Centre de Mathématiques Laurent Schwartz
Ecole Polytechnique

SEMINAIRE

Equations aux Dérivées Partielles

2006-2007

Gilles Lebeau et Persi Diaconis

Metropolis : ***Le jour où l'étoile probabilité entra dans le champ gravitationnel de la galaxie microlocale***

Séminaire É. D. P. (2006-2007), Exposé n° XIV, 11 p.

<http://sedp.cedram.org/item?id=SEDP_2006-2007____A14_0>

U.M.R. 7640 du C.N.R.S.
F-91128 PALAISEAU CEDEX
Fax : 33 (0)1 69 33 49 49
Tél : 33 (0)1 69 33 49 99

Metropolis

Le jour où l'étoile probabilité entra dans le champ gravitationnel de la galaxie microlocale

STARING : GILLES LEBEAU (†), DIRECTED BY : PERSI DIACONIS (†, ‡)

† Département de Mathématiques,
Université de Nice Sophia-Antipolis
Parc Valrose 06108 Nice Cedex 02, France
‡ Dept. of Math.
Stanford University
USA

`A GILPER Entertainment Production Company`

Résumé

Ceci n'est pas une oeuvre de fiction. Cependant, toute ressemblance avec des théorèmes connus serait purement fortuite et le fruit du hasard.

We prove sharp rates of convergence to stationarity for a simple case of the Metropolis algorithm : the placement of a single disc of radius h randomly into the interval $[-1-h, 1+h]$. We find good approximations for the top eigenvalues and eigenvectors. The analysis gives rigorous proof for the careful numerical work in [DN04]. The micro-local techniques employed offer promise for the analysis of more realistic problems.

Table des matières

1 Introduction et résultat

L'algorithme de Metropolis a été introduit par N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller et E. Teller en 1953 ([MRR$^+$53]). Leur motivation était de trouver une méthode pour placer n disques disjoints au hasard dans un domaine borné Ω. Leur algorithme a été généralisé par W. Hastings ([Has70]), et est devenu un des outils les plus utilisés du calcul scientifique.

L'algorithme de [MRR$^+$53] est simple à décrire. Soit $\underline{x} = (x_1, ..., x_n)$ une configuration admissible des centres des disques. On choisit un disque i au hasard avec probabilité uniforme $1/n$, et on déplace son centre x_i au hasard dans Ω (pour la mesure de Lebesgue). Le centre du $i$ème disque est alors en y_i. Si la nouvelle configuration $\underline{y} = (x_1, ..., y_i, ..., x_n)$ est admissible, c'est à dire si le nouveau $i$ème disque n'intersecte pas les autres et reste contenu dans Ω, alors la nouvelle configuration est $\underline{y}$, sinon on conserve la configuration initiale $\underline{x}$. En itérant ce procédé, on obtient, à partir d'une configuration initiale quelconque $\underline{x}^0$, une suite de configurations $\underline{x}^0, \underline{x}^1, \underline{x}^2, ..., \underline{x}^N, ...$ qui converge vers une configuration aléatoire.

A. Algorithme de Metropolis

Soit $(X, \mathcal{B}, \mu)$ un espace mesuré fini. Soit $f(x) > 0$, $\int f(x) d\mu(x) = 1$, une densité de probabilité sur X. On se donne un noyau de Markov symétrique de densité $p(x, y) = p(y, x) \geq 0$ pour tout x, y, et $\int p(x, y)\mu(dy) = 1$ pour tout x. Ce noyau auxiliaire p permet de construire le noyau de Metropolis comme suit

$$\begin{aligned} M(x, dy) &= m(x)\delta_x + p(x, y) min\big(\frac{f(y)}{f(x)}, 1\big)\mu(dy) \\ m(x) &= \int_{\{z: f(z) < f(x)\}} \big(1 - \frac{f(z)}{f(x)}\big) p(x, z)\mu(dz) \end{aligned} \tag{1.1}$$

L'interprétation algorithmique de la formule 1.1 est simple. On part de x, on choisit y celon la loi de probabilité $p(x, y)\mu(dy)$. Si $f(y) \geq f(x)$, on va en y. Si $f(y) < f(x)$ on lance une pièce avec probabilité de tomber sur face égale à $f(y)/f(x)$. Si la pièce retombe sur face, on va en y. Si elle retombe sur pile, on reste en x. Le point essentiel est que pour faire tourner cet algorithme dans la pratique, où par exemple f est de la forme $f(x) = Z^{-1} e^{-\beta H(x)}$, on n'a pas besoin de connaitre la constante de normalisation Z, très souvent incalculable, puisqu'il suffit de toute évidence de connaitre β et $H(x)$.

