1 \leq i < j \leq S \implies f_i > f_j$.

Il est bien connu que pour $t > 0$, m_t est absolument continue par rapport à λ et que sa densité est strictement positive et de classe C^∞ . Elle satisfait de plus sur $]0, +\infty[\times M$ l'équation de Chapman-Kolmogorov

$$\begin{aligned} \frac{d}{dt} m_t &= L_{\beta_t}^*(m_t) \\ &= \operatorname{div}(\mu_{\beta_t} \nabla f_t) \end{aligned}$$

où on convient de poser $\mu_{\beta_t} = \mu_{M, \beta_t}$ et $f_t = m_t / \mu_{\beta_t}$.

Pour $p \geq 1$ et $t > 0$ la fonctionnelle suivante est donc bien définie

$$F_p(t) = \int f_t^p d\mu_{\beta_t}$$

et on a l'équivalent de la proposition 6 :

PROPOSITION 11. — *Il existe deux constantes $K_p > 0$ et $r_p \geq 0$ (qui dépendent aussi, outre de p , de la structure riemannienne de M , du potentiel U , de k et de $F_p(2)$), telles que pour tout $t \geq 2$, on ait*

$$F_p(t) \leq K_p \beta_t^{r_p} \exp((p-1)f_1\beta_t).$$

Démonstration. — On va tout d'abord prouver la proposition pour des p de la forme $p = 2^n$, par une récurrence sur n .

Pour $n = 0$, on a $p = 1$, or $F_1(t) = \int dm_t = 1$, et la proposition est donc vérifiée avec $K_1 = 1$ et $r_1 = 0$.

Supposons maintenant que la proposition soit satisfaite pour $p/2 \geq 1$ et montrons la pour p .

L'équation de Chapman-Kolmogorov permet de voir que $t \mapsto F_p(t)$ est de classe C^1 sur $]0, +\infty[$ et que pour $t \geq 2$, on a

$$\begin{aligned} \frac{d}{dt} F_p(t) &= p \int f_t^{p-1} \frac{d}{dt} m_t d\lambda + (1-p) \int f_t^p \frac{d}{dt} \mu_{\beta_t} d\lambda \\ &= -p \int \langle \nabla f_t^{p-1}, \nabla f_t \rangle d\mu_{\beta_t} + (p-1) \frac{d\beta_t}{dt} \int f_t^p (U - \langle U \rangle_{\beta_t}) d\mu_{\beta_t} \\ &\leq -4 \frac{p-1}{p} \int |\nabla f_t^{p/2}|^2 d\mu_{\beta_t} + (p-1) \frac{1}{kt} \|U\|_{\infty} F_p(t). \end{aligned}$$

(où $\langle U \rangle_{\beta_t} = \int U d\mu_{\beta_t}$). Il faut maintenant comparer

$$\int |\nabla f_t^{p/2}|^2 d\mu_{\beta_t} \quad \text{et} \quad F_p(t).$$

On introduit pour cela, pour $1 \leq i \leq R$, les domaines

$$\tilde{A}_i = \left\{ x \in M \mid \exists y \in A_i, d(x, y) < \frac{\delta}{2} \right\}.$$

Du fait que $\max_{z \in M} |\nabla U|(z) \leq 1$, remarquons que

$$\inf_{z \in \tilde{A}_i \setminus A_i} U(z) \geq f(A_i) + k + \frac{\delta}{2},$$

ainsi $N_{\tilde{A}_i} = N_{A_i}$ et $f(\tilde{A}_i) = f(A_i)$ (on a aussi $h(\tilde{A}_i) \geq k + \delta/2$ et $c(\tilde{A}_i) \leq k - \delta/2$).

D'autre part, soient $\beta \geq 2/\delta \vee 1$, $1 \leq i \leq R$ et $x_{i,1} \in N_{A_i}$ fixés. Un résultat classique de géométrie différentielle permet de choisir une famille $(x_{i,j})_{2 \leq j \leq N_\beta(i)}$ de $N_\beta(i) - 1$ autres points de A_i , telle que les boules $(B(x_{i,j}, \beta^{-1}))_{1 \leq j \leq N_\beta(i)}$ recouvrent $\overline{A_i}$, avec $N_\beta(i)$ majoré par un terme de la forme $\tilde{K}_1 \beta^{\tilde{r}_1}$, où les constantes $\tilde{K}_1 > 0$ et $\tilde{r}_1 > 0$ ne dépendent que de la structure riemannienne de M .

Soit $\phi \in C^1(M)$. Holley, Kusuoka et Stroock ont montré (cf. les calculs des pages 337 à 340 de [9]) qu'il existe deux autres constantes $\tilde{K}_2 > 0$ et $\tilde{r}_2 > 0$ ne dépendant que de la structure riemannienne de M telles que pour tout $1 \leq i \leq R$, tout $1 \leq j \leq N_\beta(i)$ et tout $\beta \geq 2/\delta \vee 1$, on ait

$$\begin{aligned} & \int_{B(x_{i,j}, \beta^{-1}) \times B(x_{i,1}, \beta^{-1})} (\phi(x) - \phi(y))^2 \mu_{\tilde{A}_i, \beta}(dx) \mu_{\tilde{A}_i, \beta}(dy) \\ & \leq \tilde{K}_2 \beta^{\tilde{r}_2} \exp(\beta H_{A_i}(x_{i,j}, x_{i,1})) \int_{\tilde{A}_i} |\nabla \phi|^2 d\mu_{\tilde{A}_i, \beta} \\ & \leq \tilde{K}_2 \beta^{\tilde{r}_2} \exp(\beta c(A_i)) \int_{\tilde{A}_i} |\nabla \phi|^2 d\mu_{\tilde{A}_i, \beta}. \end{aligned}$$

On applique ce résultat avec $\phi = f_t^{p/2}$ et $\beta = \beta_t$, pour obtenir, pour t assez grand,

$$\begin{aligned} & \int_{\tilde{A}_i \times B(x_{i,1}, \beta_t^{-1})} (f_t^{p/2}(x) - f_t^{p/2}(y))^2 \mu_{\tilde{A}_i, \beta_t}(dx) \mu_{\tilde{A}_i, \beta_t}(dy) \leq \\ & \leq \sum_{1 \leq j \leq N_{\beta_t}(i)} \int_{B(x_{i,j}, \beta_t^{-1}) \times B(x_{i,1}, \beta_t^{-1})} (f_t^{p/2}(x) - f_t^{p/2}(y))^2 \times \\ & \quad \times \mu_{\tilde{A}_i, \beta_t}(dx) \mu_{\tilde{A}_i, \beta_t}(dy) \\ & \leq N_{\beta_t}(i) \tilde{K}_2 \beta_t^{\tilde{r}_2} \exp(\beta_t c(A_i)) \int_{\tilde{A}_i} |\nabla f_t^{p/2}|^2 d\mu_{\tilde{A}_i, \beta_t} \\ & \leq \tilde{K}_1 \tilde{K}_2 \beta_t^{\tilde{r}_1 + \tilde{r}_2} \exp(\beta_t c(A_i)) \int_{\tilde{A}_i} |\nabla f_t^{p/2}|^2 d\mu_{\tilde{A}_i, \beta_t}. \end{aligned}$$

