BULLETIN OF THE BRAZILIAN MATHEMATICAL SOCIETY

Bull Braz Math Soc, New Series 40(1), 1-52

Negative entropy, zero temperature and Markov chains on the interval

A.O. Lopes, J. Mohr, R.R. Souza and Ph. Thieullen

Abstract. We consider ergodic optimization for the shift map on the modified Bernoulli space $\sigma\colon [0,1]^{\mathbb{N}} \to [0,1]^{\mathbb{N}}$, where $[0,1]$ is the unit closed interval, and the potential $A\colon [0,1]^{\mathbb{N}} \to \mathbb{R}$ considered depends on the two first coordinates of $[0,1]^{\mathbb{N}}$. We are interested in finding stationary Markov probabilities μ_∞ on $[0,1]^{\mathbb{N}}$ that maximize the value $\int A d\mu$, among all stationary (i.e. σ-invariant) probabilities μ on $[0,1]^{\mathbb{N}}$. This problem correspond in Statistical Mechanics to the zero temperature case for the interaction described by the potential A. The main purpose of this paper is to show, under the hypothesis of uniqueness of the maximizing probability, a Large Deviation Principle for a family of absolutely continuous Markov probabilities μ_β which weakly converges to μ_∞. The probabilities μ_β are obtained via an information we get from a Perron operator and they satisfy a variational principle similar to the pressure in Thermodynamic Formalism. As the potential A depends only on the first two coordinates, instead of the probability μ on $[0,1]^{\mathbb{N}}$, we can consider its projection ν on $[0,1]^2$. We look at the problem in both ways. If μ_∞ is the maximizing probability on $[0,1]^{\mathbb{N}}$, we also have that its projection ν_∞ is maximizing for A. The hypothesis about stationarity on the maximization problem can also be seen as a transhipment problem. Under the hypothesis of A being C^2 and the twist condition, that is,

$$\frac{\partial^2 A}{\partial x \partial y}(x,y) \neq 0, \quad \text{for all } (x,y) \in [0,1]^2,$$

we show the graph property of the maximizing probability ν on $[0,1]^2$. Moreover, the graph is monotonous. An important result we get is: the maximizing probability is unique generically in Mañé's sense. Finally, we exhibit a separating sub-action for A.

Keywords: negative entropy, Markov chain on $[0,1]$, zero temperature, penalized entropy, maximizing probability, graph property.

Mathematical subject classification: 28D05, 60J10, 37C40, 82B05.

Received 24 June 2008.

1 Introduction

We consider ergodic optimization [Jen1, CG, CLT, Mo] for the shift map on the modified Bernoulli space $\sigma\colon [0,1]^{\mathbb{N}} \to [0,1]^{\mathbb{N}}$, where $[0,1]$ is the unit closed interval, and the potential $A\colon [0,1]^{\mathbb{N}} \to \mathbb{R}$ considered depends on the two first coordinates of $[0,1]^{\mathbb{N}}$. We are interested in finding stationary Markov probabilities μ_∞ on $[0,1]^{\mathbb{N}}$ that maximize the value $\int A d\mu$, among all stationary (i.e. σ-invariant) probabilities μ on $[0,1]^{\mathbb{N}}$, and study properties of this maximizing measures.

We denote by $\mathbf{x} = (x_1, x_2, \ldots)$ a point in $[0,1]^{\mathbb{N}}$, and we consider the shift map $\sigma\colon [0,1]^{\mathbb{N}} \to [0,1]^{\mathbb{N}}$ given by $\sigma((x_1, x_2, \ldots)) = (x_2, x_3, \ldots)$. The sigma-algebra we consider in $[0,1]^{\mathbb{N}}$ is the one generated by the cylinders.

By a stationary probability (or stationary measure) we mean a probability that is σ-invariant. By a stationary Markov probability we mean a stationary probability that is obtained from an initial probability θ on $[0,1]$, and a Markovian transition Kernel $dP_x(y) = P(x, dy)$, where θ is invariant for the kernel defined by P. In the next section we will present precise definitions.

We consider a continuous potential $A\colon [0,1]^{\mathbb{N}} \to \mathbb{R}$ which depends only on the two first coordinates of $[0,1]^{\mathbb{N}}$. Therefore, we can define $\tilde{A}\colon [0,1]^2 \to \mathbb{R}$, as $\tilde{A}(x_1, x_2) = A(\mathbf{x})$, where $\mathbf{x}$ is any point in $[0,1]^{\mathbb{N}}$ which has x_1 and x_2 as its two first coordinates. We will drop the symbol ˜ and the context will show if we are considering a potential in $[0,1]^2$ or in $[0,1]^{\mathbb{N}}$.

We are interested in finding stationary Markov probabilities μ_∞ on the Borel sets of $[0,1]^{\mathbb{N}}$ that maximize the value

$$\int A(x_1, x_2)\, d\mu(\mathbf{x}),$$

among all stationary probabilities μ on $[0,1]^{\mathbb{N}}$.

The maximizing probabilities μ_∞, in general, are not positive in all open sets on $[0,1]^{\mathbb{N}}$.

