Exponential decay of correlation for the Stochastic Process associated to the Entropy Penalized Method

D. A. Gomes (IST - Portugal) and A. O. Lopes (Inst. Mat - UFRGS -Brasil)

February 27, 2009

D. Gomes was partially supported by the Center for Mathematical Analysis, Geometry and Dynamical Systems through FCT Program POCTI/FEDER and also by grant POCI/FEDER/MAT/55745/2004. A. O. Lopes was partially supported by CNPq, PRONEX – Sistemas Dinâmicos, Instituto do Milênio, and beneficiary of CAPES financial support.

Abstract

In this paper we present an upper bound for the decay of correlation for the stationary stochastic process associated with the Entropy Penalized Method. Let $L(x,v) : \mathbb{T}^n \times \mathbb{R}^n \to \mathbb{R}$ be a C^1 Lagrangian of the form

$$L(x,v) \;=\; \frac{1}{2}\,|v|^2 \;-\; U(x) \;+\; \langle P, v\rangle.$$

We point out that we do not assume more differentiability of L according the the dimension of the torus $\mathbb{T}^n$.

For each value of ϵ and h, consider the operator

$$\mathcal{G}[\phi](x) := -\epsilon h \,\ln\left[\int_{\mathbb{R}^N} e^{-\frac{hL(x,v)+\phi(x+hv)}{\epsilon h}} dv\right],$$

as well as the reversed operator

$$\bar{\mathcal{G}}[\phi](x) := -\epsilon h \,\ln\left[\int_{\mathbb{R}^N} e^{-\frac{hL(x+hv,-v)+\phi(x+hv)}{\epsilon h}} dv\right],$$

both acting on continuous functions $\phi : \mathbb{T}^n \to \mathbb{R}$. Denote by $\phi_{\epsilon,h}$ the solution of $\mathcal{G}[\phi_{\epsilon,h}] = \phi_{\epsilon,h} + \lambda_{\epsilon,h}$, and by $\bar{\phi}_{\epsilon,h}$ the solution of $\bar{\mathcal{G}}[\phi_{\epsilon,h}] = \bar{\phi}_{\epsilon,h} + \lambda_{\epsilon,h}$. Let $\theta_{\epsilon,h}(x) = e^{-\frac{\bar{\phi}_{\epsilon,h}(x)+\phi_{\epsilon,h}(x)}{\epsilon h}}$. From [GV], it is known that

$$\mu_{\epsilon,h}(x,v) = \theta_{\epsilon,h}(x)\,\gamma_{\epsilon,h}(x,v) \;=\; \theta_{\epsilon,h}(x)\, e^{-\frac{hL(x,v)+\phi_{\epsilon,h}(x+hv)-\phi_{\epsilon,h}(x)-\lambda_{\epsilon,h}}{\epsilon h}},$$

is a solution to the entropy penalized problem: $\min\{\int_{\mathbb{T}^N\times\mathbb{R}^N} L(x,v)d\mu(x,v) + \epsilon S[\mu]\}$, where the entropy S is given by

$$S[\mu] = \int_{\mathbb{T}^N\times\mathbb{R}^N} \mu(x,v)\ln\frac{\mu(x,v)}{\int_{\mathbb{R}^N}\mu(x,w)dw}dxdv,$$

and the minimization is made over all holonomic probability densities on $\mathbb{T}^n\times\mathbb{R}^n$, that is probabilities that satisfy $\int \varphi(x+v)-\varphi(x)\mu(x,v)dxdv = 0$, for all $\varphi \in C^1(\mathbb{T}^n)$. The density $\gamma_{\epsilon,h}(x,v)$ defines a Markovian transition kernel on $(\mathbb{T}^n)^{\mathbb{N}}$. The invariant initial density in $\mathbb{T}^n$ is $\theta_{\epsilon,h}(x)$. In order to analyze the decay of correlation for this process we show that the operator $\mathcal{L}\,(\varphi)\,(x) \;=\; \int e^{-\frac{hL(x,v)}{\epsilon}}\,\varphi(x + h\,v)\,d\,v$, has a maximal eigenvalue isolated from the rest of the spectrum.