Pour une introduction au sujet, on pourra consulter ([HH64]), et on renvoie à ([BD01]) pour une interprétation géométrique de l'algorithme. On trouvera des exemples d'applications en physique dans ([BH02]), et en biologie et statistique dans ([Liu01]).

B. Convergence

Soit $L^2(f)$ l'espace des fonctions réelles de carré intégrable pour la mesure $f(x)\mu(dx)$. Le noyau de metropolis M opère sur $L^2(f)$ par

$$Mg(x) = m(x)g(x) + \int g(y)p(x, y) min\big(\frac{f(y)}{f(x)}, 1\big)\mu(dy) \tag{1.2}$$

Des arguments élémentaires montrent que $M : L^2(f) \to L^2(f)$ est autoadjoint et contractant. Les itérés de M sont définis comme d'habitude par

$$(M^k g)(x) = \int (M^{k-1}g)(z)M(x,dz) = \int g(y)M_x^k(dy) \tag{1.3}$$

Dans l'exemple très simple que l'on va étudier, on aura $X = \mathbb{R}$, μ est la mesure de Lebesgue et la fonction f vaut $1/2$ sur $[-1,1]$ et 0 ailleurs (le fait que $\mu(\mathbb{R}) = \infty$ est compensé par le fait que f s'annule hors de $[-1,1]$, ces détails n'ont aucune importance). On travaille dans un cadre un peu général concernant le choix du déplacement aléatoire de chaque étape élementaire. On se donne $\varphi(x) = \varphi(-x) \geq 0$ une fonction C^∞ sur $[-1,1]$ telle que $\int_{-1}^1 \varphi(x)dx = 1$ et $\varphi(\pm 1) \neq 1$. Alors $\varphi(x)dx$ est une densité de probabilité sur $[-1,1]$ par rapport à la mesure de Lebesgue . On se donne un petit paramètre h, et on pose

$$\varphi_h(x) = \frac{1}{h}\varphi(x/h), \quad -h < x < h \tag{1.4}$$

Soit $p_h(x,y) = \varphi_h(x-y)$. Alors dans notre exemple, le noyau de metropolis est l'opérateur K_h opérant sur $L^2([-1,1])$

$$\begin{aligned} K_h(g)(x) &= m_h(x)g(x) + \int_{[-1,1]} g(y)\varphi_h(x-y)dy \\ m_h(x) &= \int_{\{y<-1\}} \varphi_h(x-y)dy + \int_{\{y>1\}} \varphi_h(x-y)dy \end{aligned} \tag{1.5}$$

Le petit paramètre h mesure la taille du déplacement de chaque étape de l'algorithme, et le problème est d'estimer la vitesse de convergence de l'algorithme vers la mesure stationnaire, en fonction de h. L'intuition élémentaire dit que, partant d'un point $x \in [-1,1]$ quelconque, il faut $\mathcal{O}(h^{-2})$ itérations pour construire un point de $[-1,1]$ en position générale. La difficulté est que si $dist(x, \pm 1) \leq h$, on peut passer beaucoup de temps à faire du "sur-place", et ce phénomène se traduit par des difficultés analytiques non triviales.

Si on revient au cas général de l'algorithme 1.2, alors sous des hypothèses simples et naturelles sur $p(x,y)$, vérifiées en particulier par 1.5, on a le résultat de convergence

$$\text{pour tout } x, \quad \|M_x^k - fd\mu\|_{TV} \to 0 \quad (k \to \infty) \tag{1.6}$$

Ici la norme de variation totale est définie par

$$\|M_x^k - fd\mu\|_{TV} = \sup_A |M_x^k(A) - \int_A f(y)d\mu(y)| \tag{1.7}$$

où le sup est pris sur les parties mesurable A.

Dans les applications, par exemple, pour le problème de placement aléatoire de disques, ou pour les simulations sur réseaux comme le modèle de Ising, il est important d'avoir de l'information sur la vitesse de convergence dans 1.6. On dispose de plusieurs types de majoration. Peut être la plus faible est "l'ergodicité géometrique" , c'est à dire

$$\|M_x^k - fd\mu\|_{TV} \leq a(x)\gamma^k \tag{1.8}$$

avec $a(x) > 0$ et $|\gamma| < 1$. Si on ne sait rien dire de $a(x)$ et γ, 1.8 n'est guère plus précis dans la pratique que 1.6. Des estimées plus précises ont été développées dans des cas particuliers dans ([JH01]), . On y obtient des "bornes honnêtes " dans la mesure où $a(x) > 0$ et γ sont explicites. Bien sur, ces majorations peuvent ne pas être optimales, de sorte qu'on s'intéresse aussi à la question de minoration de la vitesse de convergence, avec si possible un encadrement qui fournisse un "équivalent" à constantes près. On trouvera une revue sur les estimations de type optimal pour les algorithmes de metropolis sur les ensembles finis dans ([DSC98]). Les techniques analytiques employées y sont très variées (inégalités de Poincaré, Cheeger, Nash, Sobolev et log-Sobolev).