We present an entropy penalized method (see [GV] for the case of Mather measures) designed to approximate a maximizing probability μ_∞ by (absolutely continuous) stationary Markov probabilities μ_β, $\beta > 0$, obtained from $\theta_\beta(x)$ and $P_\beta(x, y)$ which are continuous functions. The functions θ_β and P_β are obtained from the eigenfunctions and the eigenvalue of a pair of Perron operator (we consider the operators $\varphi \to \mathcal{L}_\beta\varphi(\cdot) = \int e^{\beta A(x,\cdot)}\varphi(x)dx$ and $\varphi \to \bar{\mathcal{L}}_\beta\varphi(\cdot) = \int e^{\beta A(\cdot,y)}\varphi(y)dy$ and we use Krein-Ruthman Theorem) in an analogous way as the case described by F. Spitzer in [Sp] for the Bernoulli space $\Omega = \{1, 2, \ldots, d\}^{\mathbb{N}}$ (see also [PP]).

We will show a large deviation principle for the sequence $\{\mu_\beta\}$ which converges to μ_∞ when $\beta \to \infty$. The large deviation principle give us important information on the rate of such convergence [DZ].

When the state space is the closed unit interval [0, 1], therefore, not countable, strange properties can occur: the natural variational problem of pressure deals with a negative entropy, namely, we have to consider the entropy penalized concept. Negative entropies appear in a natural way when we deal with a continuous state space (see [Ju] for mathematical results and also applications to Information Theory). In physical problems they occur when the spins are in a continuum space (see for instance [Lu, Cv, Ni, RRS, W, BBNg]).

Our result is similar to [BLT] which considers the states space $S = \{1, 2, \ldots, d\}$ and [GLM] which consider the entropy penalized method for Mather measures [CI, Fathi].

In a certain extent, the problem we consider here can be analyzed just by considering probabilities ν on $[0, 1] \times [0, 1]$ defined by

$$\nu\big([a_1, a_2] \times [b_1, b_2]\big) = \int_{a_1}^{a_2} \int_{b_1}^{b_2} dP_{x_1}(x_2) d\theta(x_1),$$

instead of probabilities μ on $[0, 1]^{\mathbb{N}}$ defined by corresponding θ and the markovian kernel $P_x(y)$. We say that ν is the projection of μ on $[0, 1] \times [0, 1]$.

From the point of view of Statistical Mechanics we are analyzing a system of neighborhood interactions described by $A(x, y)$ at temperature zero, where the spin x takes values on [0, 1]. This is another point of view for the meaning of the concept of maximizing probability for A. A well known example is when $A(x, y) = x\,y$, and $x, y \in [-1, 1]$ (see [Th] for references), which can be analyzed using the methods described here via change of coordinates. In the so called XY spin model, we have $A(x, y) = \cos(x - y)$, where $x, y \in (0, 2\pi]$ (see [V, Pe] and [Ta] for explicit solutions). When there is magnetic term one could consider, for instance, $A(x, y) = \cos(x - y) + l \cos(x)$, where l is constant [RRS, A]. We show, among other things, that for such model, given a generic f (in the sense of Mañé [Man]), the maximizing probability for A is unique. Our result seems to be related to section III b) in [CG].

Finally, another point of view for our main result: consider the cost A: $[0, 1] \times [0, 1] \to \mathbb{R}$, and the problem of maximizing $\int A(x, y)\, d\nu(x, y)$, among probabilities ν over $[0, 1] \times [0, 1]$ (which can be disintegrated as $d\nu(x, y) = d\theta(x) dP_x(y)$) with the property of having the same marginals in the x and y coordinates. We refer the reader to [Ra] for a broad description of the Monge-Kantorovich mass transport problem and the Kantorovich-Rubinstein mass transhipment problem. We consider here a special case of such problem. In this way we obtain a robust method (the LDP is true) to approximate the probability ν_∞, which is solution of the mass transhipment problem, via the entropy penalized method.

Under the twist hypothesis, that is

$$\frac{\partial^2 A}{\partial x \partial y}(x, y) \neq 0, \quad \text{for all } (x, y) \in [0, 1]^2,$$

we show that the probability ν_∞ on $[0, 1]^2$ is supported in a graph.

The twist condition is essential in Aubry Theory for twist maps [Ban, Go]. It corresponds, in the Mather Theory, to the hypothesis of convexity of the Lagrangian [Mat, CI, Fathi, Man]. It is also considered in discrete time for optimization problems as in [Ba, Mi]. Here, several results can be obtained without it. But, for getting results like the graph property, it is necessary.

In section 1.1 we present some basic definitions and the main results of the paper. In section 2 we present the induced Markov measures on $[0, 1]^2$ and its relation with stationary measures on $[0, 1]^{\mathbb{N}}$. In section 3 we introduce the Perron operator, the entropy penalized concept and we consider the associated variational problem. In section 4, under the hypothesis of A being C^2 and the twist condition, we show the graph property of the maximizing probability. We also show that for the potential A, in the generic sense of Mañé (see [Man, BC, CI, CLT]), the maximizing probability on $[0, 1]^2$ is unique. We get the same results for calibrated sub-actions. In section 5, we present the deviation function I and show the L.D.P. In section 6, we show the monotonicity of the graph and we exhibit a separating sub-action.

All results presented here can be easily extended to Markov Chains with state space $[0, 1]^2$, or, to more general potentials depending on a finite number of coordinates in $[0, 1]^{\mathbb{N}}$, that is, to A of the form $A(x_1, x_2, \ldots, x_n)$, $A\colon [0, 1]^n \to \mathbb{R}$.

We would like to thanks Alexandre Baraviera and Ana Ribeiro-Teixeira for references and interesting conversations on the subject of the paper.