1 Definitions and the set up of the problem

Let $\mathbb{T}^n$ be the n-dimensional torus. In this paper we assume that the Lagrangian, $L(x,v)$: $\mathbb{T}^n \times \mathbb{R}^N \to \mathbb{R}$ has the form

$$L(x,v) \,=\, \frac{1}{2}\,|v|^2 \,-\, U(x) \,+\, \langle P, v\rangle,$$

where $U \in C^1(\mathbb{T}^n)$, and $P \in \mathbb{R}^n$ is constant.

We consider here the discrete time Aubry-Mather problem [Gom] and the Entropy Penalized Mather method which provides a way to obtain approximations by continuous densities of the Aubry-Mather measure. We refer the reader to [Gom] and the last section of [GLM] for some of the main properties of Aubry-Mather measures , subactions, Peierl's barrier, etc...

The Entropy Penalized Mather problem (see [GV] for general properties of this problem) can be used to approximate Mather measures [CI] by means of absolutely continuous densities $\mu_{\epsilon,h}(x)$, when $\epsilon, h \to 0$, both in the continuous case or in the discrete case. In [GLM] it is presented a Large Deviation principle associated to this procedure. We briefly mention some definitions and results.

Consider, for each value of ϵ and h, the operators acting on continuous functions ϕ:

$$\mathcal{G}[\phi](x) := -\epsilon h\,\ln\left[\int_{\mathbb{R}^N} e^{-\frac{hL(x,v)+\phi(x+hv)}{\epsilon h}}dv\right],$$

and

$$\bar{\mathcal{G}}[\phi](x) := -\epsilon h\,\ln\left[\int_{\mathbb{R}^N} e^{-\frac{hL(x-hv,v)+\phi(x-hv)}{\epsilon h}}dv\right].$$

Denote by $\phi_{\epsilon,h}$ the solution of $\mathcal{G}[\phi_{\epsilon,h}] = \phi_{\epsilon,h}+\lambda_{\epsilon,h}$, and by $\bar{\phi}_{\epsilon,h}$ the solution of $\bar{\mathcal{G}}[\phi_{\epsilon,h}] = \bar{\phi}_{\epsilon,h} + \lambda_{\epsilon,h}$. Let

$$\theta_{\epsilon,h}(x) = e^{-\frac{\bar{\phi}_{\epsilon,h}(x)+\phi_{\epsilon,h}(x)}{\epsilon\, h}}$$

By adding a suitable constant to $\phi_{\epsilon,h}$ or $\bar{\phi}_{\epsilon,h}$, we can assume that $\theta_{\epsilon,h}(x)$ is a probability density on $\mathbb{T}^N$. From D. Gomes and E. Valdinoci, it is known that the probability measure on $\mathbb{T}^N \times \mathbb{R}^N$

$$\mu_{\epsilon,h}(x,v) = \theta_{\epsilon,h}(x)\, e^{-\frac{hL(x,v)+\phi_{\epsilon,h}(x+hv)-\phi_{\epsilon,h}(x)-\lambda_{\epsilon,h}}{\epsilon h}},$$

is a solution to the entropy penalized Mather problem:

$$\min_{\mathcal{M}_h}\left\{\int_{\mathbb{T}^N\times\mathbb{R}^N} L(x,v)d\mu(x,v) + \epsilon S[\mu]\right\},$$

where the entropy S is given by

$$S[\mu] = \int_{\mathbb{T}^N\times\mathbb{R}^N} \mu(x,v)\ln\frac{\mu(x,v)}{\int_{\mathbb{R}^N}\mu(x,w)dw}dxdv,$$

and

$$\mathcal{M}_h := \left\{\mu\in\mathcal{M};\int_{\mathbb{T}^N\times\mathbb{R}^N}\varphi(x+hv)-\varphi(x)d\mu = 0, \forall\varphi\in C(\mathbb{T}^N)\right\}. \tag{1}$$

Here $\mathcal{M}$ denotes the set of probability densities on $\mathbb{T}^N \times \mathbb{R}^N$ and we will call $\mu \in \mathcal{M}_h$ a holonomic probability measure.