Intéressons-nous au terme

$$\begin{aligned}
 & \int_{\bar{A}_i \times B(x_{i,1}, \beta_t^{-1})} (f_t^{p/2}(x) - f_t^{p/2}(y))^2 \mu_{\tilde{A}_i, \beta_t}(dx) \mu_{\tilde{A}_i, \beta_t}(dy) = \\
 & = \mu_{\tilde{A}_i, \beta_t}(B(x_{i,1}, \beta_t^{-1})) \int_{\bar{A}_i \times B(x_{i,1}, \beta_t^{-1})} (f_t^{p/2}(x) - f_t^{p/2}(y))^2 \times \\
 & \quad \times \mu_{\tilde{A}_i, \beta_t}(dx) \frac{1}{\mu_{\tilde{A}_i, \beta_t}(B(x_{i,1}, \beta_t^{-1}))} \mu_{\tilde{A}_i, \beta_t}(dy) \\
 & \geq \mu_{\tilde{A}_i, \beta_t}(B(x_{i,1}, \beta_t^{-1})) \int_{\bar{A}_i} \times \\
 & \quad \times \left[f_t^{p/2}(x) - \int_{B(x_{i,1}, \beta_t^{-1})} f_t^{p/2}(y) \frac{1}{\mu_{\tilde{A}_i, \beta_t}(B(x_{i,1}, \beta_t^{-1}))} \mu_{\tilde{A}_i, \beta_t}(dy) \right]^2 \mu_{\tilde{A}_i, \beta_t}(dx) \\
 & \geq \mu_{\tilde{A}_i, \beta_t}(B(x_{i,1}, \beta_t^{-1})) \left[\left(\int_{\bar{A}_i} f_t^p(x) \mu_{\tilde{A}_i, \beta_t}(dx) \right)^{\frac{1}{2}} + \right. \\
 & \quad \left. - \left(\int_{\bar{A}_i} \left[\int_{B(x_{i,1}, \beta_t^{-1})} f_t^{p/2}(y) \frac{1}{\mu_{\tilde{A}_i, \beta_t}(B(x_{i,1}, \beta_t^{-1}))} \mu_{\tilde{A}_i, \beta_t}(dy) \right]^2 \mu_{\tilde{A}_i, \beta_t}(dx) \right)^{\frac{1}{2}} \right]^2 \\
 & \geq \mu_{\tilde{A}_i, \beta_t}(B(x_{i,1}, \beta_t^{-1})) \left[\frac{1}{2} \int_{\bar{A}_i} f_t^p(x) \mu_{\tilde{A}_i, \beta_t}(dx) + \right. \\
 & \quad \left. - \mu_{\tilde{A}_i, \beta_t}(\bar{A}_i) \left(\int_{B(x_{i,1}, \beta_t^{-1})} f_t^{p/2}(y) \frac{1}{\mu_{\tilde{A}_i, \beta_t}(B(x_{i,1}, \beta_t^{-1}))} \mu_{\tilde{A}_i, \beta_t}(dy) \right)^2 \right].
 \end{aligned}$$

On en déduit que

$$\begin{aligned}
 & \int_{\bar{A}_i} f_t^p d\mu_{\tilde{A}_i, \beta_t} \leq \\
 & \leq \frac{2\mu_{\tilde{A}_i, \beta_t}(\bar{A}_i)}{[\mu_{\tilde{A}_i, \beta_t}(B(x_{i,1}, \beta_t^{-1}))]^2} \left(\int_{B(x_{i,1}, \beta_t^{-1})} f_t^{p/2} d\mu_{\tilde{A}_i, \beta_t} \right)^2 + \\
 & \quad + \frac{2}{\mu_{\tilde{A}_i, \beta_t}(B(x_{i,1}, \beta_t^{-1}))} \tilde{K}_1 \tilde{K}_2 \beta_t^{\tilde{r}_1 + \tilde{r}_2} \exp(\beta_t c(A_i)) \int_{\tilde{A}_i} |\nabla f_t^{p/2}|^2 d\mu_{\tilde{A}_i, \beta_t},
 \end{aligned}$$

c'est-à-dire aussi

$$\begin{aligned} \int_{A_i} f_t^p d\mu_{\beta_t} &\leq \\ &\leq \frac{2\mu_{\tilde{A}_i, \beta_t}(\overline{A_i})}{[\mu_{\tilde{A}_i, \beta_t}(B(x_{i,1}, \beta_t^{-1}))]^2} \frac{Z_{M, \beta_t}}{Z_{\tilde{A}_i, \beta_t}} \left(\int_{B(x_{i,1}, \beta_t^{-1})} f_t^{p/2} d\mu_{\beta_t} \right)^2 + \\ &\quad + \frac{2}{\mu_{\tilde{A}_i, \beta_t}(B(x_{i,1}, \beta_t^{-1}))} \tilde{K}_1 \tilde{K}_2 \beta_t^{\tilde{r}_1 + \tilde{r}_2} \exp(\beta_t c(A_i)) \int_{\tilde{A}_i} |\nabla f_t^{p/2}|^2 d\mu_{\beta_t}. \end{aligned}$$

Cependant, $\mu_{\tilde{A}_i, \beta_t}(\overline{A_i}) \leq 1$ et $Z_{M, \beta_t} \leq 1$, et il existe des constantes $\tilde{K}_3 > 0$ et $\tilde{r}_3 > 0$ ne dépendant que de la structure riemannienne de M , telles que pour tout $1 \leq i \leq R$, et tout $t \geq 0$, on ait

$$\frac{1}{\mu_{\tilde{A}_i, \beta_t}(B(x_{i,1}, \beta_t^{-1}))} \leq \tilde{K}_3 \beta_t^{\tilde{r}_3}$$

et

$$Z_{\tilde{A}_i, \beta_t}^{-1} \leq \tilde{K}_3 \beta_t^{\tilde{r}_3} \exp(f(A_i)\beta_t) \leq \tilde{K}_3 \beta_t^{\tilde{r}_3} \exp(f_1\beta_t);$$

ainsi, en notant $K = \tilde{K}_1 \tilde{K}_2 \tilde{K}_3$ et $r = \tilde{r}_1 + \tilde{r}_2 + \tilde{r}_3$, on a

$$\begin{aligned} \int_{A_i} f_t^p d\mu_{\beta_t} &\leq 2\tilde{K}_3^3 \beta_t^{3\tilde{r}_3} \exp(f_1\beta_t) \left(\int_{B(x_{i,1}, \beta_t^{-1})} f_t^{p/2} d\mu_{\beta_t} \right)^2 + \\ &\quad + 2K \beta_t^r \exp(\beta_t c(A_i)) \int_{\tilde{A}_i} |\nabla f_t^{p/2}|^2 d\mu_{\beta_t} \\ &\leq 2\tilde{K}_3^3 \beta_t^{3\tilde{r}_3} \exp(f_1\beta_t) \left(\int_M f_t^{p/2} d\mu_{\beta_t} \right)^2 + \\ &\quad + 2K \beta_t^r \exp((k - \delta)\beta_t) \int_M |\nabla f_t^{p/2}|^2 d\mu_{\beta_t}. \end{aligned}$$

On utilise l'hypothèse de récurrence pour estimer l'expression

$$\int_M f_t^{p/2} d\mu_{\beta_t} \leq K_{p/2} \beta_t^{r_{p/2}} \exp\left(\left(\frac{p}{2} - 1\right)f_1\beta_t\right)$$

puis on somme sur $1 \leq i \leq R$, pour obtenir

$$\begin{aligned} F_p(t) &\leq \sum_{i=1}^R \int_{A_i} f_t^p d\mu_{\beta_t} \leq 2R \tilde{K}_3^3 K_{p/2}^2 \beta_t^{3\tilde{r}_3} \beta_t^{2r_{p/2}} \exp((p-1)f_1\beta_t) + \\ &\quad + 2RK \beta_t^r \exp((k - \delta)\beta_t) \int_M |\nabla f_t^{p/2}|^2 d\mu_{\beta_t}. \end{aligned}$$

En regroupant les résultats précédents, on a donc montré que $F_p(t)$ satisfait pour $t \geq 2$ l'inégalité différentielle

$$\begin{aligned} \frac{d}{dt} F_p(t) &\leq -4 \frac{p-1}{p} R^{-1} K^{-1} \beta_t^{-r} \exp(-(k-\delta)\beta_t) \times \\ &\quad \times \left[F_p(t) - 2R\tilde{K}_3^3 K_{p/2}^2 \beta_t^{3\tilde{r}_3+2r_{p/2}} \exp((p-1)f_1\beta_t) \right] + \\ &\quad + (p-1) \frac{1}{kt} \|U\|_{\infty} F_p(t) \end{aligned}$$

et il est clair que l'on peut récrire ceci sous la forme suivante : il existe des constantes $A_p > 0$, $B_p > 0$ et $T_p \geq 2$ (dépendant de M , U , p , k , R et δ) telles que pour tout $t \geq T_p$, on ait

$$\frac{d}{dt} F_p(t) \leq A_p (\ln(t))^{3\tilde{r}_3+2r_{p/2}-r} t^{\frac{(p-1)f_1-(k-\delta)}{k}} - B_p (\ln(t))^{-r} t^{-\frac{k-\delta}{k}} F_p(t).$$

Une généralisation immédiate du lemme 5 permet alors d'obtenir le résultat annoncé.