1.1 Main results

Next we will give some definitions in order to state the main results of this work.

$[0, 1]^{\mathbb{N}}$ can be endowed with the product topology, and then $[0, 1]^{\mathbb{N}}$ becomes a compact metrizable topological space. We will define a distance in $[0, 1]^{\mathbb{N}}$ by

$$d(\mathbf{x}, \mathbf{y}) = \sum_{j \geq 1} \frac{|x_j - y_j|}{2^j}.$$

Definition 1.

(a) *the* shift map *in* $[0, 1]^{\mathbb{N}}$ *is defined as* $\sigma((x_1, x_2, \ldots)) = (x_2, x_3, \ldots)$.

(b) *Let* $A_1, A_2, \ldots, A_k$ *be non degenerated intervals of* $[0, 1]$. *We call a* cylinder of size k *the subset of* $\mathbb{R}^k$ *given by* $A_1 \times A_2 \times \cdots \times A_k$*, and we denote it by* $A_1 \ldots A_k$.

(c) *Let* $\mathcal{M}_{[0,1]^{\mathbb{N}}}$ *be the set of probabilities in the Borel sets of* $[0, 1]^{\mathbb{N}}$. *We define the set of* holonomic measures *in* $\mathcal{M}_{[0,1]^{\mathbb{N}}}$ *as*

$$\mathcal{M}_0 := \left\{ \mu \in \mathcal{M}_{[0,1]^{\mathbb{N}}} : \int (f(x_1) - f(x_2))\, d\mu(\mathbf{x}) = 0,\ \ \forall f \in C([0, 1]) \right\}.$$

Remark.

(i) A cylinder can also be viewed as a subset of $[0, 1]^{\mathbb{N}}$: in this case, we have

$$A_1 \ldots A_k = \left\{ \mathbf{x} \in [0, 1]^{\mathbb{N}} : x_i \in A_i,\ \forall\, 1 \le i \le i \right\}.$$

(ii) For the set of holonomic probabilities $\mathcal{M}_0$, we keep the terminology used in [Gom] and [GL]. This set has been also considered in [Man] and [FS].

(iii) $\mathcal{M}_0$ contain all σ-invariant measures. This is a consequence of the fact that invariant measures for a transformation defined in a compact metric space can be characterized by the measures μ such that $\int f d\mu = \int (f \circ \sigma)\, d\mu$ for all continuous functions defined in $[0, 1]^{\mathbb{N}}$ and taking values in $\mathbb{R}$. Note that the set of σ-invariant measures is a proper subset of $\mathcal{M}_0$.

Definition 2. *A function* $P : [0, 1] \times \mathcal{A} \to [0, 1]$ *is called a* transition probability function *on* $[0, 1]$*, where* $\mathcal{A}$ *is the Borel* σ*-algebra on* $[0, 1]$*, if*

(i) *for all* $x \in [0, 1]$*,* $P(x, \cdot)$ *is a probability measure on* $([0, 1], \mathcal{A})$*,*

(ii) *for all* $B \in \mathcal{A}$*,* $P(\cdot, B)$ *is a* $\mathcal{A}$*-measurable function from* $([0, 1], \mathcal{A}) \to [0, 1]$.

Sometimes we will use the notation $P_x(B)$ for $P(x, B)$.

Any probability ν on $[0, 1]^2$ can be disintegrated as $d\nu(x, y) = d\theta(x) dP_x(y)$, and we will denote it by $\nu = \theta P$, where θ is a probability on $([0, 1], \mathcal{A})$ [Dellach], Pg. 78, (70-III).

Definition 3. *A probability measure θ on* $([0,1],\mathcal{A})$ *is called* stationary *for a transition $P(\cdot\,,\cdot)$, if*

$$\theta(B)=\int P(x,B)d\theta(x) \;\text{ for all } B\in\mathcal{A}.$$

Given the initial probability θ and the transition P, as above, one can define a Markov process $\{X_n\}_{n\in\mathbb{N}}$ with state space $S=[0,1]$ (see [AL] section 14.2 for general references on the topic). If θ is stationary for P, then, one can prove that X_n is a stationary stochastic process. The associated probability μ over $[0,1]^{\mathbb{N}}$ is called the Markov stationary probability defined by θ and P.

Definition 4. *A probability measure $\mu\in\mathcal{M}_{[0,1]^{\mathbb{N}}}$ will be called a* stationary Markov measure *if there exist θ and P as in the Definition* 3*, such that μ is given by*

$$\mu(A_1\ldots A_n):=\int_{A_1\ldots A_n} dP_{x_{n-1}}(x_n)\ldots dP_{x_1}(x_2)\,d\theta(x_1)\,, \tag{1}$$

where $A_1\ldots A_n$ is a cylinder of size n.

We consider the following problem: to find measures that maximize, over $\mathcal{M}_0$, the value

$$\int A(x_1,x_2)\,d\mu(\mathbf{x}),$$

which is more general than the problem of maximizing $\int A d\mu$ over the stationary probabilities.

We define

$$m=\max_{\mu\in\mathcal{M}_0}\left\{\int A d\mu\right\}.$$

We will see that this two problems are equivalents, as we will construct a stationary Markov measure μ such that $m=\int A\,d\mu$. This measure will be called a maximizing stationary Markov measure.

Definition 5.

