Spectral problems in the Metropolis Algorithm

Gilles Lebeau
in collaboration with Persi Diaconis and Laurent Michel

Département de Mathématiques,
Université de Nice Sophia-Antipolis
Parc Valrose 06108 Nice Cedex 02, France

february 2016

The analysis of the rate of convergence of a Metropolis algorithm involves the study of the spectral theory of the associated Markov operator. I will illustrate what are these spectral properties on a simple variant of the historical Metropolis chain related to hard spheres.

I will explain how these spectral properties are related to classical estimates in PDE's such as Weyl estimates, Sobolev inequalities, Fourier analysis...

I will describe some of the numerous new challenges and unsolved problems.

And I will try to explain why I believe that the "Metropolis Laplacian" is a natural extension of the usual Laplacian in the framework of a metric and measured space.

Some publications

-P. Diaconis and G. Lebeau *Microlocal analysis for the Metropolis algorithm* , Mathematische Zeitschrift: Volume 262, Issue 2, (2009).

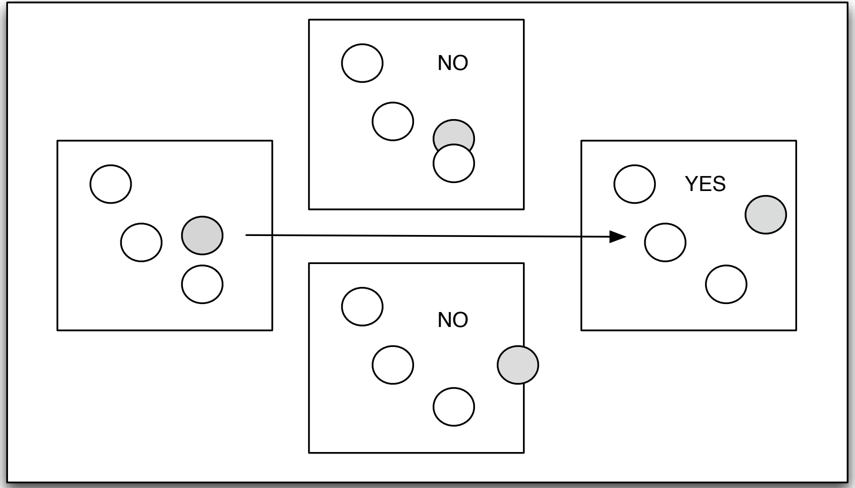
- G. Lebeau and L. Michel *Semi-classical analysis of a random walk on a manifold*, Annals of Probability , Vol. 38, No. 1, (2010).

-P. Diaconis, G. Lebeau and L. Michel *Geometric Analysis of the Metropolis Algorithm in Lipschitz Domains*, Inventiones 185 (2), (2011).

- P. Diaconis, G. Lebeau and L. Michel *Metropolis algorithm on convex polytops*, Mathematische Zeitschrift, 272, (1), (2012).

- G. Lebeau and L. Michel *Hypoelliptic random walks*, Journal de l'Institut Mathématique de Jussieu, (2014).

Hard discs in a square



Outline

Outline

Metropolis Algorithm

Let $(X, \mathcal{B}, \mu)$ be a σ-finite measure space. Let $\rho(x) > 0$, $\int \rho(x) d\mu(x) = 1$, be a probability density on X. The Metropolis algorithm gives a way of drawing samples from ρ. It requires a symmetric proposal density $p(x, y) = p(y, x) \geq 0$ for all x, y, and $\int p(x, y)\mu(dy) = 1$ for all x. This kernel allows us to define the Metropolis kernel

$$
\begin{aligned}
T(x, dy) &= m(x)\delta_x + p(x, y) min(\frac{\rho(y)}{\rho(x)}, 1)\mu(dy) \\
m(x) &= \int_{\{z:\rho(z)<\rho(x)\}} (1 - \frac{\rho(z)}{\rho(x)}) p(x, z)\mu(dz)
\end{aligned}
\tag{2.1}
$$

Metropolis Algorithm

$$
\begin{aligned}
T(x,dy) &= m(x)\delta_x + p(x,y)min(\frac{\rho(y)}{\rho(x)},1)\mu(dy) \\
m(x) &= \int_{\{z:\rho(z)<\rho(x)\}} (1-\frac{\rho(z)}{\rho(x)})p(x,z)\mu(dz)
\end{aligned}
\tag{2.2}
$$

Formula 2.2 has a simple algorithmic interpretation. From x, choose y from the density $p(x,y)$. If $\rho(y) \geq \rho(x)$, move to y. If $\rho(y) < \rho(x)$ flip a coin with probability of heads $\rho(y)/\rho(x)$. If it comes up heads, move to y. If it comes up tails, stay at x. Observe that implementing this does not require knowledge of the normalizing constant for ρ. This is a crucial feature in applications where $\rho(x)$ is given as $Z^{-1}e^{-\beta H(x)}$ with Z unknownable in practice.

Metropolis Algorithm

Let $T^n(x, dy)$ be the kernel of the iterate operator T^n. Then $\int_A T^n(x, dy)$ is the probability to be in the set A after n steps of the walk. Under mild conditions on $p(x, y)$, one has

$$\|T^n(x, dy) - \rho(y)d\mu(y)\|_{TV} \quad \text{as } n \to \infty \quad \forall x$$

The problem is to get estimates on this rate of convergence, and on the spectral theory of the operator T acting as a self adjoint contraction on $L^2(\rho d\mu)$.