Pour les problèmes continus, on sait très peu de choses. La raison semble être qu'il peut s'agir de problèmes durs (et passionnant) de théorie spectrale de type semi-classique, et que les probabilistes ne connaissent pas l'analyse microlocale (et réciproquement, d'ailleurs).

La première question est d'estimer le "gap" i.e l'écart spectral entre la valeur propre 1 et le sup $\lambda_1(h)$ du reste du spectre (1 est toujours valeur propre, associée à la mesure d'équilibre limite, qui est l'état vers lequel doit converger l'algorithme ; pour les fanatiques de l'équation de Schrödinger, c'est l'analogue de l'état fondamental). Dans notre problème 1.5, $\lambda_1(h)$ est aussi une valeur propre, le spectre de K_h étant discret hors de $[0, 1/2]$. On trouvera dans ([Kie00]), ([MR00]), ([AT87]), des estimations sur le trou spectral, pour la chaîne de Markov 1.5

$$Ah^2 \leq gap \leq Bh^2 \quad A, B \quad \text{explicites} \tag{1.9}$$

On obtient ici un équivalent (voir le théorème 2.7, formule 2.35)

$$1 - \lambda_1(h) = gap = \frac{\alpha h^2 \pi^2}{8} + \mathcal{O}(h^3), \quad \alpha = \int_{-1}^{1} z^2 \varphi(z) dz \tag{1.10}$$

La preuve donne aussi l'asymptotique de la première fonction propre non triviale ψ

$$\psi(x) = \cos(\eta_{1,h} \frac{x+1}{h}) + \frac{b(\hat{\varphi}(\eta_{1,h}))}{\eta_{1,h}} \sin(\eta_{1,h} \frac{x+1}{h}) + r_h(x) \tag{1.11}$$

où $\eta_{1,h}$ est défini dans le théorème 2.7 et la b-fonction dans le théorème 2.5. On a

$$\begin{aligned} &\eta_{1,h} = h\pi/2 + \mathcal{O}(h^2), \quad b(1) = 0 \\ &|r_h(x)| \leq c_1 |\lambda_1(h) - 1| \exp(-\frac{c_2}{h} dist(x, \{[-1, -1+h] \cup [1-h, 1]\})) \end{aligned} \tag{1.12}$$

pour des constantes universelles $c_1, c_2 > 0$. Le théorème 2.7 donne l'asymptotique complète par rapport au paramètre h pour les fonctions propres associées aux valeurs propres dans l'intervalle $[1-\delta, 1]$, uniformément en h. Cette étude précise de la théorie spectrale permet de prouver le théorème de convergence suivant pour la chaîne de Metropolis 1.5 sur $[-1, 1]$.

Théorème 1.1 *Soit $h_0 > 0$ petit. Il existe $B > 0, C > 0$ et A, A' tels que pour tout $h \in]0, h_0]$, on a :*

$$C \leq \min_{x \in [-1,1]} \|M_x^n - f d\mu\|_{TV} \quad \textit{pour } n \leq B/h^2 \tag{1.13}$$

$$A' e^{-\gamma'(h) n h^2} \leq \sup_{x \in [-1,1]} \|M_x^n - f d\mu\|_{TV} \leq A e^{-\gamma(h) n h^2} \quad \textit{pour } n \geq B/h^2 \tag{1.14}$$

Ici, $\gamma(h), \gamma'(h)$ sont des fonctions de h telles que $\gamma(h) \simeq \gamma'(h) \simeq \alpha\pi^2/8$ quand $h \to 0$, avec α donné par 1.10. Le fait que la vitesse de convergence soit en $n \simeq B/h^2$ est évidemment naturel pour un processus de diffusion.

C. Problème au limite semi-classique

Même si le problème 1-dimensionnel 1.5 est le cas particulier le plus simple de chaîne de metropolis 1.1 dans le cas continu et avec paramètre d'échelle h, il n'entre pas exactement dans la catégorie de ce qu'on savait déjà faire en analyse semi-classique. La raison est qu'il ne s'agit pas d'un problème aux limites standart, mais d'un vrai problème aux limites semi-classique.