(a) *A continuous function $u\colon[0,1]\to\mathbb{R}$ is called a* calibrated forward-subaction *if, for any y we have*

$$u(y)=\max_x\big[A(x,y)+u(x)-m\big]. \tag{2}$$

(b) *A continuous function* $u\colon [0,1] \to \mathbb{R}$ *is called a* calibrated backward-subaction *if, for any x we have*

$$u(x) = \max_{y} \big[A(x,y) + u(y) - m \big]. \tag{3}$$

Remark. If A depends on all coordinates in $[0,1]^{\mathbb{N}}$, a calibrated forward-subaction (see [BLT], but note that there they call it a strict subaction, see also [GL]) is a continuous function $u\colon [0,1]^{\mathbb{N}} \to \mathbb{R}$ satisfying

$$u(\mathbf{z}) = \max_{\mathbf{x}\colon \sigma(\mathbf{x})=\mathbf{z}} [A(\mathbf{x}) + u(\mathbf{x}) - m].$$

Hence, if A depends only on the two first coordinates of $[0,1]^{\mathbb{N}}$, Definition 5 is a particular case of this definition.

We denote by $C^2([0,1])$ the set of twice continuously differentiable maps from $[0,1]$ to the real line. The main results of this paper can be summarized by the following theorems (although in the text they will be split in several other results):

Theorem 1. *If* A *is* C^2 *and satisfies* $\dfrac{\partial^2 A}{\partial x \partial y} \neq 0$, *then there exists a generic set* $\mathcal{O}$ *in* $C^2([0,1])$ *(in Baire sense) such that:*

(a) *for each* $f \in \mathcal{O}$, *given* $\mu, \tilde{\mu} \in \mathcal{M}_0$ *two maximizing measures for* $A + f$ *(i.e.,* $m = \int (A+f)\, d\mu = \int (A+f)\, d\tilde{\mu}$*), then*

$$\nu = \tilde{\nu},$$

where ν *and* $\tilde{\nu}$ *are the projections of* μ *and* $\tilde{\mu}$ *in the first two coordinates.*

(b) *The calibrated backward-subaction (respectively, calibrated forward-subaction) for* $A + f$ *is unique.*

Theorem 2. *Let* $A\colon [0,1]^{\mathbb{N}} \to \mathbb{R}$ *be a continuous potential that depends only on the first two coordinates of* $[0,1]^{\mathbb{N}}$. *Then*

(a) *There exist a measure* $\mu_\infty \in \mathcal{M}_0$ *such that* $\int A d\mu_\infty = m$, *and a sequence of stationary Markov measures* μ_β, $\beta \in \mathbb{R}$ *such that*

$$\mu_\beta \rightharpoonup \mu_\infty,$$

where μ_β *is defined by* $\theta_\beta\colon [0, 1] \to \mathbb{R}$, $K_\beta\colon [0, 1]^2 \to \mathbb{R}$ *(see equations* (14) *and* (15)*) as*

$$\mu_\beta(A_1 \dots A_n) := \int_{A_1 \dots A_n} K_\beta(x_{n-1}, x_n) \dots K_\beta(x_1, x_2)\theta_\beta(x_1) dx_n \dots dx_1$$

for any cylinder $A_1 \dots A_n$. *Also* μ_∞ *is a stationary Markov measure.*

(b) *If* A *has only one maximizing stationary Markov measure and there exist an unique calibrated forward-subaction* V *for* A, *then the following LDP is true: for each cylinder* $D = A_1 \dots A_k$, *the following limit exists*

$$\lim_{\beta\to\infty} \frac{1}{\beta} \ln \mu_\beta(D) = - \inf_{\mathbf{x}\in D} I(\mathbf{x}) \, .$$

where $I\colon [0, 1]^{\mathbb{N}} \to [0, +\infty]$ *is a function defined by*

$$I(\mathbf{x}) := \sum_{i\geq 1} V(x_{i+1}) - V(x_i) - (A - m)(x_i, x_{i+1}) \, .$$

Remark to Theorem 2(b). We will show, in what follows, that Theorem 1(a) implies that, for $f \in \mathcal{O}$, the maximizing stationary Markov measure for $A + f$ is unique.

2 Induced stationary Markov measures

In this section we consider a special class of two-dimensional measures that is closely related to the stationary measures. We will prove that the two-dimensional measure of this class that maximizes the integral of the observable A can be extended to a Markov stationary measure that solves the problem of maximization of the integral of A among all stationary measures.

We will denote by $\mathcal{M}_{[0,1]^2}$ the set of probabilities measures in the Borel sets of $[0, 1]^2$. $\mathcal{M}_{[0,1]^2}$ can be endowed with the weak-$\star$ topology, where a sequence $\nu_n \to \nu$, iff, $\int f d\nu_n \to \int f d\nu$, for all continuous functions $f\colon [0, 1]^2 \to \mathbb{R}$. We remember that Banach-Alaoglu theorem implies that $\mathcal{M}_{[0,1]^2}$ is a compact topological space.

Definition 6.

(a) *A probability measure* $\nu \in \mathcal{M}_{[0,1]^2}$ *will be called a* induced stationary Markov measure *if its disintegration* $\nu = \theta P$ *is such that the probability measure* θ *on* $([0, 1], \mathcal{A})$ *is stationary for* P.