Outline

N-body configuration space

We suppose that Ω is a bounded, Lipschitz, quasi-regular, connected open subset of $\mathbb{R}^d$ with $d \geq 2$. Let $N \in \mathbb{N}, N \geq 2$ and $\epsilon > 0$ be given. Let $\mathcal{O}_{N,\epsilon}$ be the open bounded subset of $\mathbb{R}^{Nd}$,

$$\mathcal{O}_{N,\epsilon} = \left\{ x = (x_1, \ldots, x_N) \in \Omega^N, \forall\, 1 \leq i < j \leq N, |x_i - x_j| > \epsilon \right\}. \quad (3.1)$$

Let $\varphi(z)dz$ be the uniform probability on the unit ball of $\mathbb{R}^d$, and let $K_h(x, dy)$ be the Markov kernel

$$K_h(x, dy) = \frac{1}{N} \sum_{j=1}^{N} \delta_{x_1} \otimes \cdots \otimes \delta_{x_{j-1}} \otimes h^{-d} \varphi\left(\frac{x_j - y_j}{h}\right) dy_j \otimes \delta_{x_{j+1}} \otimes \cdots \otimes \delta_{x_N}, \quad (3.2)$$

and the associated Metropolis operator on $L^2(\mathcal{O}_{N,\epsilon})$

$$T_h(u)(x) = m_h(x)u(x) + \int_{O_{N,\epsilon}} u(y) K_h(x, dy), \quad (3.3)$$

$$m_h(x) = 1 - \int_{O_{N,\epsilon}} K_h(x, dy). \quad (3.4)$$

The operator T_h is Markov and self -adjoint on $L^2(\mathcal{O}_{N,\epsilon})$. The configuration space $\mathcal{O}_{N,\epsilon}$ is the set of N disjoint closed balls of radius $\epsilon/2$ in $\mathbb{R}^d$, with centers at the $x_j \in \Omega$. The topology of this set, and the geometry of its boundary is generally hard to understand (!), but since $d \geq 2$, $\mathcal{O}_{N,\epsilon}$ is clearly non-void and connected for a given N if ϵ is small enough. The Metropolis kernel T_h is associated to the following algorithm: at each step, we choose uniformly at random a ball, and we move its center uniformly at random in $\mathbb{R}^d$ in a ball of radius h. If the new configuration is in $\mathcal{O}_{N,\epsilon}$, the change is made. Otherwise, the configuration is kept as it started.

In order to study the random walk associated to T_h, we prove that the open set $\mathcal{O}_{N,\epsilon}$ is Lipschitz quasi-regular for $\epsilon > 0$ small enough, and we prove that the kernel of the iterated operator T_h^M (with M large, but independent of h) admits a suitable lower bound, so that we will be able to use some analytic tools on a simple "model".

Quasi-regular boundary

Definition

Let Ω be a Lipschitz open set of $\mathbb{R}^d$. We say that $\partial\Omega$ is quasi-regular if $\partial\Omega = \Gamma_{reg} \cup \Gamma_{sing}, \Gamma_{reg} \cap \Gamma_{sing} = \emptyset$ with Γ_{reg} the union of smooth hypersurfaces, relatively open in $\partial\Omega$, and Γ_{sing} a closed subset of $\mathbb{R}^d$ such that

$$v \in H^{-1/2}(\partial\Omega) \quad and \quad support(v) \subset \Gamma_{sing} \Longrightarrow v = 0. \qquad (3.5)$$

Observe that 3.5 is obviously satisfied if $\partial\Omega$ is smooth, since in that case one can take $\Gamma_{sing} = \emptyset$. More generally, the boundary is quasi-regular if it is 'piece-wise smooth' in the following sense: Ω is Lipschitz, $\partial\Omega = \Gamma_{reg} \cup \Gamma_{sing}$, $\Gamma_{reg} \cap \Gamma_{sing} = \emptyset$, Γ_{reg} is a smooth hypersurface of $\mathbb{R}^d$, relatively open in $\partial\Omega$, and Γ_{sing} a closed subset of $\mathbb{R}^d$ such that $\Gamma_{sing} = \cup_{j\geq 2} S_j$ where the S_j are smooth disjoint submanifolds of $\mathbb{R}^d$ such that $codim_{\mathbb{R}^d} S_j \geq j$ and $\cup_{k\geq j} S_k = \overline{S_j}$. This 'piece-wise smooth' condition (often called "stratified") is easy to visualize.

Proposition

There exists $C > 0$ such that for $N\epsilon < C$, the set $\mathcal{O}_{N,\epsilon}$ is connected, Lipschitz and with quasi-regular boundary.

The condition $N\epsilon^d < c$, which says that the density of the balls is sufficiently small, does not imply that the set $\mathcal{O}_{N,\epsilon}$ has Lipschitz regularity. As an example, if $\Omega =]0,1[^2$ is the unit square in the plane, then $x = (x_1, \dots, x_N),\ x_j = ((j-1)\epsilon, 0),\ j = 1, \dots, N$, with $\epsilon = \frac{1}{N-1}$ is a configuration point in the boundary $\partial\mathcal{O}_{N,\epsilon}$. However, $\partial\mathcal{O}_{N,\epsilon}$ is not Lipschitz at x: otherwise, there would exist $\nu_j = (a_j, b_j)$ such that $(x_1 + t\nu_1, \dots, x_N + t\nu_N) \in \mathcal{O}_{N,\epsilon}$ for $t > 0$ small enough, and this implies $a_1 > 0, a_{j+1} > a_j$ and $a_N < 0$ which is impossible.

For $k \in \mathbb{N}^*$ denote $B^k = B_{\mathbb{R}^k}(0,1)$ the unit Euclidean ball and $\varphi_k(z) = \frac{1}{vol(B^k)} 1_{B_k}(z)$.

Lemma

Let ϵ be small. There exists $h_0 > 0, c_0, c_1 > 0$ and $M \in \mathbb{N}^*$ such that for all $h \in]0, h_0]$, one has

$$T_h^M(x, dy) = \mu_h(x, dy) + c_0 h^{-Nd} \varphi_{Nd}\left(\frac{x-y}{c_1 h}\right) dy, \tag{3.6}$$

where for all $x \in \mathcal{O}_{N,\epsilon}$, $\mu_h(x, dy)$ is a positive Borel measure.

This allows to reduce most of the analysis of the N-body problem in $\mathbb{R}^d$ to a more simple model, of type 1-body problem in $\mathbb{R}^{Nd}$.