Pour la chaîne de Markov 1.5, le rôle de la condition aux limites est tenu par la fonction $m_h(x)$ dont le support est égal à la couche limite $[-1, -1+h] \cup [1-h, 1]$. Ceci montre que l'espace de trace de notre problème 1-dimensionnel n'est pas de dimension finie, mais est égal à l'espace de Hilbert $L^2(-1, -1+h) \oplus L^2(1-h, 1)$. Ceci introduit évidemment des complications techniques notables. Il s'agit d'un authentique "problème aux limites semi-classique pseudodifférentiel ", de sorte que la frontière n'est plus une sous-variété, mais une vraie couche limite de taille h. Remarquer aussi que la fonction $m_h(x)$ n'est pas C^∞, mais seulement lipschitz en $x = -1 + h$ et $x = 1 - h$. On doit donc travailler avec des coefficients à régularité limitée. Le fait que la probabilité de rester dans la couche limite pendant n itérations de la marche aléatoire décroît exponentiellement en n a une contre-partie analytique : l'usage d'opérateurs pseudodifférentiels d'ordre variable dans l'étude des espaces de "scattering" W_λ^c (voir 2.14).

La première chose à faire est bien sur d'etudier la variété caractéristique. Comme dans tout problème aux limites qui se respecte, on doit étudier les solutions ζ complexes de l'équation caractéristique $\hat{\varphi}(\zeta) = \lambda$. Cette étude doit être menée avec soin, car $\hat{\varphi}(\zeta) = \lambda$ n'est pas une équation algébrique, et il est fondamental de comprendre l'asymptotique de ses zéros dans le plan complexe, pour pouvoir in-fine se ramener à la technique usuelle de réduction à un "hamiltonien effectif".

Dans notre travail, les projecteurs de Calderon de la théorie usuelle des problèmes aux limites, sont remplacés par les applications S_λ^c (voir 2.25). La notation S est choisie en référence à " Scattering operator" : S_λ^c relie le comportement au bord aux propriétés globales.

Finalement, la condition de quantification qui sélectionne les valeurs propres est heureusement facile pour ce problème 1-dimensionel : on utilise juste la méthode de Bohr-Sommerfeld.

2 Eléments de preuve

A. Les espaces W_λ^c

On prolonge φ par 0 hors de $[-1, 1]$ et on pose $\hat{\varphi}(\zeta) = \int_{-1}^{1} e^{-iz\zeta}\varphi(z)dz$. Pour tout

$\xi \in \mathbb{R}$, on a $|\hat{\varphi}(\xi)| \leq 1$, avec égalité ssi $\xi = 0$, $\lim_{|\xi|\to\infty} \varphi(\xi) = 0$, et il existe $\nu_0 \in]0, 1[$ tel que

$$\hat{\varphi}(\mathbb{R}) = [-\nu_0, 1] \tag{2.1}$$

Pour $\lambda \in \mathbb{C} \setminus 0$, soit

$$Z_\lambda = \{\zeta \in \mathbb{C};\ \hat{\varphi}(\zeta) = \lambda\} \tag{2.2}$$

Pour tout $\zeta \in Z_\lambda$, on a $-\zeta \in Z_\lambda$ et $\overline{\zeta} \in Z_{\overline{\lambda}}$ car φ est une fonction paire et réelle. Le lemme suivant donne en particulier la description du comportement asymptotique des points de Z_λ pour $|\zeta| \to \infty$.

Lemme 2.1 *Soit $0 < \rho_0 < \rho_1$. Il existe $M > 0$ tel que pour tout $\zeta = a + ib$ on a*

$$\begin{aligned} |a| \geq Me^{|b|} &\Rightarrow |\hat{\varphi}(\zeta)| \leq \rho_0/2 \\ |a| \leq \frac{1}{M}e^{|b|} - 1 &\Rightarrow |\hat{\varphi}(\zeta)| \geq 2\rho_1 \end{aligned} \tag{2.3}$$

Pour tout $\lambda_0 \in \mathbb{C}$ dans l'anneau $\rho_0 \leq |\lambda_0| \leq \rho_1$, il existe $R > 0, l_0 > 0, \tau_0 > 0, r_0 > 0$, avec $2l_0\pi + \pi/2 > \tau_0$ et un symbole $\Theta(\tau, \lambda)$ défini pour $\tau \geq \tau_0$, holomorphe dans $|\lambda - \lambda_0| \leq r_0$, vérifiant les estimations (on fixe une détermination de $log(\lambda)$ dans $|\lambda - \lambda_0| \leq r_0$)

$$\begin{aligned} &\Theta(\tau, \lambda) = \frac{i}{\tau}\Big(log(\frac{\tau}{\widetilde{\varphi}(1)}) + log\lambda + \frac{\widetilde{\varphi}'(1)}{\widetilde{\varphi}(1)}\Big) + \Theta_1(\tau, \lambda) \\ &\forall k \quad \exists C_k \quad |\partial_\tau^k \Theta_1(\tau, \lambda)| \leq C_k \tau^{-k}(\frac{log\tau}{\tau})^2 \end{aligned} \tag{2.4}$$

tel que pour tout $\zeta = a + ib$ avec $|\zeta| \geq R, a > 0, b > 0$, on a $\zeta \in Z_\lambda$ ssi il existe un entier $l \geq l_0$ tel que $\zeta = \zeta_l$ avec

$$\zeta_l = \tau_l + ilog(\lambda) + ilog(\frac{\tau_l}{\widetilde{\varphi}(1)}) + i\Theta(\tau_l, \lambda), \quad \tau_l = 2l\pi + \pi/2 \tag{2.5}$$