In this case for each set $(a, b) \times (c, d) \in [0, 1]^2$ *we have*

$$\nu((a, b) \times (c, d)) = \int_{(a,b)} \int_{(c,d)} dP_x(y) d\theta(x)$$

(b) *We will denote by* $\mathbf{M}$ *the set of induced stationary Markov measures.*

Definition 7.

(a) *A probability measure* ν *will be called an* induced absolutely continuous stationary Markov measure*, if* ν *is in* $\mathbf{M}$ *and can be disintegrated as* $\nu = \theta K$*, where* θ *is an absolutely continuous measure given by a continuous density* $\theta(x)dx$*, and also for each* $x \in [0, 1]$ *the measure* $K(x, .)$ *is an absolutely continuous measure given by a continuous density* $K(x, y)dy$.

(b) *We will denote by* $\mathbf{M}_{ac}$ *the set of induced absolutely continuous stationary Markov measures.*

We can see that the above continuous densities $K : [0, 1]^2 \to [0, +\infty)$ and $\theta : [0, 1] \to [0, +\infty)$ satisfy the following equations:

$$\int K(x, y)\, dy = 1, \qquad \forall x \in [0, 1], \tag{4}$$

$$\int \theta(x)\, K(x, y)\, dxdy = 1, \tag{5}$$

$$\int \theta(x)\, K(x, y)\, dx = \theta(y), \qquad \forall y \in [0, 1]. \tag{6}$$

Moreover, any pair of non-negative continuous functions satisfying the three equations above define an induced absolutely continuous stationary Markov measure.

Let $C[0, 1]$ denotes the set of continuous functions defined in $[0, 1]$ and taking values on $\mathbb{R}$, and $C([0, 1]^2)$ denotes the set of continuous functions defined in $[0, 1]^2$ and taking values on $\mathbb{R}$.

Lemma 1.

(a) $\mathbf{M} = \left\{\nu \in \mathcal{M}_{[0,1]^2} : \int f(x) - f(y)\, d\nu(x, y) = 0\ ,\ \forall f \in C[0, 1]\right\}$,

(b) $\mathbf{M}$ *is a closed set in the weak-*$\star$ *topology.*

Proof.

(a) Suppose that $\nu = \theta P \in \mathbf{M}$ is a induced stationary Markov measure. Remembering that $\int f d\nu$ is defined by the limit of integrals of simple functions, it is enough to show that $\int f(x)d\nu(x, y) = \int f(y)d\nu(x, y)$ for $f = \chi_B$ where B is a Borel set. We have

$$\begin{aligned}\int_{[0,1]}\int_{[0,1]} \chi_B(x)dP_x(y)d\theta(x) &= \int_{[0,1]} \chi_B(x)\int_{[0,1]} dP_x(y)d\theta(x)\\ &= \int_{[0,1]} \chi_B(x)d\theta(x) = \theta(B) = \int P(x, B)d\theta(x)\\ &= \int_{[0,1]}\int_B dP_x(y)d\theta(x) = \int_{[0,1]}\int_{[0,1]} \chi_B(y)dP_x(y)d\theta(x)\,.\end{aligned}$$

Now we will suppose that ν is a measure in $\mathcal{M}_{[0,1]^2}$ which satisfies $\int f(x)d\nu(x, y) = \int f(y)d\nu(x, y)$ for any $f \in C[0, 1]$. Let $\nu = \theta P$ be the disintegration of ν. To prove that ν belongs to $\mathbf{M}$, we can use the fact that $\mathcal{A}$ is generated by the intervals, and thus we just have to prove that $\theta(B) = \int P(x, B)d\theta(x)$ if B is an interval.

Therefore, Let B be an interval, and $f_n \in C[0, 1]$ a sequence of $[0, 1]$-valued continuous functions that converges pointwise to χ_B (such a sequence always exists). By the dominated convergence theorem we have that

$$\begin{aligned}\theta(B) &= \int_{[0,1]} \chi_B(x)d\theta(x) = \lim_{n\to+\infty}\int_{[0,1]} f_n(x)d\theta(x)\\ &= \lim_{n\to+\infty}\int_{[0,1]}\int_{[0,1]} f_n(x)dP_x(y)d\theta(x)\\ &= \lim_{n\to+\infty}\int_{[0,1]}\int_{[0,1]} f_n(y)dP_x(y)d\theta(x)\\ &= \lim_{n\to+\infty}\int_{[0,1]} \varphi_n(x)d\theta(x)\,,\end{aligned}$$

where $\varphi_n(x) \equiv \int_{[0,1]} f_n(y)dP_x(y)$. Now, defining

$$\varphi(x) \equiv \int_{[0,1]} \chi_B(y)dP_x(y),$$

we can use again the dominated convergence theorem to get that $\varphi_n(x) \to \varphi(x)$. Hence the function φ_n is pointwise convergent and uniformly bounded. Using the dominated convergence theorem once more, we have that

$$\begin{aligned}\theta(B) &= \lim_{n\to+\infty}\int_{[0,1]}\varphi_n(x)d\theta(x) = \int_{[0,1]}\varphi(x)d\theta(x)\\ &= \int_{[0,1]}\int_{[0,1]}\chi_B(y)\,dP_x(y)d\theta(x) = \int_{[0,1]}\int_B dP_x(y)d\theta(x)\\ &= \int P(x,B)d\theta(x)\,.\end{aligned}$$