Open problems

1. What can be said about the connected components of the configuration space $\mathcal{O}_{N,\epsilon}$ even in the regime of low density $N\epsilon^d$ small ?

2. What is the geometric structure of the boundary $\partial\mathcal{O}_{N,\epsilon}$ of the configuration space for $N\epsilon$ large ?

3. If K is a compact subspace of the configuration space $\mathcal{O}_{N,\epsilon}$, is it true that $T_h^N(x, dy)$ converges exponentially in total variation, with a rate independent of $x \in K$, to the uniform probability on the h-connected component $C(h, K)$ of K in $\mathcal{O}_{N,\epsilon}$ at least on compact subset of $C(h, K)$?

Outline

model

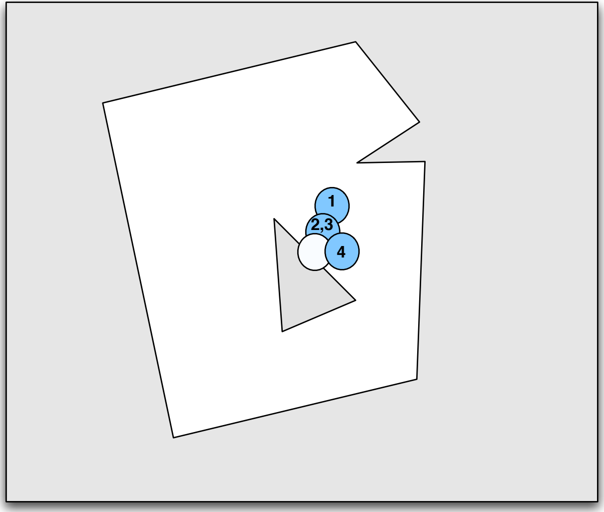
Let Ω be an open, bounded, connected, subset of $\mathbb{R}^d$ with Lipschitz boundary. Let B_1 be the unit ball of $\mathbb{R}^d$ and $\varphi(z) = \frac{1}{Vol(B_1)} 1_{B_1}(z)$ so that $\int \varphi(z)dz = 1$. Let $\rho(x)$ be a measurable bounded positive function on $\overline{\Omega}$ such that $\int_\Omega \rho(x)dx = 1$ and such that $1/\rho$ is bounded. For $h \in]0, h_0]$, set

$$K_{h,\rho}(x,y) = h^{-d}\varphi(\frac{x-y}{h})min(\frac{\rho(y)}{\rho(x)}, 1) \tag{4.1}$$

and let $T_{h,\rho}$ be the (local) Metropolis operator associated to these data, that is

$$\begin{aligned} T_{h,\rho}(u)(x) &= m_{h,\rho}(x)u(x) + \int_\Omega K_{h,\rho}(x,y)u(y)dy \\ m_{h,\rho}(x) &= 1 - \int_\Omega K_{h,\rho}(x,y)dy \geq 0 \end{aligned} \tag{4.2}$$

model



model

Then the Metropolis kernel $T_{h,\rho}(x,dy) = m_{h,\rho}(x)\delta_{x=y} + K_{h,\rho}(x,y)dy$ is a Markov kernel, the operator $T_{h,\rho}$ is self-adjoint on $L^2(\Omega, \rho(x)dx)$, and thus the probability measure $\rho(x)dx$ on Ω is stationary. For $n \geq 1$, we denote by $T^n_{h,\rho}(x,dy)$ the kernel of the iterate operator $(T_{h,\rho})^n$. For any $x \in \Omega$, $T^n_{h,\rho}(x,dy)$ is a probability measure on Ω, and our main goal is to get some estimates on the rate of convergence, when $n \to +\infty$, of the probability $T^n_{h,\rho}(x,dy)$ toward the stationary probability $\rho(y)dy$.

Observe that h is the size of each elementary step of the algorithm, and we want to estimate the rate of convergence uniformly with respect to

$$h \in]0, h_0]$$

.

If $h > 0$ is fixed, estimates are easy: the problem is uniformity when $h \to 0$.

The limit diffusion-operator

Let $\nu(\rho)$ be the best constant such that the following Poincaré inequality holds true for all u in the Sobolev space $H^1(\Omega)$

$$\|u\|^2_{L^2(\rho)} - (u|1)^2_{L^2(\rho)} \leq \frac{1}{\nu(\rho)} \frac{1}{2(d+2)} \int_\Omega |\nabla u|^2(x)\rho(x)dx \tag{4.3}$$

For a smooth density ρ, $\nu(\rho) > 0$ is the first non zero eigenvalue of the unbounded self-adjoint operator L_ρ acting on $L^2(\Omega, \rho(x)dx)$ with domain $D(L_\rho) = \{u \in H^1(\Omega),\ \triangle u \in L^2(\Omega)\ \partial_n u|_{\partial\Omega} = 0\}$

$$L_\rho(u)(x) = \frac{-1}{2(d+2)}(\triangle u + \frac{\nabla\rho}{\rho}.\nabla u) \tag{4.4}$$

$2(d+2)L_\rho$ is the (positive) Laplacian associated to the Dirichlet form $\int_\Omega |\nabla u(x)|^2\rho(x)dx$. It has a compact resolvant and we denote its spectrum by $\nu_0 = 0 < \nu_1 = \nu(\rho) < \nu_2 < ...$ and by m_j the multiplicity of ν_j. Observe that $m_0 = 1$ since $Ker(L)$ is spanned by the constant function equal to 1 .