Soit maintenant $p > 1$ quelconque et notons q le plus petit entier de la forme 2^n qui lui soit supérieur. On a

$$\begin{aligned} F_p(t) &= \int f_t^{p-1} dm_t \leq \left(\int f_t^{q-1} dm_t \right)^{\frac{p-1}{q-1}} \\ &\leq K_q^{(p-1)/(q-1)} \beta_t^{(p-1)/(q-1)r_q} \exp((p-1)f_1\beta_t) \end{aligned}$$

et il suffit donc de prendre $K_p = K_q^{(p-1)/(q-1)}$ et $r_p = (p-1)/(q-1)r_q$, ce qui termine la démonstration de la proposition. \square

Désormais les preuves sont très proches de celles données dans le cas discret et on se contentera donc de les esquisser.

Définissons les ensembles suivants :

$$\begin{aligned} I_1 &= \{1 \leq i \leq R \mid f(A_i) = f_1\} \\ D_1 &= \left\{ x \in M \mid U(x) \leq f_1 + k + \frac{\delta}{2} \right\} \\ D'_1 &= \left\{ x \in M \mid U(x) > f_1 + k + \frac{\delta}{4} \right\} \end{aligned}$$

et soit $\phi_1 \in C^\infty(M)$ telle que

$$\phi_1(x) = \begin{cases} 1 & \text{si } x \notin D_1 \\ 0 & \text{si } x \notin D'_1. \end{cases}$$

La proposition précédente permet de montrer, d'une part que la semi-martingale (continue) $t \mapsto \phi_1(X_t)$ converge p.s. en temps grand, et d'autre part que

$$\lim_{t \rightarrow +\infty} m_t(D'_1) = 0$$

ce qui prouve qu'on a en fait p.s.

$$\lim_{t \rightarrow +\infty} \phi_1(X_t) = 0.$$

Pour interpréter ceci, considérons pour $1 \leq i \leq R$, les compacts

$$K_i = \left\{ x \in A_i \mid U(x) \leq f(A_i) + k + \frac{\delta}{2} \right\}$$

$$K'_i = \left\{ x \in M \mid \inf_{y \in K_i} d(x, y) \geq \frac{\delta}{4} \right\}$$

et

$$R_1 = D_1 \setminus \bigcup_{i \in I_1} K_i$$

$$R'_1 = \left\{ x \in M \mid \inf_{y \in R_1} d(x, y) \geq \frac{\delta}{4} \right\}.$$

Le résultat précédent montre que p.s. la trajectoire d'un k -algorithme finit par ne plus sortir, soit de R_1 , soit d'un K_i , avec $i \in I_1$.

Notons $U_{R_1} \in C^\infty(M)$ un potentiel tel que

$$U_{R_1}(x) = \begin{cases} U(x) & \text{si } x \in R_1 \\ f_1 + k + \frac{3\delta}{4} & \text{si } x \in R'_1 \end{cases}$$

et tel que

$$\forall x \notin R_1 \cup R'_1, \quad f_1 + k + \frac{\delta}{2} \leq U_{R_1}(x) \leq f_1 + k + \frac{3\delta}{4}.$$

Il est facile de construire à partir des ensembles $(A_i)_{i \notin I_1}$ une décomposition de M en k -cellules pour le potentiel U_{R_1} dont l'ensemble des fonds est $\{f_2, \dots, f_S\}$.

De même pour $1 \leq i \leq R$, notons $U_{K_i} \in C^\infty(M)$ un potentiel tel que

$$U_{K_i}(x) = \begin{cases} U(x) & \text{si } x \in K_i \\ f(A_i) + k + \frac{3\delta}{4} & \text{si } x \in K'_i \end{cases}$$

et tel que

$$\forall x \notin K_i \cup K'_i, \quad f(A_i) + k + \frac{\delta}{2} \leq U_{K_i}(x) \leq f(A_i) + k + \frac{3\delta}{4}.$$

Remarquons que la constante $c(M, U_{K_i})$, associée sur M au potentiel U_{K_i} , satisfait

$$c(M, U_{K_i}) < k.$$

Soit $B \in \{K_1, \dots, K_R, R_1\}$ et $T > 0$, on appelle $Q^{B,T}$ la probabilité sur $C([T, +\infty[; M)$ qui est solution du problème de martingales associé à la famille

$$(\Delta \cdot -\beta_t \langle \nabla U_B, \nabla \cdot \rangle)_{t \geq T}$$

d'opérateurs sur $C^\infty(M)$, et qui admet

$$\frac{1}{m_T(B)} m_T(\cdot \cap B)$$

pour probabilité initiale (à l'instant T). Notons aussi P^T la projection canonique de $P_{m,\beta}$ sur $C([T, +\infty[; M)$.

Le résultat précédent sur le comportement des trajectoires implique que pour tout $\varepsilon > 0$, il existe un $T > 0$ tel que

$$\left\| P^T - m_T(R_1)Q^{R_1,T} - \sum_{i \in I_1} m_T(K_i)Q^{K_i,T} \right\| \leq \varepsilon.$$

Ceci nous permet, comme dans le cas discret, de considérer une récurrence sur S pour obtenir que p.s. la trajectoire d'un k -algorithme finit par rester dans l'un des K_i , avec $1 \leq i \leq R$.

De plus, en reprenant la démonstration de la proposition 8 (voir aussi [9]) on voit que pour $1 \leq i \leq R$, la quantité $m_t(K_i)$ tend en temps grand vers une limite strictement positive λ_i .

Le théorème 9 découle alors facilement du fait que pour tout $\varepsilon > 0$, il existe un $T > 0$ tel que

$$\left\| P^T - \sum_{1 \leq i \leq R} m_T(K_i) Q^{K_i, T} \right\| \leq \varepsilon$$

et du résultat classique de convergence des algorithmes admissibles (qui dit que $\|m_t - \mu_{M, \beta_t}\|$ tend vers zéro en temps grand, pour un algorithme du recuit simulé pour lequel l'évolution de l'inverse de la température est donnée par $\beta_t = k^{-1} \ln(A + t)$, avec $A \geq 1$ et $k > c(M)$).

6. Récurrence locale et transition de phases

Soit $k \notin \mathcal{K}$ fixé et $(A_i)_{1 \leq i \leq R}$ une décomposition de M (qui est ici, soit une variété riemannienne compacte et connexe, soit un graphe fini irréductible et réversible) en k -cellules. On a vu dans les sections précédentes que la trajectoire d'un k -algorithme finit p.s. (sous $P_{m, \beta}$) par rester dans une k -cellule. En fait, la méthode utilisée permet d'obtenir un résultat un peu plus précis : Notons pour $1 \leq i \leq R$, Ω_i l'ensemble des trajectoires qui finissent par rester dans A_i , et \hat{A}_i la composante connexe de $\{x \in A_i \mid U(x) \leq f(A_i) + k\}$ qui contient N_{A_i} (qui est bien inclus dans l'une des composantes connexes, du fait que $c(A_i) < k$). Alors pour tout voisinage V_i de \hat{A}_i donné (dans le cas fini, M est muni de la topologie discrète), on peut prouver, en reprenant les démonstrations précédentes, que les trajectoires de Ω_i finissent par rester p.s. dans V_i , c'est-à-dire que p.s., l'ensemble des valeurs d'adhérence en temps grand d'une trajectoire de Ω_i est inclus dans \hat{A}_i .