We will be interested in measures that minimize the functional bellow (under the holonomic constrain)

$$\int_{\mathbb{T}^N \times \mathbb{R}^N} L(x,v) d\mu(x,v) + \epsilon S[\mu]. \qquad (2)$$

Note that, for a probability $\mu(x,v)$ the value

$$-S[\mu] = -\int_{\mathbb{T}^N \times \mathbb{R}^N} \mu(x,v) \ln \frac{\mu(x,v)}{\int_{\mathbb{R}^N} \mu(x,w) dw} dx\, dv$$

is not necessarily positive.

This is the entropy penalized version of the discrete time Aubry-Mather problem, see [Gom], where we look for probability measures $\mu \in \mathcal{M}_h$ that minimize the action

$$\int_{\mathbb{T}^N \times \mathbb{R}^N} L(x,v) d\mu(x,v) \qquad (3)$$

Definition 1: The forward (non-normalized) Perron operator $\mathcal{L}$ is defined

$$x \to \varphi(x) \Rightarrow x \to \mathcal{L}(\varphi)(x) = \int e^{-\frac{L(x,v)}{\epsilon}} \varphi(x + h\, v)\, d\, v,$$

In [GV] it is shown that $\mathcal{L}$ has a unique eigenfunction $e^{-\frac{\phi_{\epsilon,h}}{h\,\epsilon}}$ with eigenvalue $e^{-\frac{\lambda_{\epsilon,h}}{h\,\epsilon}}$

Definition 2: The backward operator $\mathcal{N}$ is given by

$$x \to \varphi(x) \Rightarrow x \to \mathcal{N}(\varphi)(x) = \int e^{-\frac{L(x-hv,v)}{\epsilon}} \varphi(x - h\, v)\, d\, v,$$

In [GV] it is shown that $\mathcal{N}$ has a unique eigenfunction $e^{-\frac{\bar{\phi}_{\epsilon,h}}{h\,\epsilon}}$ with eigenvalue $e^{-\frac{\lambda_{\epsilon,h}}{h\,\epsilon}}$

Definition 3: The operator

$$g(x) \to \mathcal{F}(g)(x) = \int e^{-\frac{hL(x,v)+\phi_{\epsilon,h}(x+hv)-\phi_{\epsilon,h}(x)-\lambda_{\epsilon,h}}{\epsilon h}} g(x + hv)\, dv,$$

is the normalized forward Perron operator.

From [GV] we have that given a continuous function $g : \mathbb{T}^n \to \mathbb{R}$, then $\mathcal{F}^m(g)$ converges to the unique eigenfunction k as $m \to \infty$. We show in this paper that for ϵ and h fixed, the convergence is exponentially fast.

Our notation:

$$\theta = \theta_{\epsilon,h}(x) = e^{-\frac{\bar{\phi}_{\epsilon,h}(x)+\phi_{\epsilon,h}(x)}{\epsilon\, h}},$$

$$\gamma(x,v) = \gamma_{\epsilon,h}(x,v) = e^{-\frac{hL(x,v)+\phi_{\epsilon,h}(x+hv)-\phi_{\epsilon,h}(x)-\lambda_{\epsilon,h}}{\epsilon h}},$$

in such way that $\mu_{\epsilon,h} = \theta_{\epsilon,h}(x)\gamma_{\epsilon,h}(x,v)$.

2 Reversed Markov Process and Adjoint Operator

In this section we define the reversed Markov process and compute the adjoint of $\mathcal{F}$ in $\mathcal{L}^2(\theta)$. We assume $h = 1$ from now on.