Spectral theory: easy facts

Since $T_{h,\rho}$ is self-adjoint on $L^2(\Omega, \rho dx)$ and Markov, one has

$$Spect(T_{h,\rho}) \subset [-1, +1]$$

and since $T_h(1) = 1$, 1 is an eigenvalue of T_h. Moreover, since

$$T_{h,\rho}(u)(x) = m_{h,\rho}(x)u(x) + \int_\Omega K_{h,\rho}(x,y)u(y)dy$$

and the operator $u \mapsto \int_\Omega K_{h,\rho}(x,y)u(y)dy$ is compact, the essential spectrum of $T_{h,\rho}$ is **equal** to the closure of the range of the function $x \mapsto m_{h,\rho}(x)$. Since Ω is Lipschitz, there exists $\delta_0 \in]0, 1/2[$ **independent of** h such that $0 \leq m_{h,\rho}(x) \leq 1 - \delta_0$ for all $x \in \Omega$. Therefore

The spectrum of $T_{h,\rho}$ in $[-1, 0[\cup]1 - \delta_0, 1]$ is discrete

Spectral theory: easy facts

Moreover, there exists $\delta_0 \in]0, 1/2[$, **independent of** h $Spec(T_{h,\rho}) \subset [-1 + \delta_0, 1]$ for all $h \in]0, h_0]$. To see this point, write

$$(u+T_{h,\rho}u|u)_{L^2(\rho)} = \frac{1}{2}\int_{\Omega\times\Omega} K_{h,\rho}(x,y)|u(x)+u(y)|^2\rho(x)dxdy+2(m_{h,\rho}u|u)_{L^2(\rho)}$$

Therefore, it is sufficient to prove that there exists $C_0 > 0$ such that the following inequality holds true for all $h \in]0, h_0]$ and all $u \in L^2(\Omega)$

$$\int_{\Omega\times\Omega} h^{-d}\varphi(\frac{x-y}{h})|u(x) + u(y)|^2 dxdy \geq C_0\|u\|^2_{L^2(\Omega)}$$

This is an easy exercise.

Spectral theory: easy facts

Finally, 1 **is a simple eigenvalue of** $T_{h,\rho}$, since one has

$$(u - T_{h,\rho}u|u)_{L^2(\rho)} = \frac{1}{2}\int_{\Omega\times\Omega} K_{h,\rho}(x,y)|u(x)-u(y)|^2\rho(x)dxdy$$

Therefore, $u = T_{h,\rho}u$ implies $|u(x)-u(y)| = 0$ a.e. as soon as $|x-y| < h$, and since Ω is connected, this shows $u = Cte$.
Thus one has the following lemma

Lemma

There exists $\delta_0 \in]0,1/2[$ **independent of** h **such that** $Spect(T_{h,\rho})$ **is a subset of** $[-1+\delta_0, 1]$**. The spectrum of** $T_{h,\rho}$ **is discrete in** $[-1+\delta_0, 0[\cup]1-\delta_0, 1]$**.** 1 **is a simple eigenvalue of** $T_{h,\rho}$ **and** $Ker(T_{h,\rho} - Id)$ **is the one dimensional space spanned by constant functions.**

Spectral gap

We shall denote by $g(h,\rho)$ the **spectral gap** of the Metropolis operator $T_{h,\rho}$. It is defined as the best constant such that the following **Poincaré inequality** holds true for all $u \in L^2(\rho) = L^2(\Omega, \rho(x)dx)$

$$\|u\|^2_{L^2(\rho)} - (u|1)^2_{L^2(\rho)} \leq \frac{1}{g(h,\rho)}(u - T_{h,\rho}u|u)_{L^2(\rho)} \tag{4.5}$$

or equivalently

$$\begin{aligned} &\int_{\Omega\times\Omega} |u(x)-u(y)|^2 \rho(x)\rho(y)dxdy \\ &\leq \frac{1}{g(h,\rho)} \int_{\Omega\times\Omega} K_{h,\rho}(x,y)|u(x)-u(y)|^2 \rho(x)dxdy \end{aligned} \tag{4.6}$$

By the previous lemma, one has always $g(h,\rho) > 0$, but the previous lemma doesn't give a positive lower bound on $g(h,\rho) > 0$.

Spectral gap

Proposition

There exists positive constants C_1, C_2 independent of $h \in]0, h_0]$ such that the spectral gap satisfies

$$C_1 h^2 \leq g(h, \rho) \leq C_2 h^2$$

Observe that the upper bound is obvious, since for any Lipschitz function u, one has

$$(u - T_{h,\rho} u | u)_{L^2(\rho)} \leq Cte \|\nabla u\|^2_{L^\infty} \; h^2$$

I will show how to get the lower bound in the next slides.

It is also obvious that $\lim_{h \to +\infty} g(h, \rho) = 0$, since for h large, the function $m_{h,\rho}(x)$ takes values arbitrary close to 1.

Question: For a given density ρ, find h such that $g(h, \rho)$ is maximal.

h-rough calculus

In all situations, the following theorem is a crucial step.

Theorem

There exists a family $A_{h,L}, h \in]0, h_0]$ of linear operators such that:
1. For any $q \in [1, \infty]$, $A_{h,L}$ is uniformly in h bounded on $L^q(\Omega, \rho dx)$.
2. There exists $C > 0$ such that for any $u \in L^2(\Omega, \rho dx)$ such that $\|u\|^2_{L^2} + h^{-2}(u - T_{h,\rho}u|u)_{L^2} \leq 1$, one has

$$u = u_L + u_H, \quad u_L = A_{h,L}(u)$$

$$\|u_L\|_{H^1} \leq C, \quad \|u_H\|_{L^2} \leq Ch$$

In other words, $u_L = A_{h,L}(u)$ is **the low frequency part of** u, i.e. it retains the behavior of u at scales $\geq h$, and $u_H = (I - A_{h,L})(u)$ is **the high frequency part of** u. Observe that this theorem is **a regularity result**. Observe also that as in the classical h-pseudodifferential calculus, the cut-off on low frequencies $u \mapsto A_{h,L}(u)$ is bounded on all L^q spaces.

Dirichlet forms

For simplicity, I will now assume that the density ρ is given and smooth on $\overline{\Omega}$. I erase the subscript ρ in all the notations.