Pour tout $c_0 > 0$, il existe $C > 0$ tel que pour ζ vérifiant $|\zeta|$ grand et $|\zeta - \zeta_l| \geq c_0$ pour tout l, on a

$$|\hat{\varphi}(\zeta) - \lambda| \geq C\frac{e^{Im(\zeta)}}{|\zeta|} \tag{2.6}$$

Pour $\zeta \in \mathbb{C}$, on pose $< \zeta >= \sqrt{1 + |\zeta|^2}$. Une fonction $\zeta \in \mathbb{C} \to d(\zeta) \in \mathbb{C}$ est modérée s'il existe $A, B > 0$ tels que

$$|d(\zeta)| \leq A < \zeta >^B \tag{2.7}$$

Pour tout $\zeta \in \mathbb{C}$, comme φ est à support compact, on a

$$\varphi * e^{iz\zeta} = \hat{\varphi}(\zeta)e^{iz\zeta} \tag{2.8}$$

Il est donc naturel d'étudier les sommes d'exponentielles, pour $\lambda \in \mathbb{C} \setminus 0$ donné,

$$f(z) = \sum_{\zeta \in Z_\lambda} d(\zeta)e^{iz\zeta} \tag{2.9}$$

Le lemme 2.1 implique que pour toute fonction modérée d, la formule 2.9 définit une distribution $f \in \mathcal{D}'(\mathbb{R})$. On utilise la convention suivante : si $\zeta \in Z_\lambda$ est un zéro d'ordre m

de l'équation $\hat{\varphi}(\zeta) = \lambda$, $d(\zeta)e^{iz\zeta}$ désigne toute fonction de la forme $\sum_0^{m-1} d_j z^j e^{iz\zeta}$. D'après le lemme 2.1, les zéros multiples n'interviennent que pour ζ dans un compact de $\mathbb{C}$ qui dépend de λ, mais reste dans un compact fixe pour $0 < \rho_0 \leq |\lambda| \leq \rho_1 < \infty$.

On note $\mathcal{D}'_\lambda$ l'espace vectoriel des distributions sur $\mathbb{R}$ de la forme 2.9. D'après 2.8, tout $f \in \mathcal{D}'_\lambda$ vérifie l'équation de convolution

$$\forall f \in \mathcal{D}'_\lambda \quad \varphi * f - \lambda f = 0 \tag{2.10}$$

car f est définie par une série convergente dans $\mathcal{D}'$. Pour $c \in \mathbb{R}$, on note Z^c_λ

$$Z^c_\lambda = \{\zeta \in Z_\lambda, \quad Im(\zeta) > c\} \tag{2.11}$$

On note $\mathcal{D}'_{\lambda,c}$ le sous espace de $\mathcal{D}'_\lambda$

$$f(z) = \sum_{\zeta \in Z^c_\lambda} d(\zeta)e^{iz\zeta}, \qquad d \text{ modérée} \tag{2.12}$$

Soit L^2_c l'espace de Hilbert

$$L^2_c = \{f \in L^2_{\text{loc}}(]0,\infty[), \quad \int_0^\infty e^{2cz}|f(z)|^2 dz < \infty\} \tag{2.13}$$

On note W^c_λ l'espace de distributions sur $]0,\infty[$

$$W^c_\lambda = \{f \in \mathcal{D}'(]0,\infty[), \exists g \in \mathcal{D}'_{\lambda,c} \text{ tel que } g|_{]0,\infty[} = f \quad \text{et} \quad f|_{]0,2[} \in L^2\} \tag{2.14}$$

La proposition suivante est cruciale dans notre analyse. Elle précise la structure de la fibration $W^c_\lambda \to \lambda$ de $\mathbb{C} \setminus 0$ par les espaces de dimension infinie W^c_λ, et dit pourquoi les espaces W^c_λ sont naturels dans notre problème.