(b) Suppose $\nu_n \in \mathbf{M}$, and $\nu_n \to \nu \in \mathcal{M}_{[0,1]^2}$ in the weak-$\star$ topology. We have that $\int f d\nu_n \to \int f d\nu\ \forall f \in C([0,1]^2)$. In particular, if $f \in C[0,1]$, we have $\int f(x)d\nu_n(x,y) \to \int f(x)d\nu(x,y)$ and $\int f(y)d\nu_n(x,y) \to \int f(y)d\nu(x,y)$. Therefore, $\nu \in \mathbf{M}$ because

$$\int f(x) - f(y)d\nu(x,y) = \lim_{n\to\infty}\int f(x) - f(y)d\nu_n(x,y) = 0\,,$$

for all $f \in C[0,1]$. □

The above formulation of the set $\mathbf{M}$ is more convenient for the duality of Fenchel-Rockafellar (see [Roc] and the discussion on section 3) required by Proposition 4. It just says that both marginals in the x and y coordinates are the same.

Sometimes we consider μ over $[0,1]^{\mathbb{N}}$ and sometimes the corresponding projected ν over $[0,1]^2$ (Proposition 1 below deals with projections of measures from $\mathcal{M}_0$ to $\mathbf{M}$). We will forget the word projected from now on, and the context will indicate which one we are working with. Note that, to make the lecture easier, we are using the following notation: ν when we want to refer to a measure in $[0,1]^2$ and μ for the measures in $[0,1]^{\mathbb{N}}$.

Remark. We point out that maximizing $\int A d\nu$ for probabilities on $\nu \in \mathbf{M}$, means a Kantorovich-Rubinstein (mass transhipment) problem where we assume the two marginals are the same (see [Ra], Vol. I, section 4 for a related problem). The methods presented here can be used to get approximations of the optimal probability by absolutely continuous ones. These probabilities are obtained via the eigenfunctions of a Perron operator.

In the case we are analyzing, where the observable depends only on the two first coordinates, we will establish some connections between the measures in $[0,1]^2$ and the measures in $[0,1]^{\mathbb{N}}$, and we will see that the problem of maximization can be analyzed as a problem of maximization among induced Markov measures in $[0, 1]^2$.

Proposition 1. *Let $A\colon [0,1]^{\mathbb{N}} \to \mathbb{R}$ be a potential which depends only in the first two coordinates of $[0,1]^{\mathbb{N}}$. Then the following is true:*

(a) *There exists a map, not necessarily surjective, from $\mathbf{M}$ to $\mathcal{M}_0$.*

(b) *There exists a map, not necessarily injective, from $\mathcal{M}_0$ to $\mathbf{M}$.*

(c) $\max_{\mu \in \mathcal{M}_0} \int A(x_1, x_2)\, d\mu(\mathbf{x}) = \max_{\nu \in \mathbf{M}} \int A(x, y)\, d\nu(x, y)$

Proof.

(a) A measure $\nu \in \mathbf{M}$ can be disintegrated as $\nu = \theta P$, and then can be extended to a measure $\mu \in \mathcal{M}_0$ by

$$\mu(A_1 \dots A_n) := \int_{A_1 \dots A_n} dP_{x_{n-1}}(x_n) \dots dP_{x_1}(x_2)\, d\theta(x_1)\,, \tag{7}$$

Also, we have

$$\int_{[0,1]^{\mathbb{N}}} A(x_1, x_2) d\mu(\mathbf{x}) = \int_{[0,1]^2} A(x, y) d\nu(x, y)\,.$$

(b) A measure $\mu \in \mathcal{M}_0$ can be projected in a measure $\nu \in \mathcal{M}_{[0,1]^2}$, defined by, for each Borel set B of $[0, 1]^2$,

$$\nu(B) = \mu(\Pi^{-1}(B))\,,$$

where $\Pi\colon [0,1]^{\mathbb{N}} \to [0,1]^2$ is the projection in the two first coordinates. Note that, by Lemma 1, $\nu \in \mathbf{M}$. Then we have

$$\int_{[0,1]^{\mathbb{N}}} A(x_1, x_2) d\mu(\mathbf{x}) = \int_{[0,1]^2} A(x, y) d\nu(x, y)\,.$$

(c) It follows easily by (a) and (b). □

Remark. Note that in the item (a), in the particular case where $\nu \in \mathbf{M}_{ac}$, we have that ν can be disintegrated as $\nu = \theta K$, and then the stationary Markov measure μ is given by

$$\mu(A_1 \dots A_n) := \int_{A_1 \dots A_n} K(x_{n-1}, x_n) \dots K(x_1, x_2)\, \theta(x_1)\, dx_n \dots dx_1\,, \tag{8}$$

where $A_1 \dots A_n$ is a cylinder.

3 The maximization problem

We are interested in finding stationary Markov probabilities μ_∞ on $[0, 1]^{\mathbb{N}}$ that maximize the value

$$\int A(x_1, x_2) d\mu(\mathbf{x}),$$

over $\mathcal{M}_0$.

By item (c) of Proposition 1:

$$\max_{\mu \in \mathcal{M}_0} \int A\, d\mu = \max_{\nu \in \mathbf{M}} \int A\, d\nu.$$

Hence, the problem we are analyzing is equivalent to the problem of finding ν_∞ which is maximal for $\int A d\nu$, among all $\nu \in \mathbf{M}$. Because once we have ν_∞, by item (a) of Proposition 1, we obtain a maximizing Markov measure μ_∞ among the holonomic measures.