Let $\mathcal{E}_h, \mathcal{B}_h$ be the Dirichlet forms

$$\mathcal{E}_h(u) = h^{-2}(u - T_h u|u)_{L^2}, \quad \mathcal{B}_h(u,v) = h^{-2}(u - T_h u|v)_{L^2}$$

Let $\mathcal{E}, \mathcal{B}$ be the associated "formal" limit objects when $h \to 0$

$$\mathcal{E}(u) = \frac{1}{2(d+2)} \int_\Omega |\nabla u|^2 \rho dx, \quad \mathcal{B}(u,v) = \frac{1}{2(d+2)} \int_\Omega \nabla u \overline{\nabla v}\ \rho dx$$

Proposition

Let $(f,u) \in H^1 \times H^1$, $(u_h, v_h) \in H^1 \times L^2$. Assume that u_h converges weakly to u in H^1 and v_h bounded in L^2. Then

$$\lim_{h \to 0} \mathcal{B}_h(f, u_h + h v_h) = \mathcal{B}(f,u)$$

The low frequency space

Take $\delta_1 < \delta_0$. We know that the spectrum of T_h in $[1-\delta_1, 1]$ is discrete. Let $e_{j,h}$ be the L^2 normalized orthogonal eigenfunctions of T_h with eigenvalues $1 - h^2\lambda_{j,h}$ such that $h^2\lambda_{j,h} \le \delta_1$.

$$T_h(e_{j,h}) = (1 - h^2\lambda_{j,h})e_{j,h}, \quad \|e_{j,h}\|_{L^2} = 1$$

Definition

The low frequency space $E_{h,L}$ is the finite dimensional space spanned by the eigenfunctions $e_{j,h}$, with $h^2\lambda_{j,h} \le \delta_1$. The family $(e_{j,h})$ is an orthonormal basis of $E_{h,L}$.

Observe that there exists C independent of $e_{j,h}$ such that one has the crude bound

$$\|e_{j,h}\|_{L^\infty} \le Ch^{-d/2}$$

This follows easily from the eigenfunction equation

$$(1 - h^2\lambda_{j,h} - m_h(x))e_{j,h}(x) = \int_\Omega e_{j,h}(y)K_h(x, dy)$$

The Weyl estimate

Observe that the "formal" limit of $h^{-2}(1 - T_h)$ is the Laplacian with Neumann boundary condition

$$L = \frac{-1}{2(d+2)}(\triangle + \frac{\nabla \rho}{\rho}.\nabla)$$

The spectrum of L is discrete, and the "classical" Weyl estimate asserts that the number of eigenvalues μ of L s.t. $\mu \leq \lambda$ is equivalent to $c\lambda^{d/2}$.

Lemma

For any $0 \leq \lambda \leq \delta_1/h^2$, the number of eigenvalues of T_h in $[1 - h^2\lambda, 1]$ (with multiplicity) is bounded by $C(1+\lambda)^{d/2}$. In particular, $dim(E_{h,L}) \leq Ch^{-d}$.

Abstract Weyl estimate

Lemma

Let $s > 0$ *and* $A_h = A_h^* \geq 0$, $h \in]0,1]$ *a family of non negative self-adjoint bounded operators acting on* $L^2(M,\mu)$. *Assume that there exists a constant* $C_0 > 0$ *independent of* h *such that for all* $u \in L^2(M,\mu)$ *such that* $((Id + A_h)u|u) \leq 1$, *the following holds true:*

$$\exists (v,w) \in H^s \times L^2 \text{ such that } u = v + w,\ \|v\|_{H^s} \leq C_0,\ \|w\|_{L^2} \leq C_0 h. \quad (4.7)$$

Let $C_1 < \frac{1}{4C_0^2}$. *There exists* $C_2 > 0$ *independent of* h *such that* $Spec(A_h) \cap [0, \lambda - 1]$ *is discrete for all* $\lambda \leq C_1 h^{-2}$ *and*

$$\#(Spec(A_h) \cap [0, \lambda - 1]) \leq C_2 < \lambda >^{dim(M)/2s}, \quad \forall \lambda \leq C_1 h^{-2}. \quad (4.8)$$

Here, $\#(Spec(A_h) \cap [0,r])$ *is the number of eigenvalues of* A_h *in the interval* $[0,r]$ *with multiplicities, and* $< \lambda >= \sqrt{1 + \lambda^2}$.

The Sobolev estimate

Lemma

There exist $p > 2$ and C independent of $h \in]0, h_0]$ such that for all $u \in E_{h,L}$, the following inequality holds true

$$\|u\|^2_{L^p(M)} \leq C(\mathcal{E}_h(u) + \|u\|^2_{L^2}). \tag{4.9}$$

Proof Let $u \in E_{h,L}$ such that $\mathcal{E}_h(u) + \|u\|^2_{L^2} \leq 1$. One has $u = v_h + w_h$ with $\|v_h\|_{H^1} \leq C$ and $\|w_h\|_{L^2} \leq Ch$. From the imbedding $H^1 \subset L^q(M)$ for $1 > d(1/2 - 1/q)$, we get $\|v_h\|_{L^q} \leq C$. One has $u = \sum_{\lambda_{j,h} \leq \delta_1 h^{-2}} z_{j,h} e_{j,h}$ with $\sum_{\lambda_{j,h} \leq \delta_1 h^{-2}} |z_{j,h}|^2 \leq 1$. By Cauchy-Schwarz, the Weyl estimate and $\|e_{j,h}\|_{L^\infty} \leq Ch^{-d/2}$ we get

$$\|u\|_{L^\infty} \leq Ch^{-d/2}(\sum_{\lambda_{j,h} \leq \delta_1 h^{-2}} |z_{j,h}|^2)^{1/2}(dim(E_{h,L}))^{1/2} \leq Ch^{-d}.$$

From the rough calculus theorem, one has $\|v_h\|_{L^\infty} \leq C\|u\|_{L^\infty} \leq Ch^{-d}$. Thus we get $\|w_h\|_{L^\infty} = \|u - v_h\|_{L^\infty} \leq Ch^{-d}$. Since $\|w_h\|_{L^2} \leq Ch$ we conclude by interpolation.