Proposition 2.2 *Pour $\lambda \in \mathbb{C} \setminus 0$ et $c \in \mathbb{R}$, W^c_λ est un sous espace fermé de L^2_c , et*

$$\bigcap_{c \in \mathbb{R}} W^c_\lambda = \{0\} \tag{2.15}$$

De plus, pour tout c fixé et tout $\lambda_0 \in \mathbb{C} \setminus 0$, il existe $r > 0$ tel que $\lambda \to W^c_\lambda \subset L^2_c$ est analytique pour $|\lambda - \lambda_0| < r$. Pour tout $\lambda \in \mathbb{C} \setminus [-\nu_0, 1]$, et tout $g \in L^2(\mathbb{R})$ à support dans $[-1,1]$, l'unique solution $f \in L^2(\mathbb{R})$ de l'équation de convolution $\varphi * f - \lambda f = g$ vérifie

$$f|_{]0,\infty[} \in W^0_\lambda \tag{2.16}$$

On utilise aussi le lemme élémentaire suivant

Lemme 2.3 *Pour $\lambda \in \mathbb{C} \setminus 0$ et $c \in \mathbb{R}$, l'opérateur*

$$\begin{aligned} W^c_\lambda &\longrightarrow L^2_c \\ f &\longrightarrow (\varphi * 1_+ f)|_{]0,\infty[} \end{aligned} \tag{2.17}$$

est compact.

B. Le noyau K sur la demi-droite $[0,\infty[$.

Soit m la fonction définie pour $z \geq 0$ par

$$m(z) = \int_0^\infty \varphi(z+t)dt \tag{2.18}$$

Alors m est lipschitzienne, positive décroissante et à support dans $[0,1]$. On a $m(0) = 1/2$. Pour $f \in L^1_{\text{loc}}([0,\infty[)$, on pose pour $z \geq 0$

$$K(f)(z) = m(z)f(z) + (\varphi * 1_+ f)(z) = m(z)f(z) + \int_0^\infty \varphi(z-t)f(t)dt \tag{2.19}$$

Alors K opère sur $L^1_{\text{loc}}([0,\infty[)$, et est la localisation de l'opérateur de metropolis sur la demi-droite $[0,\infty[$. Comme $(\varphi * 1_+ e^{iz\zeta})(z) = (\varphi * e^{iz\zeta})(z) - \int_0^\infty \varphi(z+t)e^{-it\zeta}dt$, pour tout $\zeta \in \mathbb{C}$, on a la formule

$$K(1_+ e^{iz\zeta})(z) = \hat{\varphi}(\zeta)e^{iz\zeta} + \int_0^\infty \varphi(z+t)(e^{iz\zeta} - e^{-it\zeta})dt \tag{2.20}$$

On pose

$$F(z,\zeta) = \int_0^\infty \varphi(z+t)(e^{iz\zeta} - e^{-it\zeta})dt \tag{2.21}$$

Alors $F(z,\zeta)$ est holomorphe en $\zeta \in \mathbb{C}$ à valeurs dans les fonctions continues de $z \geq 0$ à support dans $[0,1]$, et on a la formule de calcul symbolique

$$K(1_+ e^{iz\zeta})(z) = \hat{\varphi}(\zeta)e^{iz\zeta} + F(z,\zeta) \tag{2.22}$$

On a $K(1) = 1$, $K(f) \geq 0$ pour $f \geq 0$, et puisque $\varphi(z) = \varphi(-z)$, pour toutes fonctions mesurables positives f, g

$$\int_0^\infty K(f)g dz = \int_0^\infty f K(g) dz \tag{2.23}$$

Il en résulte

$$\begin{aligned} &\|K(f)\|_{L^\infty} \leq \|f\|_{L^\infty} \quad \forall f \in L^\infty([0,\infty[) \\ &\|K(f)\|_{L^1} \leq \|f\|_{L^1} \quad \forall f \in L^1([0,\infty[) \end{aligned} \tag{2.24}$$

Donc K est autoadjoint sur $L^2([0,\infty[)$, et $\|K\|_{L^2} \leq 1$.

On note S^c_λ l'opérateur (c'est l'analogue d'un opérateur de trace)

$$\begin{aligned} W^c_\lambda &\longrightarrow L^2([0,1]) \\ f &\longrightarrow S^c_\lambda(f) = (K-\lambda)f \end{aligned} \tag{2.25}$$

On a le théorème fondamental suivant :

Théorème 2.4 *Pour tout $\lambda \in \mathbb{C} \setminus [0,1/2]$ et tout $c \in \mathbb{R}$, S^c_λ est Fredholm.*

Soit $c_1 > 0$ tel qu'on ait $Im(\zeta) \geq 2c_1$ pour tout $\zeta \in Z^0_1$. Soit $U_\delta = \{\lambda, |\lambda - 1| < \delta\}$. Pour $\delta > 0$ petit, et $\lambda \in U_\delta$, l'équation $\hat{\varphi}(\zeta) = \lambda$ possède exactement deux solutions $\pm\eta(\lambda)$ telles que $|Im(\eta(\lambda))| \leq c_1$, $\lambda \to \eta(\lambda)^2$ est holomorphe sur U_δ et $\eta(1) = 0$. De plus

$$\begin{aligned} &W^{-c_1}_\lambda = W^{c_1}_\lambda \oplus V^0_\lambda \\ &V^0_\lambda = \text{vect}(\cos(\eta(\lambda)z), \frac{\sin(\eta(\lambda)z)}{\eta(\lambda)}\ ) \end{aligned} \tag{2.26}$$

Le point clé de la preuve est le théorème suivant. Il permet de pouvoir utiliser dans la suite une réduction à un "hamiltonien effectif" .