On peut se demander s'il est possible d'améliorer ce résultat, en redécoupant par exemple la cellule A_i en d'autres sous-parties "stables" en temps grand. La réponse est négative : soit \tilde{A}_i la fermeture de la composante connexe de l'ensemble $\{x \in A_i \mid U(x) < f(A_i) + k\}$ qui contient N_{A_i} , la proposition suivante (on se ramène au cas $k > c(M)$ par le procédé utilisé à la fin des sections 4 et 5) permet de voir que p.s., l'ensemble des valeurs d'adhérence en temps grand d'une trajectoire de Ω_i contient \tilde{A}_i . Ainsi un phénomène de récurrence locale apparaît, qui montre que dans le cas continu, on n'a jamais convergence p.s. d'un k -algorithme. Par contre, dans le cas discret, on a convergence p.s. si k est strictement inférieur à la plus

petite différence de potentiel entre un élément de $\bigcup_{1 \leq i \leq R} N_{A_i}$ et un de ses voisins (pour q) qui n'y appartient pas. Cette propriété est notamment satisfaite pour k suffisamment petit, mais remarquons que le fait qu'elle soit satisfaite pour un $k_1 > 0$ n'entraîne pas qu'elle le soit pour tout $k \in]0, k_1[$ (et $k \notin \mathcal{K}$).

PROPOSITION 12. — Soit $(X_t)_{t \geq 0}$ un k -algorithme avec $k > c(M)$, soit \widetilde{M} la composante connexe de $\{x \in M \mid U(x) < k\}$ qui contient N_M (on suppose que la condition de normalisation $\min U = 0$ est satisfaite), soit $x_0 \in \widetilde{M}$ et soit V un voisinage de x_0 . Alors p.s., pour tout $t > 0$, il existe $t' > 0$ tel que $X_{t'} \in V$.

Démonstration. — Supposons par l'absurde que $P_{m,\beta}(\widetilde{\Omega}) > 0$, où $\widetilde{\Omega}$ est l'ensemble des trajectoires qui finissent par ne plus visiter V . Commençons par traiter le cas continu.

Quitte à réduire V , on peut supposer qu'il est fermé et que $\varepsilon \stackrel{\text{def}}{=} k - \max_{x \in V} U(x) > 0$. Soit $\delta = \varepsilon \wedge (k - c(M))$. Il est facile de construire un potentiel régulier \widetilde{U} sur M tel que

$$\begin{aligned} \forall x \notin V, \quad \widetilde{U}(x) &= U(x), \\ \forall x \in V \setminus \{x_0\}, \quad -\frac{\delta}{2} &< \widetilde{U}(x) \leq k - \varepsilon, \\ U(x_0) &= -\frac{\delta}{2}. \end{aligned}$$

Remarquons que $c(M, \widetilde{U}) < k$ et que x_0 est l'unique minimum global de \widetilde{U} . Notons $\widetilde{P}_{m,\beta}$ la probabilité sur (Ω, \mathcal{F}) correspondant à ce nouveau potentiel et $(\widetilde{m}_t)_{t \geq 0}$ la famille associée des lois des positions du processus canonique $(X_t)_{t \geq 0}$. D'après un résultat classique on a

$$\lim_{t \rightarrow +\infty} \widetilde{m}_t(V) = 1. \quad (\text{C})$$

Cependant, l'hypothèse $P_{m,\beta}(\widetilde{\Omega}) > 0$ implique qu'il existe un ouvert non vide $W \subset M$ et un temps $T > 0$ tels que pour tout $x \in W$ on ait

$$P_{m,\beta}(\forall t \geq T, X_t \notin V \mid X_T = x) > 0.$$

On en déduit que pour tout $x \in W$ on a

$$\widetilde{P}_{m,\beta}(\forall t \geq T, X_t \notin V \mid X_T = x) > 0$$

(en effet, si $x \notin V$ et si τ est le temps d'entrée dans V après T , alors le processus $(X_{t \wedge \tau})_{t \geq T}$ a même loi sous $P_{m,\beta}(\cdot \mid X_T = x)$ et sous $\tilde{P}_{m,\beta}(\cdot \mid X_T = x)$).

Ainsi, puisque $\tilde{m}_T(W) > 0$, il en découle que $\tilde{P}_{m,\beta}(\tilde{\Omega}) > 0$, ce qui est en contradiction avec (C).

Intéressons-nous maintenant au cas discret, pour lequel il suffit de prendre $V = \{x_0\}$.

On rajoute à M un point Δ et on considère $M' = M \sqcup \{\Delta\}$, muni du noyau de probabilités de transitions q' défini par

$$q'(x, y) = \begin{cases} q(x, y) & \text{si } x \in M \setminus \{x_0\} \text{ et } y \in M \\ \frac{1}{2} q(x_0, y) & \text{si } x = x_0 \text{ et } y \in M \setminus \{x_0\} \\ \frac{1}{2} \sum_{y \in M \setminus \{x_0\}} q(x_0, y) & \text{si } x = x_0 \text{ et } y = \Delta \\ q'(x_0, \Delta) & \text{si } x = \Delta \text{ et } y = x_0 \\ 0 & \text{si } x = \Delta \text{ et } y \in M \setminus \{x_0\} \\ q(x_0, x_0) & \text{si } x = y = x_0 \\ 1 - q'(\Delta, x_0) & \text{si } x = y = \Delta. \end{cases}$$

Il est clair que q' est irréductible et réversible par rapport à la mesure π définie par

$$\pi(x) = \begin{cases} \mu(x) & \text{si } x \in M \setminus \{x_0\} \\ 2\mu(x_0) & \text{si } x = x_0 \text{ ou si } x = \Delta \end{cases}$$

et donc aussi par rapport à la probabilité $\mu'(\cdot) = \pi(M')^{-1} \pi(\cdot)$.

Soit $\delta = (k - U(x_0)) \wedge (k - c(M))$. On considère le potentiel \tilde{U} défini sur M' par

$$\tilde{U}(x) = \begin{cases} U(x) & \text{si } x \in M \\ -\frac{\delta}{2} & \text{si } x = \Delta. \end{cases}$$

Remarquons que $c(M', \tilde{U}) < k$ et que Δ est l'unique minimum global de \tilde{U} sur M' . On conclut alors de la même manière que dans le cas continu. \square

La seconde question que l'on peut se poser est de savoir si la loi limite (i.e. les coefficients $(\lambda_i)_{1 \leq i \leq R}$ qui apparaissent dans les théorèmes 1 et 11) dépend de la loi initiale.

Il est commode d'élargir la définition d'un k -algorithme; on dira qu'un algorithme du recuit simulé est un (k, A) -algorithme, où $k > 0$ et $A \geq 1$, si pour tout $t \geq 0$, on a

$$\beta_t = k^{-1} \ln(A + t).$$

Soit $k \notin \mathcal{K}$ fixé et $(A_i)_{1 \leq i \leq R}$ une décomposition de M en k -cellules. On suppose de plus que $k < c(\bar{M})$, ce qui assure que $R > 1$.

Soit $A \geq 1$ et m une probabilité (initiale) sur M . On note $m_t^{(A)}$ son image à l'instant $t \geq 0$ par un (k, A) -algorithme. En reprenant les démonstrations des sections précédentes, il est clair qu'il existe pour $1 \leq i \leq R$ des réels $0 < \lambda_i(A, m) < 1$ tels que

$$\lim_{t \rightarrow +\infty} \left\| m_t^{(A)} - \sum_{1 \leq i \leq R} \lambda_i(A, m) \mu_{A_i, \beta_t} \right\| = 0.$$

On convient d'écrire $\lambda_i(A, x)$ pour $\lambda_i(A, \delta_x)$. Notons que l'application $m \mapsto \lambda_i(A, m)$ est affine : pour toute probabilité m , on a

$$\lambda_i(A, m) = \int \lambda_i(A, x) m(dx).$$

La proposition suivante montre que ces applications $\lambda_i(A, \cdot)$ ne sont jamais constantes, si $k < c(M)$ (et $k \notin \mathcal{K}$). Il apparaît donc l'équivalent d'un phénomène de transition de phases; le processus conserve une mémoire de sa distribution initiale par l'intermédiaire des $(\lambda_i(A, \cdot))_{1 \leq i \leq R}$.