The spectrum of $T_{h,\rho}$ near 1

Recall that $\nu(\rho) > 0$ is the first non zero eigenvalue of

$$L = \frac{-1}{2(d+2)}(\triangle + \frac{\nabla \rho}{\rho}.\nabla), \quad \text{Neumann}$$

The spectrum of L is denoted by $0 = \nu_0 < \nu_1 = \nu(\rho) \leq \nu_2 \leq ...$ and m_j is the multiplicity of the eigenvalue ν_j.

Theorem

One has

$$lim_{h \to 0} h^{-2} g(h, \rho) = \nu(\rho) \tag{4.10}$$

Moreover, for any $R > 0$ and $\varepsilon > 0$, there exists $h_1 > 0$ such that one has

$$Spec(\frac{1 - T_h}{h^2}) \cap]0, R] \subset \cup_{j \geq 1} [\nu_j - \varepsilon, \nu_j + \varepsilon], \quad \forall h \in]0, h_1] \tag{4.11}$$

and the number of eigenvalues of $\frac{1-T_h}{h^2}$ in the interval $[\nu_j - \varepsilon, \nu_j + \varepsilon]$ is equal to m_j.

Taylor expansion

The proof of the above theorem uses an elementary Taylor expansions lemma:

Lemma

Let $\theta \in C^\infty(\overline{\Omega})$ such that support$(\theta) \cap \Gamma_{sing} = \emptyset$ and $\partial_n \theta|_{\partial\Omega} = 0$. Then

$$(Id - T_h)(\theta) = h^2 L(\theta) + r, \quad \|r\|_{L^2} \in 0(h^{5/2})$$

This lemma implies that for any given interval I, and with $I_\varepsilon = \{x, dist(x, I) \leq \varepsilon\}$ the number of eigenvalues of $h^{-2}(Id - T_h)$ in I_ε is at least equal to the number of eigenvalues of L in I if $h \leq h_\varepsilon$. Also, if $h^{-2}(Id - T_h)(u_h) = \nu_h u_h$, with $\nu_h \to \nu$, then u_h is compact in L^2, so we may assume $u_h \to u$ and we get $(u|(L - \nu)\theta) = 0$ for all $\theta \in C^\infty(\overline{\Omega})$ such that support$(\theta) \cap \Gamma_{sing} = \emptyset$ and $\partial_n \theta|_{\partial\Omega} = 0$. With the quasi regularity of the boundary, this implies that ν is an eigenvalue of L, and u an eigenfunction. The equality of the dimensions follows by compacity of the sequence u_h.

Total Variation

The total variation distance between the two probability measure $T^n_{h,\rho}(x, dy)$ and $\rho(y)dy$ is defined by

$$\| T^n_{h,\rho}(x, dy) - \rho(y)dy \|_{TV} = \sup_A | T^n_{h,\rho}(x, A) - \int_A \rho(y)dy| \tag{4.12}$$

where the sup is taken over all Borel set $A \subset \Omega$. Equivalently

$$\| T^n_{h,\rho}(x, dy) - \rho(y)dy \|_{TV} = \frac{1}{2} \sup_{\|g\|_\infty \leq 1} | T^n_{h,\rho}(g)(x) - \int g(y)\rho(y)dy| \tag{4.13}$$

Convergence in total variation distance

In the following theorem, we assume h_0 small enough to have $g(h,\rho) < dist(Spect(T_h), -1)$.

Theorem

There exists a constant C, such that for all $h \in]0, h_0]$ the following estimate holds true for all integer n

$$sup_{x\in\Omega} \| T_h^n(x,dy) - \rho(y)dy\|_{TV} \leq Ce^{-ng(h,\rho)} \tag{4.14}$$

The above estimate is good for "large n". But it is not enough to answer with some reasonable precision the question:

What is the minimal value of n to get

$$sup_{x\in\Omega} \| T_h^n(x,dy) - \rho(y)dy\|_{TV} \leq 0.1$$

Observe that the following weaker L^2-estimate is obvious

$$\| T_h^n - \Pi_0 \|_{L^2} \leq e^{-ng(h,\rho)}$$

Total variation

Let Π_0 be the orthogonal projector in $L^2(\rho)$ on the space of constant functions

$$\Pi_0(u)(x) = 1_\Omega(x) \int_\Omega u(y)\rho(y)dy \tag{4.15}$$

Then

$$2sup_{x\in\Omega}\| T^n_{h,x} - \rho(y)dy\|_{TV} = \| T^n_h - \Pi_0\|_{L^\infty \to L^\infty} \tag{4.16}$$

Thus, we have to prove that there exist C, such that for any n and any $h \in]0, h_0]$, one has

$$\| T^n_h - \Pi_0\|_{L^\infty \to L^\infty} \leq Ce^{-ng_{h,\rho}} \tag{4.17}$$

Total variation

Observe that since $g_{h,\rho} \simeq h^2$, we may assume $n \geq Ch^{-2}$. In order to prove our estimate, we split T_h in 2 pieces, according to the spectral theory. Let $0 < \lambda_{1,h} \leq ... \leq \lambda_{j,h} \leq \lambda_{j+1,h} \leq ... \leq h^{-2}\delta_1$ be such that the eigenvalues of T_h in the interval $[1-\delta_1, 1[$ are the $1 - h^2\lambda_{j,h}$, with associated orthonormals eigenfunctions $e_{j,h}$

$$T_h(e_{j,h}) = (1 - h^2\lambda_{j,h})e_{j,h}, \quad (e_{j,h}|e_{k,h})_{L^2(\rho)} = \delta_{j,k} \tag{4.18}$$

Then we write $T_h - \Pi_0 = T_{h,1} + T_{h,2}$ with

$$\begin{aligned} T_{h,1}(x,y) &= \sum_{\lambda_{1,h} \leq \lambda_{j,h} \leq \delta_1 h^{-2}} (1 - h^2\lambda_{j,h})e_{j,h}(x)e_{j,h}(y) \\ T_{h,2} &= T_h - \Pi_0 - T_{h,1} \end{aligned} \tag{4.19}$$