As we only consider potentials of the form $A(x, y)$, it is not possible to have uniqueness of the maximizing measure on $\mathcal{M}_0$. We just take into account the information of the measure on cylinders of size two. In any case, the stationary Markov probability we will describe below will also solve this maximizing problem.

One of the main results we will get in this section is to be able to approximate singular probabilities by absolutely continuous probabilities (depending on a parameter β) by means of eigenfunctions of a kind of Perron operator.

Now we will concentrate on the maximizing problem in $[0, 1]^2$.

Let $A\colon [0, 1] \times [0, 1] \to \mathbb{R}$ be a continuous function. We will denote by

$$\mathfrak{M}_0 := \left\{ \nu \in \mathbf{M} \colon \int A(x, y)\, d\nu(x, y) = m \right\}$$

where

$$m = \max_{\nu \in \mathbf{M}} \left\{ \int A(x, y) d\nu(x, y) \right\}.$$

A measure in $\mathfrak{M}_0$ will be called a maximizing measure on $\mathbf{M}$.

Consider now the variational problem

$$\max_{\theta K \in \mathbf{M}_{ac}} \left\{ \int \beta A(x, y) \theta(x) K(x, y) dx dy - \int \theta(x) K(x, y) \log(K(x, y)) dx dy \right\} \tag{9}$$

In some sense we are considering above a kind of pressure problem (see [PP]).

Definition 8. *We define the* term of entropy *of an absolutely continuous probability measure* $\nu \in \mathcal{M}_{[0,1]^2}$*, given by a density* $\nu(x, y)dxdy$*, as*

$$S[\nu] = -\int \nu(x, y) \log\left(\frac{\nu(x, y)}{\int \nu(x, z)dz}\right) dxdy\,. \tag{10}$$

We remark that, in the case where A depends on all coordinates in $[0, 1]^{\mathbb{N}}$, the natural entropy (similar to Kolmogorov entropy for the case of the usual shift on the Bernoulli space) to be considered would be infinity. Therefore, it does not make sense to consider the associated concept of pressure (using Kolmogorov entropy) and we believe it is not possible to go further in our reasoning to this more general setting. The bottom line is: we want to approximate singular probabilities by absolutely continuous probabilities (depending on a parameter β) by means of eigenfunctions of a kind of Perron operator. We want to take limits in a parameter β and this is easier to do if we have a variational principle (like the one considered above).

It is easy to see that any $\nu = \theta K \in \mathbf{M}_{ac}$ satisfies

$$S[\theta K] = -\int \theta(x)K(x, y) \log(K(x, y))\, dxdy\,. \tag{11}$$

We call $S[\nu] = S[\theta K]$ the entropy penalized of the probability $\nu = \theta K \in \mathbf{M}_{ac}$.

Lemma 2. *If* $\nu = \theta K \in \mathbf{M}_{ac}$ *and* K *is positive, then* $S[\nu] \leq 0$.

Proof. log is a concave function. Hence, by Jensen inequality, we have

$$\begin{aligned} &-\int \theta(x)K(x, y) \log(K(x, y))\, dxdy \\ &= \int \theta(x)K(x, y) \log\left(\frac{1}{K(x, y)}\right) dxdy \\ &\leq \log \int \theta(x)K(x, y) \frac{1}{K(x, y)} dxdy \\ &= \log(1) = 0. \end{aligned}$$

□

For each β fixed, we will exhibit a measure ν_β in $\mathbf{M}_{ac}$ which maximizes (9). After, we will show that such ν_β will approximate in weak convergence the probabilities ν_∞ which are maximizing for A in the set $\mathbf{M}$.

In order to do that, we need to define the following operators:

Definition 9. *Let $\mathcal{L}_\beta, \bar{\mathcal{L}}_\beta : C([0,1]) \to C([0,1])$ be given by*

$$\mathcal{L}_\beta\varphi(y) = \int e^{\beta A(x,y)}\,\varphi(x)dx, \tag{12}$$

$$\bar{\mathcal{L}}_\beta\varphi(x) = \int e^{\beta A(x,y)}\,\varphi(y)dy. \tag{13}$$

We refer the reader to [Ka] and [Sch] chapter IV for a general reference on positive integral operators.

The above definitions are quite natural and extend the usual Ruelle-Perron operator definition. In the present situation the state space is continuous and an integral should take place of the sum. We are interested in approximating singular measures (which are maximizing for A) by absolutely continuous probabilities, therefore, it is natural to integrate with respect to Lebesgue measure.

Theorem 3. *The operators $\mathcal{L}_\beta$ and $\bar{\mathcal{L}}_\beta$ have the same positive maximal eigenvalue λ_β, which is simple and isolated. The eigenfunctions associated are positive functions.*

Proof. We can see that $\mathcal{L}_\beta$ is a compact operator, because the image by $\mathcal{L}_\beta$ of the unity closed ball of $C([0,1])$ is a equicontinuous family in $C([0,1])$: we know that $e^{\beta A}$ is a uniformly continuous function, and then, if φ is in the closed unit ball, we have

$$\begin{aligned} |\mathcal{L}_\beta\varphi(y) - \mathcal{L}_\beta\varphi(z)| &\le \int |e^{\beta A(x,y)} - e^{\beta A(x,z)}|\,|\varphi(x)|dx \\ &\le |e^{\beta A(x,y)} - e^{\beta A(x,z)}| < \delta\,, \end{aligned}$$

if, y and z are close enough. Thus, we can use Arzela-Ascoli Theorem to prove the compactness of $\mathcal{L}_\beta$ (see also [Sch, Chapter IV, section 1]).