PROPOSITION 13. — *Pour $A \geq 1$ et $1 \leq i \leq R$ fixés, l'application*

$$M \ni x \longmapsto \lambda_i(A, x) \in]0, 1[$$

n'est jamais constante.

Pour la démonstration, on distingue la situation où M est un graphe fini de celle où M est une variété compacte.

Le cas discret

Rappelons que f_1 est la valeur la plus élevée des fonds des k -cellules de la décomposition choisie. En reprenant la procédure de récurrence décrite dans la section 4, le lemme suivant permet de voir que pour tout $1 \leq i \leq R$ et tout $x \in N_{A_i}$, on a

$$\lim_{A \rightarrow +\infty} P_{\delta_x}^{(A)} [\exists t \geq 0 \mid X_t \notin A_i] = 0$$

où $P_{\delta_x}^{(A)}$ désigne la loi (sur (Ω, \mathcal{F})) d'un (k, A) -algorithme dont la loi initiale est δ_x .

LEMME 14. — Soit $1 \leq i \leq R$ tel que $f(A_i) = f_1$. Si $x \in N_{A_i}$, on a

$$\lim_{A \rightarrow +\infty} P_{\delta_x}^{(A)} [\exists t \geq 0 \mid X_t \notin A_i] = 0$$

et si $x \notin N_{A_i}$ et $U(x) \leq f_1$, on a

$$\lim_{A \rightarrow +\infty} P_{\delta_x}^{(A)} [\exists t \geq 0 \mid X_t \in A_i] = 0.$$

Démonstration. — Soit $x \in M$ tel que $U(x) \leq f_1$. On considère pour $p > 1$ et $t \geq 0$,

$$F_p^{(A)}(t) = \sum_{y \in M} \left(\frac{(\delta_x)_t^{(A)}}{\mu_{\beta_t}}(y) \right)^p \mu_{\beta_t}(y)$$

(où il est évidemment sous-entendu que $\beta_t = k^{-1} \ln(A + t)$).

La démonstration de la proposition 6 permet de se rendre compte qu'il existe une constante $K'(p) > 0$ indépendante de $A \geq 1$, telle que pour tout $t \geq 0$ on ait

$$F_p^{(A)}(t) \leq F_p^{(A)}(0) + K'(p) \exp((p-1)f_1\beta_t).$$

Cependant par choix de x , il existe une constante $K''(p) > 0$ telle que

$$F_p^{(A)}(0) \leq K''(p) \exp((p-1)f_1\beta_0) \leq K''(p) \exp((p-1)f_1\beta_t)$$

pour tout $t \geq 0$. Ainsi, en prenant $K(p) = K'(p) + K''(p)$, on a pour tout $A \geq 1$ et tout $t \geq 0$,

$$F_p^{(A)}(t) \leq K(p) \exp((p-1)f_1\beta_t).$$

De ce résultat on déduit, à l'aide d'inégalités de Hölder pour un bon choix de p , que pour $x \in M$ tel que $U(x) \leq f_1$ et $1 \leq i \leq R$ tel que $f(A_i) = f_1$,

$$\lim_{A \rightarrow +\infty} P_{\delta_x}^{(A)} \left[\int_0^\infty |L_{\beta_t} 1_{A_i}|(X_t) dt \right] = 0$$

et

$$\lim_{A \rightarrow +\infty} P_{\delta_x}^{(A)} \left[\int_0^\infty \Gamma_{\beta_t}(1_{A_i}, 1_{A_i})(X_t) dt \right] = 0;$$

ce qui implique facilement que

$$\lim_{A \rightarrow +\infty} P_{\delta_x}^{(A)} \left[\exists t \geq 0 \mid |1_{A_i}(X_t) - 1_{A_i}(X_0)| > \frac{1}{2} \right] = 0 ,$$

propriété qui peut encore se récrire sous la forme énoncée dans le lemme. \square

Le résultat de convergence qui précède le lemme montre notamment que pour tous $1 \leq i, j \leq R$, on a

$$\forall x \in N_{A_j} , \quad \lim_{A \rightarrow +\infty} \lambda_i(A, x) = \delta_{i,j}$$

où $\delta_{i,j}$ est le symbole de Kronecker.

Ainsi pour A assez grand, les applications $x \mapsto \lambda_i(A, x)$ ne sont jamais constantes sur M .

Pour étendre ce résultat à tout $A \geq 1$, on cherche à voir comment la fonction $\lambda_i(A, m)$ évolue avec A (à $1 \leq i \leq R$ et m fixés).

Notons que l'application λ_i satisfait, pour tout $t \geq 0$

$$\lambda_i(A, m) = \lambda_i(A + t, m_t^{(A)}) .$$

Ainsi, pour $t \geq 0$

$$\begin{aligned} \lambda_i(A + t, m) - \lambda_i(A, m) &= \lambda_i(A + t, m) - \lambda_i(A + t, m_t^{(A)}) \\ &= \int \lambda_i(A + t, x) (m - m_t^{(A)})(dx) \\ &\leq \|m - m_t^{(A)}\| , \end{aligned}$$

ce qui permet de voir que la fonction $[1, +\infty[\ni A \mapsto \lambda_i(A, m)$ est continue. Écrivons alors pour $t > 0$

$$\begin{aligned} \frac{\lambda_i(A + t, m) - \lambda_i(A, m)}{t} &= \\ &= \int \lambda_i(A, x) \frac{m - m_t^{(A)}}{t} (dx) + \\ &\quad + \sum_{x \in M} (\lambda_i(A + t, x) - \lambda_i(A, x)) \frac{m(x) - m_t^{(A)}(x)}{t} . \end{aligned}$$

Cependant, du fait que pour tout $x \in M$,

$$\lim_{t \rightarrow 0^+} \frac{m(x) - m_t^{(A)}(x)}{t} \text{ existe ,}$$

on a

$$\lim_{t \rightarrow 0^+} \sum_{x \in M} (\lambda_i(A, x) - \lambda_i(A + t, x)) \frac{m(x) - m_t^{(A)}(x)}{t} = 0.$$

D'autre part,

$$\lim_{t \rightarrow 0^+} \int \lambda_i(A, x) \frac{m_t^{(A)} - m}{t} (dx) = \int \tilde{L}_A [\lambda_i(A, \cdot)](x) m(dx)$$

où on a convenu de noter $\tilde{L}_A = L_{k^{-1} \ln(A)}$, le générateur à la température $\beta_0^{-1} = k \ln^{-1}(A)$.

On obtient donc que la fonction $[1, +\infty[\ni A \mapsto \lambda_i(A, m)$ est dérivable (à droite, mais il est facile de voir qu'elle est aussi dérivable à gauche sur $]1, +\infty[$ et de même dérivée) et de dérivée $-\int \tilde{L}_A [\lambda_i(A, \cdot)](x) m(dx)$. Ainsi, on a

$$\forall x \in M, \quad \frac{d}{dA} \lambda_i(A, x) = \tilde{L}_A [\lambda_i(A, \cdot)](x).$$

Considérons $(\lambda_i(A, x))_{x \in M}$ comme un vecteur de \mathbb{R}^M , les équations ci-dessus forment alors un système d'équations différentielles linéaires du premier ordre (à coefficients non constants) pour lequel il est bien connu que l'on a unicité de la solution si l'on se donne $(\lambda_i(A_0, x))_{x \in M}$ pour un $A_0 \geq 1$ fixé.

Or, si pour un $A_0 \geq 1$ l'application $M \ni x \mapsto \lambda_i(A, x)$ devait être constante, on aurait que pour tout $x \in M$, $\tilde{L}_{A_0} [\lambda_i(A_0, \cdot)](x) = 0$ et on en déduirait, par unicité de la solution, que pour tout $A \geq 1$, $\lambda_i(A, \cdot) = \lambda_i(A_0, \cdot)$. En particulier, on aurait que pour tout $A \geq 1$, l'application $M \ni x \mapsto \lambda_i(A, x)$ est constante. Mais on a vu que ceci est impossible, ce qui termine de montrer la proposition 13 dans le cas discret. \square

Le cas continu

Comme précédemment, soit $f_1 = \max_{1 \leq i \leq R} f(A_i)$. Notons également

$$\delta = \left(k - \max_{1 \leq i \leq R} c(A_i) \right) \wedge \left(\min_{1 \leq i \leq R} h(A_i) - k \right).$$

Pour $1 \leq i \leq R$, soit m_i une proposition sur M qui admet une densité régulière par rapport à λ et dont le support est inclus dans $\{x \in A_i \mid U(x) \leq f_1 + (\delta/2)\}$.