$T_{h,2}$

One has

$$\|T_{h,2}^n\|_{L^\infty\to L^\infty} = \|T_h^n - T_{h,1}^n - \Pi_0\|_{L^\infty\to L^\infty} \leq 2 + \|T_{h,1}^n\|_{L^\infty\to L^\infty} \leq Ch^{-3d/2}$$

where the last estimate follows from the Weyl estimate and the crude L^∞ bound on the eigenfunctions $e_{j,h} \in E_{h,L}$, $\|e_{j,h}\| \leq Ch^{-d/2}$.
Next we use $T_h = m_h + R_h$ with $\|m_h\|_{L^\infty\to L^\infty} \leq \gamma < 1$ and $\|R_h\|_{L^2\to L^\infty} \leq C_0 h^{-d/2}$ and we iterate by $T_h^p = m_h^p + B_{p,h}$, $B_{p+1,h} = T_h^p R_h + B_{p,h} m_h$. Next, use $\|T_{2,h}^n\|_{L^2\to L^2} \leq (1-\delta_1)^n$ and

$$\begin{aligned} \|T_{2,h}^{p+n}\|_{L^\infty\to L^\infty} &= \|T_h^p T_{2,h}^n\|_{L^\infty\to L^\infty} \\ &\leq \|m_h^p T_{2,h}^n\|_{L^\infty\to L^\infty} + \|B_{p,h} T_{2,h}^n\|_{L^2\to L^\infty} \\ &\leq C\gamma^p h^{-3d/2} + C_0 h^{-d/2}(1+\gamma+\ldots+\gamma^{p-1})(1-\delta_1)^n \end{aligned} \tag{4.20}$$

Thus we get for some $C>0, \mu>0$,

$$\|T_{2,h}^n\|_{L^\infty\to L^\infty} \leq Ce^{-\mu n}, \quad \forall h, \quad \forall n \geq 1/h \tag{4.21}$$

This is neglectible.

Nash inequality

From the Sobolev lemma, using the interpolation inequality $\|u\|_{L^2}^2 \leq \|u\|_{L^p}^{\frac{p}{p-1}} \|u\|_{L^1}^{\frac{p-2}{p-1}}$, we deduce the **Nash inequality**, with $1/D = 2 - 4/p > 0$

$$\|u\|_{L^2}^{2+1/D} \leq C(\mathcal{E}_h(u) + \|u\|_{L^2}^2)\|u\|_{L^1}^{1/D}, \quad \forall u \in E_{h,L} \tag{4.22}$$

For $\lambda_{j,h} \leq \delta_1 h^{-2}$, one has $h^2\lambda_{j,h} \leq \delta_1 < 1$, and thus for any $u \in E_{h,L}$, one gets $h^2\mathcal{E}_h(u) \leq \|u\|_{L^2}^2 - \|T_h u\|_{L^2}^2$, thus we get from 4.22

$$\|u\|_{L^2}^{2+1/D} \leq Ch^{-2}(\|u\|_{L^2}^2 - \|T_h u\|_{L^2}^2 + h^2\|u\|_{L^2}^2)\|u\|_{L^1}^{1/D}, \quad \forall u \in E_{h,L} \tag{4.23}$$

Nash inequality

From the estimate on $T_{2,h}$, there exists C_2 such that for $n \geq h^{-1}$ one has $\| T_{1,h}^n \|_{L^\infty \to L^\infty} \leq C_2$ and thus since $T_{1,h}$ is self adjoint on L^2 $\| T_{1,h}^n \|_{L^1 \to L^1} \leq C_2$. Fix $p \simeq h^{-1}$. Take $g \in E_{h,L}$ such that $\|g\|_{L^1} \leq 1$ and consider the sequence $c_n, n \geq 0$

$$c_n = \| T_{1,h}^{n+p} g \|_{L^2}^2 \tag{4.24}$$

Then, $0 \leq c_{n+1} \leq c_n$ and from 4.23, we get

$$\begin{aligned} c_n^{1+\frac{1}{2D}} &\leq Ch^{-2}(c_n - c_{n+1} + h^2 c_n) \| T_{1,h}^{n+p} g \|_{L^1}^{1/D} \\ &\leq CC_2^{1/D} h^{-2}(c_n - c_{n+1} + h^2 c_n) \end{aligned} \tag{4.25}$$

Thus there exist A which depends only on C, C_2, D, such that for all $0 \leq n \leq h^{-2}$, one has

$$c_n \leq (\frac{Ah^{-2}}{1+n})^{2D} \tag{4.26}$$

Thus there exist C_0 , such that for $N \simeq h^{-2}$, one has $c_N \leq C_0$. This implies

$$\| T_{1,h}^{N+p} g\|_{L^2} \leq C_0 \|g\|_{L^1} \tag{4.27}$$

and thus taking adjoints

$$\| T_{1,h}^{N+p} g\|_{L^\infty} \leq C_0 \|g\|_{L^2} \tag{4.28}$$

and so we get for any n and with $N + p \simeq h^{-2}$

$$\| T_{1,h}^{N+p+n} g\|_{L^\infty} \leq C_0 (1 - h^2 \lambda_{1,h})^n \|g\|_{L^2} \tag{4.29}$$

And thus for $n \geq h^{-2}$, since by definition $h^2 \lambda_{1,h} = g(h, \rho)$

$$\| T_{1,h}^{n}\|_{L^\infty \to L^\infty} \leq C_0 e^{-(n-h^{-2}) h^2 \lambda_{1,h}} = C_0 e^{\lambda_{1,h}} e^{-ng(h,\rho)} \tag{4.30}$$

QED

Elementary Fourier Analysis

As a byproduct of the estimate (4.26), we get for $g \in E_{h,L}$ by duality

$$\| T_{h,1}^{n+h^{-1}} g \|_{L^\infty} \leq (\frac{Ah^{-2}}{1+n})^D \|g\|_{L^2}$$

With $n \simeq h^{-2} < \lambda >^{-1}$ we get the following polynomial bound .