We can consider the stationary forward Markovian process X_n according to the initial probability $\theta(x)$ and transition $\gamma(x, v)$. For example

$$P(X_0 \in A_0) = \int_{x\in\mathbb{T}^n\cap A_0} \theta(x)dx,$$

$$P(X_0 \in A_0, X_1 \in A_1) = \int_{x\in\mathbb{T}^n\cap A_0,\ (x+v)\in A_1} \theta(x)\gamma(x,v)\, dx\, dv,$$

and so on. Define the backward transfer operator $\mathcal{F}^*$ acting on continuous functions $f(x)$ by

$$\mathcal{F}^*(f)(x) = \int \frac{\theta(x-v)\,\gamma(x-v,v)}{\theta(x)}\, f(x-v)\, dv.$$

The backward transition kernel is given by

$$Q(x,v) = \frac{\theta(x-v)\,\gamma(x-v,v)}{\theta(x)}.$$

The fact that for any x we have $\int Q(x,v)\ dv = 1$ follows from Theorem 32 in [GV]. We will show in Corollary 1 that θ is an invariant measure for the process with transition kernel Q, more precisely, that

$$\int g d\theta = \int \mathcal{F}^*(g) d\theta,$$

for any $g \in \mathcal{L}^2(d\theta)$.

Theorem 1. $\mathcal{F}^*$ is the adjoint of $\mathcal{F}$ in $\mathcal{L}^2(\theta)$, that is for all $f, g \in \mathcal{L}^2(\theta)$ then

$$\int f(x)\mathcal{F}g(x)\theta(x)dx = \int g(x)\mathcal{F}^* f(x)\theta(x)dx.$$

Proof. Consider $f, g \in \mathcal{L}^2(\theta)$, then

$$\begin{aligned}
\int g(x)\ [\mathcal{F}^*(f)(x)]\,\theta(x)dx &= \int g(x)\ [\int \frac{\theta(x-v)\,\gamma(x-v,v)}{\theta(x)}\, f(x-v)\, dv\,]\ \theta(x)\ dx \\
&= \int g(x)\ [\int \theta(x-v)\,\gamma(x-v,v)\ f(x-v)\, dv\,]\ dx \\
&= \int [\int [g(x)\ \theta(x-v)\,\gamma(x-v,v)\ f(x-v)\,]\, dx\,]\ dv \\
&= \int [\int g(x+v)\ \theta(x)\,\gamma(x,v)\ f(x)\, dx\,]\ dv \\
&= \int f(x)\,[\int \gamma(x,v)\ g(x+v)\, dv\,]\ \theta(x)\, dx \\
&= \int f(x)\ [\int e^{-\frac{L(x,v)+\phi_{\epsilon,1}(x+v)-\phi_{\epsilon,1}(x)-\lambda_{\epsilon,1}}{\epsilon}}\ g(x+v)\, dv\,]\ \theta(x)\, dx \\
&= \int f(x)\,[\mathcal{F}(g)(x)\,]\,\theta(x)\, dx,
\end{aligned}$$

where we use above the change of coordinates $x \to x - v$ and the fact that μ is holonomic. $\square$

Corollary 1. Consider the inner product $\langle \cdot, \cdot \rangle$ in $\mathcal{L}^2(\theta)$. Then $\mathcal{F}$ leaves invariant the orthogonal space to the constant functions: $\{g \,|\, \langle g, 1 \rangle = \int g\, 1\, d\theta = 0\}$. Furthermore

$$\int g d\theta = \int \mathcal{F}^*(g) d\theta.$$

Proof. Note that $\mathcal{F}(1) = 1$, therefore

$$\int g 1 d\theta = \int g \mathcal{F}(1) d\theta = \int \mathcal{F}^*(g) d\theta.$$

Thus if $\int g 1 d\theta = 0$ it follows $\int \mathcal{F}^*(g) d\theta = 0$. $\square$

3 Spectral gap, exponential convergence and decay of correlations

From [GV] it is known that $\mathcal{L}$ has a unique (normalized) eigenfunction $e^{-\frac{\phi_{\epsilon,h}}{h\,\epsilon}}$ corresponding to the largest eigenvalue $e^{-\frac{\lambda_{\epsilon,h}}{h\,\epsilon}}$, in the next theorem we prove the this eigenvalue is separated from the rest of the spectrum.