Une démonstration similaire à celle du lemme 14 (pour laquelle il faut reprendre la proposition 13) et la procédure de récurrence décrite dans la section 5 permettent de voir que pour tout $1 \leq i \leq R$, on a

$$\lim_{A \rightarrow +\infty} P_{m_i}^{(A)} [\exists t \geq 0 \mid X_t \notin A_i] = 0$$

où $P_{m_i}^{(A)}$ désigne la loi (sur (Ω, \mathcal{F})) d'un (k, A) -algorithme dont la loi initiale est m_i .

On en déduit que pour tous $1 \leq i, j \leq R$,

$$\lim_{A \rightarrow +\infty} \lambda_i(A, m_j) = \delta_{i,j} ;$$

ce qui montre que pour A assez grand, les applications $x \mapsto \lambda_i(A, x)$ ne sont jamais constantes sur M .

Pour continuer à suivre la méthode utilisée dans le cas discret, il faut tout d'abord vérifier que les λ_i sont réguliers en $(A, x) \in [1, +\infty[\times M$. Considérons donc pour $1 \leq i \leq R$, une application $\varphi_i \in C^\infty(M)$ prenant la valeur 1 sur $\{x \in A_i \mid U(x) \leq f(A_i) + k + (\delta/4)\}$ et s'annulant sur le complémentaire de $\{x \in A_i \mid U(x) \leq f(A_i) + k + (\delta/2)\}$, et posons, pour $T > 0$

$$\lambda_i(A, m, T) = m_T^{(A)}(\varphi_i).$$

Il est bien connu que l'application

$$]1, +\infty[\times M \times]0, +\infty[\longrightarrow \mathbb{R}$$

$$(A, x, T) \longmapsto \lambda_i(A, x, T) \quad (\stackrel{\text{déf}}{=} \lambda_i(A, \delta_x, T))$$

est de classe C^∞ , et on a clairement que

$$\lim_{T \rightarrow +\infty} \lambda_i(A, m, T) = \lambda_i(A, m)$$

(soit K un compact de $[1, +\infty[$, les démonstrations de la section 5 montrent en fait que cette limite est uniforme en m et $A \in K$, ce qui permet notamment de voir que l'application $(A, x) \mapsto \lambda_i(A, x)$ est continue sur $[1, +\infty[\times M$).

De plus, pour tout $0 \leq t < T$, on a

$$\lambda_i(A, m, T) = \lambda_i(A + t, m_t^{(A)}, T - t) .$$

Ainsi, pour $0 < t < T$,

$$\begin{aligned} & \frac{\lambda_i(A + t, m, T) - \lambda_i(A, m, T)}{t} = \\ &= \frac{\lambda_i(A + t, m, T) - \lambda_i(A + t, m_t^{(A)}, T)}{t} + \\ &+ \frac{\lambda_i(A + t, m_t^{(A)}, T) - \lambda_i(A + t, m_t^{(A)}, T - t)}{t} \\ &= \frac{1}{t} \int \lambda_i(A + t, x, T) (m - m_t^{(A)})(dx) + \frac{1}{t} \int \varphi_i(x) (m_{T+t}^{(A)} - m_T^{(A)})(dx) \\ &= \frac{1}{t} \mathbb{E}_m^{(A)} [\lambda_i(A + t, X_0, T) - \lambda_i(A + t, X_t, T)] + \\ &+ \frac{1}{t} \mathbb{E}_m^{(A)} [\varphi_i(X_{T+t}) - \varphi_i(X_T)] \\ &= -\frac{1}{t} \mathbb{E}_m^{(A)} \left[\int_0^t L_{\beta_s} [\lambda_i(A + t, \cdot, T)](X_s) ds \right] + \\ &+ \frac{1}{t} \mathbb{E}_m^{(A)} \left[\int_T^{T+t} L_{\beta_s} \varphi_i(X_s) ds \right] , \end{aligned}$$

expression qui converge, quand t tend vers 0^+ , vers

$$- \int L_{\beta_0} [\lambda_i(A + h, \cdot, T)](x) m(dx) + \int L_{\beta_T} \varphi_i(x) m_T^{(A)}(dx) .$$

Ainsi notamment pour $x \in M$, la dérivée de la fonction $]1, +\infty[\ni A \mapsto \lambda_i(A, x, T)$ est

$$-L_{\beta_0} [\lambda_i(A + h, \cdot, T)](x) + (\delta_x)_T^{(A)}(L_{\beta_T} \varphi_i) .$$

Cependant, du fait que $\mu_\beta(L_\beta \varphi_i) = 0$ et qu'il existe une constante $K > 0$ telle que pour tout $\beta \geq 0$ on ait

$$|L_\beta \varphi_i(x)| \leq K \beta 1_{\{x \in A_i, f(A_i) + k + \frac{\varepsilon}{4} \leq U(x) \leq f(A_i) + k + \frac{\varepsilon}{2}\}} ,$$

on obtient, en reprenant les calculs de la section 5 (notamment la proposition 13), que pour tous $A_1 \geq A_0 \geq 1$ fixés,

$$\lim_{T \rightarrow +\infty} (\delta_x)_T^{(A)}(L_{\beta_T} \varphi_i) = 0$$

uniformément en $x \in M$ et $A \in [A_0, A_1]$.

Il est alors facile d'en déduire que l'application $[1, +\infty[\times M \ni (A, x) \mapsto \lambda_i(A, x)$ est une solution faible de l'edp.

$$\partial_A \lambda_i = -\tilde{L}_A \lambda_i \quad (E)$$

où $\tilde{L}_A = L_{k^{-1} \ln(A)} (= L_{\beta_0})$ et où on a convenu de noter $\partial_A = \partial/\partial A$.

Remarquons que l'opérateur $-\tilde{L}_A$ est positif, ainsi on n'a pas a priori unicité (et régularité, ni même existence) de la solution de l'équation précédente, pour une condition initiale $\lambda_i(1, \cdot)$ donnée. Mais soient $1 \leq i \leq R$ et $A_0 > 1$ fixés, on considère pour $(A, x) \in [0, A_0 - 1] \times M$, $g(A, x) = \lambda_i(A_0 - A, x)$. Cette application est une solution faible de l'edp. parabolique

$$\partial_A g = \hat{L}_A g \quad (P)$$

où \hat{L}_A est l'unique opérateur fermé sur $L^2(\lambda)$ dont la restriction à $C^2(M)$ est $\Delta \cdot -\beta(A) \langle \nabla U, \nabla \cdot \rangle$ (on convient désormais de noter $\beta(A) = k^{-1} \ln(A_0 - A)$). Il est bien connu (voir [5]) que ceci implique que g est de classe C^∞ sur $]0, A_0 - 1] \times M$ et qu'elle est en fait une solution au sens usuel de l'équation ci-dessus, ce qui montre aussi que λ_i est une solution classique de (E).