Théorème 2.5 *Il existe $\delta > 0$ tel que pour tout $\lambda \in U_\delta$, l'opérateur $S_\lambda^{-c_1}$ est surjectif sur $L^2([0,1])$, et son noyau $Ker(S_\lambda^{-c_1}) \subset W_\lambda^{c_1} \oplus V_\lambda^0$ est de dimension 1 et engendré par la fonction*

$$e_\lambda(z) = \cos(\eta(\lambda)z) + b(\lambda)\frac{\sin(\eta(\lambda)z)}{\eta(\lambda)} + g_\lambda(z) \tag{2.27}$$

où $b(\lambda)$ est une fonction holomorphe sur U_δ, $b(\lambda) \in \mathbb{R}$ pour $\lambda \in \mathbb{R}$, $b(1) = 0$, et où g_λ est une fonction holomorphe sur U_δ à valeurs dans $W_\lambda^{c_1}$ et $g_1 = 0$. De plus, il existe $C > 0$ tel que

$$|g_\lambda(z)| \leq C|\lambda - 1|e^{-c_1 z} \quad \forall z \geq 0,\ \forall \lambda \in U_\delta \tag{2.28}$$

On déduit alors du théorème 2.5, le théorème suivant qui dit que si f est approximativement fonction propre de K, associée à la valeur propre λ, alors f est proportionnelle à e_λ modulo une erreur exponentiellement petite.

Théorème 2.6 *Soit $\alpha \in]0,1[$. Il existe $\delta > 0, C_1 > 0$ tel que pour tout $M > 1$ grand, tout $\lambda \in U_\delta$ et tout $f \in L^2([0,\infty[)$ à support dans $[0, M+1]$ tel que $(K-\lambda)f = g$ vérifie $g(z) = 0$ pour tout $z \in [0, M[$, on a*

$$\begin{aligned} f(z) &= ae_\lambda(z) + r(z) \\ \|r\|_{L^2([0,\alpha M])} &\leq e^{-C_1 M}\|f\|_{L^2([0,M+1])} \end{aligned} \tag{2.29}$$

C. Le noyau K_h sur le segment $[-1, 1]$.

Soit $L^2 = L^2([-1,1], dx/2)$. L'opérateur K_h donné par 1.5 est autoadjoint sur L^2, de norme $\|K_h\|_{L^2} = 1$, et puisque l'opérateur $f \mapsto \varphi_h * f$ est compact sur L^2, et comme $m_h([0,\infty[) = [0, 1/2]$, le spectre essentiel de K_h est l'intervalle $[0, 1/2]$. On note $Spec(K_h)$ le spectre de K_h . Pour $\eta \in \mathbb{C}$ près de 0, soit

$$2i\Gamma(\eta) = \log\Big(\frac{1 + \frac{b(\hat{\varphi}(\eta))}{i\eta}}{1 - \frac{b(\hat{\varphi}(\eta))}{i\eta}}\Big) \tag{2.30}$$

Comme $b(1) = 0$, $\frac{b(\hat{\varphi}(\eta))}{i\eta}$ est une fonction holomorphe de η près de $\eta = 0$, nulle en $\eta = 0$. Donc Γ est holomorphe près de 0, $\Gamma(-\eta) = -\Gamma(\eta)$, et $\Gamma(\eta) \in \mathbb{R}$ pour $\eta \in \mathbb{R}$. On note $B(\mu, h)$ la fonction holomorphe définie près de $(0,0)$ par

$$\eta = B(\mu, h) \quad \text{ssi} \quad \eta + h\Gamma(\eta) = \mu \tag{2.31}$$

Le théorème de convergence 1.1 est conséquence du théorème spectral suivant.