Theorem 2. The largest eigenvalue of $\mathcal{L}$ is at a positive distance from the rest of the spectrum.

Proof. We will prove the result for the normalized operator

$$g(x) \to \mathcal{F}(g)\,(x) = \int e^{-\frac{hL(x,v) + \phi_{\epsilon,h}(x+hv) - \phi_{\epsilon,h}(x) - \lambda_{\epsilon,h}}{\epsilon h}}\, g(x + \,h\,v)\, dv.$$

Recall from [GV] that the functions $\phi_{\epsilon,h}(x)$ and $\bar{\phi}_{\epsilon,h}(x)$ are differentiable. In this way we consider a new Lagrangian (adding $\phi_{\epsilon,h}(x + hv) - \phi_{\epsilon,h}(x) - \lambda_{\epsilon,h}$) in such way $\mathcal{L} = \mathcal{F}$. We also assume $\epsilon = 1$ and $h = 1$ from now on.

Therefore,

$$g(x) \to \mathcal{F}(g)\,(x) = \int e^{-L(x,v)}\, g(x + v)\, dv,$$

the eigenvalue is 1, and, by the results in [GV], the corresponding eigenspace is one-dimensional and is generated by the constant functions.

Suppose there exist a sequence of $f_p \in \mathcal{L}^2(\theta)$, $p \in \mathbb{N}$. such that

$$\mathcal{F}(f_p) = \lambda_p(f_p),$$

$\langle f_p, 1 \rangle = 0$, $\lambda_p \to 1$ and $||f_p|| = 1$. If the operator is compact, then the theorem follows from the classical argument: through a subsequence $f_p \to f$, and since $\lambda_p \to 1$ we have $\mathcal{F}(f) = f$. Furthermore, since $\langle f_p, 1 \rangle = 0$, it follows $\langle f, 1 \rangle = 0$, which is a contradiction. Therefore we proceed to establish the compactness of the operator $\mathcal{F}$.

To establish compactness, consider $g \in \mathcal{L}^2(\theta)$. We claim that $f = \mathcal{F}(g)$ is in the Sobolev space $\mathcal{H}^1$ (see [E] for definition and properties). Indeed, for a fixed x, we will compute the derivative of f. Integrating by parts we have

$$\begin{aligned}
\frac{d}{dx} f(x) &= \frac{d}{dx}\,(\mathcal{F}(g)\,(x)) \\
&= \int (\,[\frac{d}{dx} g(x+v)]\; e^{-L(x,v)} \;- L(x,v)\,[\,\frac{d}{dx}\, e^{-L(x,v)}\,]\; g(x+v)\,)\,dv \\
&= \int (\,[\,\frac{d}{dv} g(x+v)\,]\; e^{-L(x,v)} \;- L(x,v)\,[\,\frac{d}{dx} e^{-L(x,v)}\,]\; g(x+v)\,)\,dv \\
&= \int (\,[\frac{d}{dv} e^{-L(x,v)}\,]\, g(x+v) \;- L(x,v)\,[\,\frac{d}{dx}\, e^{-L(x,v)}\,]\; g(x+v)\,)\,dv \\
&= \int (\,[\frac{d}{dv} e^{-L(x,v)}\,] \;\; - L(x,v)\,[\,\frac{d}{dx}\, e^{-L(x,v)}\,]\,)\, g(x+v)\; dv.
\end{aligned}$$

From the hypothesis about L, if $g \in \mathcal{L}^2(\theta)$, then indeed $\frac{d}{dx} f$ is also in $\mathcal{L}^2(\theta)$ (with the above derivative).

Note that, for v uniformly in a bounded set

$$\left\| \frac{d}{dx} f \right\|_2 \le \left\| \frac{d}{dx} f \right\|_\infty \le \left\| [\,\frac{d}{dv} e^{-L(x,v)}\,] \;\; - L(x,v)\,[\,\frac{d}{dx}\, e^{-L(x,v)}\,] \right\|_2 \|g\|_2.$$

Therefore, f is in the Sobolev space $\mathcal{H}^1$.