De plus, il existe pour l'équation (P) un théorème d'unicité rétrograde (voir [5, p. 181], dont les hypothèses (i) et (ii) sont bien satisfaites ici, ou [13]) : si g_1 et g_2 sont des solutions de (P) telles que $g_1(A_0 - 1, \cdot) = g_2(A_0 - 1, \cdot)$ alors $g_1 = g_2$ sur $[0, A_0 - 1] \times M$. Cependant, toute fonction constante est solution de (P), ainsi, si $\lambda_i(1, \cdot)$ devait être constante sur M , alors il en serait de même pour l'application $\lambda_i(A_0, \cdot)$ pour tout $A_0 > 1$ (plus généralement, s'il existait $A_1 \geq 1$ tel que $\lambda_i(A_1, \cdot)$ est constante, alors par l'unicité classique de la solution de (P) si on se donne une condition initiale, on en déduirait que $\lambda_i(1, \cdot)$ est constante, puis que toutes les applications $\lambda_i(A, \cdot)$ sont constantes pour $A \geq 1$). Or, on a vu que ceci n'était pas possible, d'où la proposition 13 dans le cas continu. \square

Remarques sur l'équation rétrograde (E)

Soient v_1, \dots, v_R des réels non tous nuls, le début de la démonstration précédente permet de voir que l'application $\sum_{1 \leq i \leq R} v_i \lambda_i(A, \cdot)$ n'est pas constante sur M , pour A suffisamment grand, ce qui montre notamment que les applications $(\lambda_i)_{1 \leq i \leq R}$, considérées sur $[1, +\infty[\times M$, sont linéairement indépendantes (plus précisément, on peut aussi en déduire, à l'aide de l'unicité rétrograde de (P), que les applications $(M \ni x \mapsto \lambda_i(A, x))_{1 \leq i \leq R}$ sont linéairement indépendantes, pour tout $A \geq 1$ fixé).

En fait, les applications $(\lambda_i)_{1 \leq i \leq R}$ forment une base de l'espace \mathcal{E}_b des solutions bornées de (E) sur $[1, +\infty[\times M$. En effet, soit g une solution bornée de E. Considérons, comme précédemment, une diffusion $(X_t)_{t \geq 1}$ sur M dont le générateur à l'instant t est \tilde{L}_t . L'équation (E) montre que $t \mapsto g(t, X_t)$ est une martingale. On a donc pour tous $t \geq s \geq 1$ et $x \in M$,

$$\begin{aligned} g(s, x) &= \mathbb{E}[g(t, X_t) \mid X_s = x] \\ &= \int g(t, y) (\delta_x)_{t-s}^{(s)}(dy) \end{aligned}$$

(en reprenant les notations introduites avant la proposition 13). Cependant, vu qu'en variation totale,

$$\lim_{t \rightarrow +\infty} \left\| (\delta_x)_{t-s}^{(s)} - \sum_{1 \leq i \leq R} \lambda_i(s, x) \mu_{A_i, k-1 \ln(t)} \right\| = 0$$

et que g est supposé bornée, on a

$$\lim_{t \rightarrow +\infty} \left| \int g(t, y) (\delta_x)_{t-s}^{(s)}(dy) - \sum_{1 \leq i \leq R} \lambda_i(s, x) \int g(t, y) \mu_{A_i, k-1 \ln(t)}(dy) \right| = 0.$$

D'autre part, notons que les expressions $\int g(t, x) \mu_{A_i, k-1 \ln(t)}(dy)$ sont bornées, pour $1 \leq i \leq R$, ainsi, il existe une suite $(t_n)_{n \geq 0}$ de réels strictement croissante vers l'infini, telle que pour tout $1 \leq i \leq R$,

$$\lim_{n \rightarrow +\infty} \int g(t_n, y) \mu_{A_i, k-1 \ln(t_n)}(dy) = v_i$$

où les v_i sont des réels (indépendants de $(s, x) \in [1, +\infty[\times M$).

En regroupant les résultats ci-dessus, on a donc

$$\begin{aligned} g(s, x) &= \lim_{n \rightarrow +\infty} \int g(t_n, x) (\delta_x)_{t_n-s}^{(s)}(dy) \\ &= \sum_{1 \leq i \leq R} v_i \lambda_i(s, x); \end{aligned}$$

ce qui montre que les $(\lambda_i)_{1 \leq i \leq R}$ engendrent \mathcal{E}_b et donc que $R = \dim(\mathcal{E}_b)$ (remarquons d'ailleurs que les $(\lambda_i)_{1 \leq i \leq R}$ sont exactement les points extrémaux de l'ensemble convexe compact $\tilde{\mathcal{E}}_b$ des solutions de (E) à valeurs dans $[0, 1]$).

On a considéré ci-dessus le cas continu, mais les remarques précédentes se généralisent immédiatement au cas discret.

On retrouve ainsi que le nombre R des k -cellules qui interviennent dans une décomposition de M (pour $k > 0$, $k \notin \mathcal{K}$ fixé) ne dépend pas de la décomposition choisie. Ce résultat découle plus simplement du fait que R est le nombre de k -cycles ou de manière équivalente du fait que la famille $(N_{A_i})_{1 \leq i \leq R}$ est indépendante de la décomposition $(A_i)_{1 \leq i \leq R}$ choisie.

Remarque. — Décrivons une démarche qui aurait évité la construction d'une partition en k -cellules (dans le cas discret, mais qui s'adapte facilement au cas continu).

Soit $k > 0$ fixé. Considérons les k -cycles $C_1, \dots, C_{R'}$, notons $N_1, \dots, N_{R'}$ les fonds de ces cycles et $x_1, \dots, x_{R'}$ des points y appartenant. À tout point $x \in M$, associons pour $1 \leq i \leq R'$, $h_i(x) = \min_{p \in C_{x, x_i}} e(p)$ et $h(x) = \min_{1 \leq i \leq R'} h_i(x)$. Remarquons que l'on a toujours $h(x) \leq k$. S'il existe $x \in M$ avec $h(x) = k$, on dit que $k \in \mathcal{K}$ (on vérifie facilement qu'il s'agit bien de l'ensemble considéré précédemment). Notons

$$A'_i = \{x \in M \mid h_i(x) = h(x)\}$$

et considérons une partition $M = \bigsqcup_{i=1}^{R'} A_i$ telle que pour tout $1 \leq i \leq R'$, on ait $A_i \subset A'_i$. Les ensembles ainsi construits ne sont pas nécessairement des k -cellules (ni même des sous-graphes connexes), néanmoins cette partition est suffisante pour pouvoir démontrer la proposition 4, si $k \notin \mathcal{K}$. Puis on montre, en reprenant la récurrence de la section 3 pour enlever les k -cycles les uns après les autres, qu'une trajectoire finit p.s. par rester dans l'un des C_i ou dans le complémentaire de $\bigcup_{i=1}^{R'} C_i$. Or ce dernier événement est de probabilité nulle, comme on peut le voir en utilisant la démonstration de la proposition 12.

Appendice

On va vérifier que si l'ensemble \mathcal{L} des classes d'équivalence de la relation \sim sur L est infini, alors l'ensemble \mathcal{K} des $k > 0$ pour lesquels il n'existe pas de décomposition de M en k -cellules admet zéro pour point d'adhérence.

Pour voir qu'un $k > 0$ appartient à \mathcal{K} , on utilisera le résultat suivant, qui termine aussi la démonstration de la proposition 10.

LEMME 15. — *S'il existe un fermé A de M tel que*

$$\inf_{\substack{x \in N_A \\ y \in D(A)}} \tilde{H}(x, y) = k > 0$$

alors $k \in \mathcal{K}$.

En fait la fermeture de A n'est pas vraiment nécessaire, car si A satisfait l'égalité précédente, le fermé $A' = \overline{N_A}$ la vérifie également (du fait que $k > 0$).

Démonstration. — Par continuité de \tilde{H} sur $M \times M$ et du fait que N_A est compact, il existe $x \in N_A$ et $y \in M$ tels que $\tilde{H}(x, y) = k$.