Théorème 2.7 *Il existe $h_0 \in]0,1]$ tel que pour tout $h \in]0, h_0]$ on a :*

Il existe $\delta_0 < 1$ indépendant de h, tel que $Spec(K_h) \subset [-\delta_0, 1]$, et $Spec(K_h)$est discret dans $[-\delta_0, 0[$ et $]1/2, 1]$. Il existe $\delta_1 \in]0, 1/2[$ indépendant de h, tel que les valeurs propres de K_h dans l'intervalle $[1-\delta_1, 1]$ sont toutes simples. Pour $0 < \delta \leq \delta_1$, soit $k_h(\delta)$ le *nombre de valeurs propres de K_h dans l' intervalle $[1-\delta, 1[$, et pour $0 \leq k \leq k_h(\delta_1)$, soit $\{\lambda_k(h)\}$ les valeurs propres de K_h dans l'intervalle $[1-\delta_1, 1]$*

$$Spec(K_h) \cap [1-\delta, 1] = \{\lambda_{k_h(\delta)}(h) < ... < \lambda_{k+1}(h) < \lambda_k(h) < ... < \lambda_1(h) < \lambda_0(h) = 1\} \tag{2.32}$$

Alors pour $\delta_1 > 0$ petit, $k_h(\delta)$ vérifie la loi de Weyl

$$k_h(\delta) \simeq \frac{1}{2\pi} vol([-1,1]) \int_{\hat{\varphi}(h\xi) \geq 1-\delta} d\xi = \frac{1}{\pi h} \int_{\hat{\varphi}(\zeta) \geq 1-\delta} d\zeta \tag{2.33}$$

Il existe $c_1 > 0, c_2 > 0$ indépendant de h tels que pour tout $0 \leq k \leq k_h(\delta_1)$, on a

$$|\lambda_k(h) - \hat{\varphi}(B(hk\pi/2, h))| \leq c_1 \exp(-c_2/h) \tag{2.34}$$

En particulier, puisque $\hat{\varphi}(\zeta) = 1 - \alpha\zeta^2/2 + ...$ avec $\alpha = \int_{-1}^{1} z^2 \varphi(z) dz > 0$, et puisque, d'après 2.31 on a $B(h\pi/2, h) = \frac{h\pi}{2} + \mathcal{O}(h^2)$, le trou spectral $1 - \lambda_1(h)$ vérifie

$$1 - \lambda_1(h) = \frac{\alpha h^2 \pi^2}{8} + \mathcal{O}(h^3) \tag{2.35}$$

De plus , le sous-espace propre de K_h associée à la valeur propre $\lambda_k(h)$ est engendré par la fonction propre $\varepsilon_{k,h}$ qui vérifie avec $\eta_{k,h} = B(hk\pi/2, h)$, et pour tout $x \in [-1, 1]$

$$\begin{aligned} &\varepsilon_{k,h}(x) = \cos(\eta_{k,h} \frac{x+1}{h}) + \frac{b(\hat{\varphi}(\eta_{k,h}))}{\eta_{k,h}} \sin(\eta_{k,h} \frac{x+1}{h}) + r_{k,h}(x) \\ &|r_{k,h}(x)| \leq c_1 |\lambda_k(h) - 1| \exp(-\frac{c_2}{h} dist(x, \{[-1, -1+h] \cup [1-h, 1]\})) \end{aligned} \tag{2.36}$$

Bibliographie

[AT87] M. Allen and D Tildesly. *Computer Simulation of liquids.* Oxford University Press, 1987. 4

[BD01] L. Billera and P. Diaconis. A geometric interpretation of the metropolis algorithm. *Stat. Sci.*, 20 :1–5, 2001. 2

[BH02] K. Binder and J. Heermann. *Monte Carlo Simulation in Statistical Physics 4th ed.* Springer, Berlin, 2002. 2

[DN04] P. Diaconis and J.W. Neuberger. Numerical results for the metropolis algorithm. *Experimental Mathematics*, 13 :207–213, 2004. 1

[DSC98] P. Diaconis and L. Saloff-Coste. What do we know about the metropolis algorithm. *Jour. Comp. and Syst. Sci.*, 57 :20–36, 1998. 4

[Has70] W. Hastings. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57 :97–109, 1970. 2

[HH64] J. Hammersley and D. Handscomb. *Monte Carlo Methods.* Wiley, London, 1964. 2

[JH01] B. Jones and J. Hobert. Honest exploration of intractable probability distributions via markov chain monte carlo. *Stat. Sci.*, 16 :317–334, 2001. 4

[Kie00] J. Kienetz. Convergence of markov chains via analytic and isoperimetric inequalities. *Ph.D diss. Univ. Bielefeld*, 2000. 4

[Liu01] J. Liu. *Monte Carlo Strategies in Scientific Computing.* Springer, New York, 2001. 2

[MR00] L. Miclo and C. Roberto. Trous spectraux pour certains algorithmes de metropolis sur $\mathbb{R}$. Lectures notes in math. 1729, pages 336–352. Springer, 2000. 4

[MRR+53] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller. Equations of state calculations by fast computing machines. *J. Chem. Phys.*, 21 :1087–1092, 1953. 2