By iterating the procedure described above, we have that

$$g_j = \mathcal{F}^j(g) \in \mathcal{H}^j.$$

It is known that if $j > \frac{n}{2}$, where n is the dimension of the torus $\mathbb{T}^n$, then g_j is continuous Hölder continuous[E]. Thus the operator $\mathcal{F}$ is compact and g_j is differentiable for a much more larger j. From the reasoning described before, $f_p \to f$, and $\mathcal{F}(f) = f$, $\langle f, 1\rangle = 0$ and f is differentiable. It is easy to see that the modulus of concavity of f is bounded (the iteration by $\mathcal{F}$ does not decrease it). We can add a constant to f and by linearity of $\mathcal{F}$ we also get a new fixed point for $\mathcal{F}$ (note that $\mathcal{F}(1) = 1$). Therefore, we can assume $f = e^{-g}$ for some g.

In this way, we obtain a contradiction with the uniqueness in Theorem 26 in[GV]. $\square$

Suppose $\int g(x)\,\theta(x)dx = 0$. For ϵ, h fixed, then it follows from above that $\mathcal{F}^m(g) \to 0$ with exponential velocity (according to the spectral gap).

Consider the backward stationary Markov process Y_n according to the transition $Q(x, v)$ and initial probability θ as above.

Theorem 3. Given $f(x), g(x)$ with $\int f(x)\,\theta(x)dx = \int g(x)\,\theta(x)dx = 0$, it follows

$$\int g(Y_0)\, f(Y_n)\, d\,P \to 0,$$

with exponential velocity.

Proof. Note that

$$\int g(Y_0)\, f(Y_1)\, d\, P = \int g(x)\, (\int Q(x,v)\, f(x-v)\, dv)\, \theta(x)\, dx =$$

$$\int g(x)\ [\mathcal{F}^*(f)\,(x)\,]\,\theta(x)dx\ =\ \int f(x)\,[\,\mathcal{F}(g)\,(x)\,]\,\theta(x)\,dx.$$

In the same way, for any n

$$\int g(Y_0)\, f(Y_n)\, d\, P = \int f(x)\,[\,\mathcal{F}^n(g)\,(x)\,]\,\theta(x)\,dx.$$

The exponential decay of correlation follows from this. □

Theorem 4. Let $f(x), g(x) \in \mathcal{L}^2(\theta)$ be such that $\int f(x)\,\theta(x)dx = \int g(x)\theta(x)dx = 0$. Then

$$\int g(X_0)\, f(X_n)\, d\, P \to 0,$$

with exponential velocity.

Proof. Now, for analyzing the decay of the forward system, X_n, with transition $\gamma(x,v)$, we have to consider the backwark operator $\mathcal{F}^*$, use the fact that its exponential convergent, that is $(\mathcal{F}^*)^n(g) \to 0$, if $\int g(x)\,\theta(x)dx = 0$, and the result follows in the same way. □

References

[CS] Piermarco Cannarsa and Carlo Sinestrari. *Semiconcave functions, Hamilton-Jacobi equations, and optimal control.* Progress in Nonlinear Differential Equations and their Applications, 58. Birkhäuser Boston Inc., Boston, MA, 2004.

[CI] G Contreras and R. Iturriaga. *Global Minimizers of Autonomous Lagrangians.* AMS 2004. To appear.

[E] L. C. Evans. *Partial Differential equations.* AMS 2004. .

[Gom] D. A. Gomes. *Viscosity Solution methods and discrete Aubry-Mather problem.* Discrete Contin. Dyn. Syst., 13(1): 103-116, 2005.

[GLM] D. A. Gomes, A. O. Lopes and J. Mohr. The Mather measure and a Large Deviation Principle for the Entropy Penalized Method, (2007) to appear

[GV] Diogo A. Gomes and Enrico Valdinoci. *Entropy Penalization Methods for Hamilton-Jacobi Equations.* To appear Adv. in Math.