Lemma

There exists C independent of h such that for any eigenfunction $e_{j,h} \in E_{h,L}$, $\|e_{j,h}\|_{L^2} = 1$, associated to the eigenvalue $1 - h^2\lambda_{j,h}$ of T_h the following inequality holds true

$$\|e_{j,h}\|_{L^\infty} \leq C < \lambda_{j,h} >^D . \tag{4.31}$$

Problem: In PDE's, the natural estimate from Sobolev embeddings for the eigenfunctions of L gives $D = d/4$. In fact, we have better L^∞ estimates for the eigenfunctions of an elliptic operators that the one given by Sobolev, namely, the **Sogge estimates**. So, what can be said about the L^q norms , $q > 2$ of the Metropolis eigenvalues?

Elementary Fourier Analysis

Recall that ν_k is the $k$ieme eigenvalue of the limit operator L with Neumann boundary condition. Let $F_k = Ker(L - \nu_k)$. Recall $m_k = dim(F_k)$ is the multiplicity of the eigenvalue ν_k of L. Let us denote by $\mathcal{J}_k$ the set of indices j such that for h small, $\lambda_{j,h}$, the $j$eme eigenvalue of the Metropolis is close to ν_k, and $F_{h,k} = span(e_{j,h}, j \in \mathcal{J}_k)$. By theorem 4.9 and his proof, the set $\mathcal{J}_k$ is independent of $h \in]0, h_k]$ for h_k small, and one has $\sharp(\mathcal{J}_k) = dim(F_{h,k}) = m_k$ for $h \in]0, h_k]$. Let Π_{F_k} and $\Pi_{F_{h,k}}$be the L^2-orthogonal projectors on F_k and $F_{h,k}$.

Lemma

For all $f \in F_k$ one has

$$\lim_{h \to 0} \|f - \Pi_{F_{h,k}}(f)\|_{L^\infty} = 0. \tag{4.32}$$

This lemma is used to prove a weak convergence result of the Metropolis walk to the continuous diffusion associated to the limit operator L with Neumann boundary condition.

Outline

Convergence to the Wiener measure

bla-bla-bla....

Outline

A canonical random walk

Let $(X, dist, \mu)$ be a metric space equipped with a Borel measure μ , and let $h \in]0, \infty[$ be a parameter. We assume that μ satisfies

$$\mu(B(x,h)) \in]0,\infty[, \quad \forall x \in X, \quad \forall h \in]0,\infty[\tag{6.1}$$

The canonical random walk on $(X, dist, \mu)$ is associated to the Markov kernel

$$T_h f(x) = \frac{1}{\mu(B(x,h))} \int_{B(x,h)} f(y) d\mu \tag{6.2}$$

This means that if the walk is at x_n, then x_{n+1} is choosen randomly in the ball $B(x,h)$ equipped with the probability induced by μ.
Unfortunately, T_h is not in general self-adjoint on $L^2(X, d\mu)$.
It is self-adjoint for the measure $d\nu_h = \mu(B(x,h))d\mu$.

The h-Laplace operator on functions

By definition, a 1-form on X at scale h is an antisymmetric function on

$$Diag_h = \{(x,y) \in X \times X, dist(x,y) < h\}.$$

If $f(x)$ is a function on X, its h-differential is the 1-form

$$\delta_h(f)(x,y) = h^{-1}(f(x) - f(y)) \tag{6.3}$$

Let $C(x,y,h) > 0$ be a symmetric function on $Diag_h$. One defines a measure $\mu_h^{(1)}$ on $Diag_h$ by the formula

$$d\mu_h^{(1)}(x,y) = C(x,y,h)d\mu_h(x)d\mu_h(y) \tag{6.4}$$

The normalization factor $C(x,y,h)$ is a parameter in the theory. The (positive) Laplacian is then defined as usually by the formula

$$|\triangle_h| = \delta_h^* \delta_h$$

$$\delta_h^*(g)(x) = \frac{2}{h}\int_{y\in B(x,h)} g(x,y)C(x,y,h)d\mu(y)$$

$$|\triangle_h|(f)(x) = \frac{2}{h^2}\int_{y\in B(x,h)} \Big(f(x)-f(y)\Big)C_0(x,y,h)d\mu(y)$$

If one defines M_h by the formula

$$\frac{Id - M_h}{h^2} = \frac{1}{2}|\triangle_h| \tag{6.5}$$

one has obviously $M_h(1) = 1$, since $|\triangle_h|(1) = 0$, and

$$M_h f(x) = \int_X f(y)M_h(x,dy) \tag{6.6}$$

where for all $x \in X$, $M_h(x,dy)$ is the measure on X

$$M_h(x,dy) = \Big(1 - \int_{B(x,h)} C_0(x,z,h)d\mu(z)\Big)\delta_x + 1_{B(x,h)}C_0(x,y,h)d\mu(y)$$

Metropolis

Now, observe that M_h is Markov iff for all $x \in X$, the measure $M_h(x, dy)$, which has total mass 1, is positive. This means

$$\forall x \in X, \quad F(x,h) = \int_{B(x,h)} C_0(x,z,h) d\mu(z) \leq 1 \qquad (6.7)$$

A natural choice to get the inequalite 6.7 is to choose the symmetric normalization factor

$$C_0(x,y,h) = \min\left(\frac{1}{\mu(B(x,h))}, \frac{1}{\mu(B(y,h))}\right) \qquad (6.8)$$

This means in that case that M_h, a reversible Markov kernel for the measure μ, is exactly the Metropolis operator associated to the Markov kernel T_h !!!

The associated random walk is simple : if the walk is at x_n, choose $z \in B(x_n, h)$ uniform for the probability $1_{B(x,h)} \frac{1}{\mu(B(x,h))} d\mu(z)$. If $\mu(B(z,h)) \geq \mu(B(x_n,h))$, go to z ($x_{n+1} = z$), else, flip a coin with probability $\mu(B(z,h))/\mu(B(x_n,h))$, go to $z = x_{n+1}$ if you win, else, stay at x_n.