Par l'absurde, supposons qu'il existe une décomposition $(A_i)_{1 \leq i \leq R}$ de M en k -cellules, et soit $1 \leq i \leq R$ tel que $x \in A_i$. Commençons par voir que $x \in N_{A_i}$. En effet, supposons que $U(x) > f(A_i)$ et soit $z \in N_{A_i}$. L'élément z n'appartient alors pas à A , puisque $U(z) < U(x)$, d'où $z \in D(A)$, puis

$$\begin{aligned} H_{A_i}(x, z) &= \inf_{\phi \in \mathcal{C}_{x,z}^{A_i}} e(\phi) - U(x) - U(z) + f(A_i) \\ &\geq \inf_{\phi \in \mathcal{C}_{x,z}^M} e(\phi) - U(x) \\ &= \tilde{H}(x, z) \geq \inf_{\substack{x' \in N_A \\ y' \in D(A)}} \tilde{H}(x', y') = k, \end{aligned}$$

ce qui est en contradiction avec $c(A_i) < k$. Ainsi $f(A_i) = U(x)$. Le fait que $h(A_i) > k$ et $\tilde{H}(x, y) = k$ entraîne alors que $y \in A_i$, mais, comme ci-dessus, ceci est incompatible avec $c(A_i) < k$, car $H_{A_i}(x, y) \geq \tilde{H}(x, y) = k$. \square

Remarquons que pour $\ell \in \mathcal{L}$ et $y \in M$, les quantités $U(x_\lambda)$ et $\tilde{H}(x_\ell, y)$ ne dépendent pas du choix du représentant $x_\ell \in \ell$; on les notera dorénavant $U(\ell)$ et $\tilde{H}(\ell, y)$. Considérons

$$K(\ell) = \inf_{y \in D(\ell)} \tilde{H}(\ell, y)$$

(notons que cette quantité s'écrit aussi $\inf_{x \in N_\ell, y \in D(\ell)} \tilde{H}(x, y)$, car $N_\ell = \ell$).

Pour $\varepsilon > 0$ fixé, on a déjà vu que l'ensemble $\{\ell \in \mathcal{L} \mid K(\ell) \geq \varepsilon\}$ est fini, car les éléments de \mathcal{L} sont deux à deux disjoints. Ainsi, des deux situations suivantes, l'une au moins est satisfaite :

- soit il existe une suite $(\ell_n)_{n \geq 0}$ d'éléments de \mathcal{L} telle que la suite $(K(\ell_n))_{n \geq 0}$ décroît strictement vers zéro, auquel cas il suffit d'appliquer le lemme 15 avec les fermés ℓ_n pour obtenir que \mathcal{K} est infini ;
- soit il existe $\ell \in \mathcal{L}$ tel que $K(\ell) = 0$.

Il est facile de vérifier que pour tout $y \in D(\ell)$, $\tilde{H}(\ell, y) > 0$. Il existe donc une suite $(y_n)_{n \geq 0}$ d'éléments de $D(\ell)$ telle que la suite $(\tilde{H}(\ell, y_n))_{n \geq 0}$ décroît strictement vers zéro. Pour $n \geq 0$, notons

$$\begin{aligned} A(\ell, n) &= \{x \in D(\ell) \mid \tilde{H}(\ell, x) = \tilde{H}(\ell, y_n)\} \\ N(\ell, n) &= N_{A(\ell, n)} \\ D(\ell, n) &= D(A(\ell, n)). \end{aligned}$$

Remarquons qu'il existe $N \in \mathbb{N}$ tel que pour $n \geq N$, on a

$$\forall x \in N(\ell, n), \quad U(x) = U(\ell).$$

En effet, on a a priori pour $x \in N(\ell, n)$, $U(x) \leq U(y_n) \leq U(\ell)$, cependant s'il existait une suite strictement croissante d'entiers $(\varphi(n))_{n \geq 0}$ telle que pour tout $n \geq 0$ on ait

$$\forall x \in N(\ell, \varphi(n)), \quad U(x) < U(\ell)$$

on obtiendrait, puisque $\lim_{n \rightarrow \infty} \tilde{H}(\ell, y_{\varphi(n)}) = 0$, une contradiction avec $\ell \in \mathcal{L}$, par définition de \mathcal{L} .

Pour $n \geq N$, on a donc $y_n \in N(\ell, n)$ et $\ell \in D(\ell, n)$, d'où

$$\inf_{\substack{x \in N(\ell, n) \\ y \in D(\ell, n)}} \tilde{H}(x, y) \leq \tilde{H}(\ell, y_n).$$

D'autre part, si $y \in M$ est tel que $U(y) \leq U(x)$ et $\tilde{H}(x, y) < \tilde{H}(\ell, y_n)$ pour un $x \in N(\ell, n)$, alors $\tilde{H}(\ell, y) \leq \tilde{H}(\ell, y_n)$, cependant, si on suppose que $\tilde{H}(\ell, y) < \tilde{H}(\ell, y_n)$, l'inégalité $\tilde{H}(x, y) < \tilde{H}(\ell, y_n)$ implique que $\tilde{H}(x, \ell) < \tilde{H}(\ell, y_n)$, ce qui est absurde, d'où $y \in A(\ell, n)$. On en déduit que $x \in N(\ell, n)$ et $y \in D(\ell, n)$ impliquent que $\tilde{H}(x, y) \geq \tilde{H}(\ell, y_n)$, d'où

$$\inf_{\substack{x \in N(\ell, n) \\ y \in D(\ell, n)}} \tilde{H}(x, y) \geq \tilde{H}(\ell, y_n) ;$$

puis, pour $n \geq N$,

$$\inf_{\substack{x \in N(\ell, n) \\ y \in D(\ell, n)}} \tilde{H}(x, y) = \tilde{H}(\ell, y_n).$$

Pour obtenir que \mathcal{K} est infini il suffit donc d'appliquer le lemme 15 avec les $A(\ell, n)$.

Remerciements

Je tiens tout particulièrement à remercier Marc Arnaudon pour les discussions très constructives que nous avons eu sur ce sujet.

Références

- [1] CATONI (O.) . — *Sharp large deviations estimates for simulated annealing algorithms*, Annales de l'IHP **27** n° 3 (1991), pp. 291-383.
- [2] CONCORDET (D.) . — *Estimation de la densité du recuit simulé*, à paraître dans les Annales de l'I.H.P.
- [3] CHIANG (T. S.) et CHOW (Y.) . — *On the Convergence Rate of Annealing Processes*, SIAM J. Control and Optimization **26**, n° 6 (1988), pp. 1455-1470.
- [4] FREIDLIN (M.I.) et WENTZELL (A.D.) . — *Random Perturbations of Dynamical Systems*, Springer-Verlag, 1984.
- [5] FRIEDMAN (A.) . — *Partial Differential Equations*, R. Krieger Publishing Company, 1976.
- [6] GEMAN (S.) et GEMAN (D.) . — *Stochastic relaxation, Gibbs Distributions, and the Bayesian Restoration of Images* I.E.E.E. Transactions on Pattern Analysis and Machine Intelligence **6** (1984), pp. 721-741.
- [7] GÖTZE (F.) . — *Rate of convergence of simulated annealing processes*, preprint, Universität Bielefeld, December 1991.
- [8] HAJEK (B.) . — *Cooling Schedules for Optimal Annealing*, Mathematics of Operations Research **13**, n° 2 (1988), pp. 311-329.
- [9] HOLLEY (R.), KUSUOKA (S.) et STROOCK (D.) . — *Asymptotics of the Spectral Gap with Applications to the Theory of Simulated Annealing*, J.F.A. **83** (1989), pp. 333-347.
- [10] HOLLEY (R.) et STROOCK (D.) . — *Annealing via Sobolev Inequalities*, C.M.P. **115** (1988), pp. 553-569.
- [11] HWANG (C. R.) . — *Laplace's Method Revisited : Weak Convergence of Probability Measures*, The Annals of Probability **8**, n° 6 (1980), pp. 1177-1182.
- [12] KIRKPATRICK (S.), GELATT (C. D.) et VECCHI (M. P.) . — *Optimization by simulated annealing*, Science **220** (1983), pp. 621-680.

- [13] LIONS (J.-L.) et MALGRANGE (B .) . — *Sur l'unicité rétrograde dans les problèmes mixtes paraboliques*, Math. Scand. **8** (1960), pp. 277-286.
- [14] MICLO (L.) . — *Recuit simulé sans potentiel sur un ensemble fini*, Séminaire de Probabilités XXVI, LNM 1526 (1992), pp. 47-60.
- [15] MICLO (L.) . — *Comportement de spectres d'opérateurs de Schrödinger à basse température*, à paraître dans le Bulletin des Sciences Mathématiques.
- [16] TROUVÉ (A.) . — *Parallélisation massive du recuit simulé*, Thèse de doctorat, Université Paris 11, janvier 1993.