The spectrum of a compact operator is a sequence of eigenvalues that converges to zero, possibly added by zero. This implies that any non-zero eigenvalue of $\mathcal{L}_\beta$ is isolated (i.e. there are no sequence in the spectrum of $\mathcal{L}_\beta$ that converges to some non-zero eigenvalue).

The definition of $\mathcal{L}_\beta$ now shows that it preserves the cone of positive functions in $C([0,1])$, indeed, sending a point in this cone to the interior of the cone. This means that $\mathcal{L}_\beta$ is a positive operator.

The Krein-Ruthman Theorem ([De, Theorem 19.3]) implies that there exists a positive eigenvalue λ_β, which is maximal (i.e. $\lambda_\beta > |\lambda|$, if $\lambda \neq \lambda_\beta$ is in the spectrum of $\mathcal{L}_\beta$) and simple (i.e. the eigenspace associated to λ_β is one-dimensional). Moreover λ_β is associated to a positive eigenfunction φ_β.

If we proceed in the same way, we get the same conclusions about the operator $\bar{\mathcal{L}}_\beta$, and we get the respective eigenvalue $\bar{\lambda}_\beta$ and eigenfunction $\bar{\varphi}_\beta$.

In order to prove that $\bar{\lambda}_\beta = \lambda_\beta$, we use the positivity of φ_β and $\bar{\varphi}_\beta$ and the fact that $\bar{\mathcal{L}}_\beta$ is the adjoint of $\mathcal{L}_\beta$ (here we see that our operators can be, in fact, defined in the Hilbert space $L^2([0, 1])$, which contains $C([0, 1])$). We have

$$\langle \varphi_\beta, \bar{\varphi}_\beta \rangle = \int \varphi_\beta(x)\bar{\varphi}_\beta(x)dx > 0, \quad \text{and}$$

$$\lambda_\beta \langle \varphi_\beta, \bar{\varphi}_\beta \rangle = \langle \mathcal{L}_\beta \varphi_\beta, \bar{\varphi}_\beta \rangle = \langle \varphi_\beta, \bar{\mathcal{L}}_\beta \bar{\varphi}_\beta \rangle = \bar{\lambda}_\beta \langle \varphi_\beta, \bar{\varphi}_\beta \rangle. \qquad \square$$

An estimate on the spectral gap for the operator $\mathcal{L}_\beta$, where $\beta > 0$, is given in [Os, Hop]: suppose

$$\tilde{M} = \sup_{(x,y)\in[0,1]^2} A(x, y), \quad \text{and} \quad \tilde{m} = \inf_{A(x,y)\in[0,1]^2} A(x, y).$$

If λ_β is the main eigenvalue, then, by Theorem 4 of [Hop], any other λ in the spectrum of $\mathcal{L}_\beta$ satisfies

$$\lambda_\beta \left(\frac{\tilde{M}^\beta - \tilde{m}^\beta}{\tilde{M}^\beta + \tilde{m}^\beta} \right) > \lambda.$$

With this information one can give an estimate of the decay of correlation for functions evolving under the probability of the Markov Chain associated to such value β (see next proposition). The proof of this claim is similar to the reasoning in chapter 2 page 26 in [PP], which deals with the case where the state space is discrete.

Let us call $\varphi_\beta, \bar{\varphi}_\beta$ the positive eigenfunctions for $\mathcal{L}_\beta$ and $\bar{\mathcal{L}}_\beta$ associated to λ_β, which satisfy $\int \varphi_\beta(x)\, dx = 1$ and $\int \bar{\varphi}_\beta(x)\, dx = 1$.

We will define a density $\theta_\beta \colon [0, 1] \to \mathbb{R}$ by

$$\theta_\beta(x) := \frac{\varphi_\beta(x)\ \bar{\varphi}_\beta(x)}{\pi_\beta}, \tag{14}$$

where $\pi_\beta = \int \varphi_\beta(x)\bar{\varphi}_\beta(x)dx$, and a transition $K_\beta \colon [0, 1]^2 \to \mathbb{R}$ by

$$K_\beta(x, y) := \frac{e^{\beta A(x,y)}\ \bar{\varphi}_\beta(y)}{\bar{\varphi}_\beta(x)\, \lambda_\beta}. \tag{15}$$

Let $\nu_\beta \in \mathcal{M}_{[0,1]^2}$ be defined by

$$d\nu_\beta(x, y) := \theta_\beta(x) K_\beta(x, y)dxdy. \tag{16}$$

It is easy to see that θ_β, K_β satisfy equations (4), (5) and (6), hence $\nu_\beta \in \mathbf{M}_{ac}$.

Proposition 2. *The Markov measure $\nu_\beta = \theta_\beta K_\beta$ defined above maximize*

$$\int \beta\, A(x,y)\, \theta(x) K(x,y) dx dy - \int \theta(x) K(x,y) \log\left(K(x,y)\right) dx dy$$

over all absolutely continuous Markov measures. Also

$$\log \lambda_\beta = \int \beta A\, \theta_\beta K_\beta dx dy + S[\theta_\beta K_\beta].$$

Proof. By the definition of the functions θ_β, K_β, we have

$$S[\theta_\beta K_\beta] = -\int (\beta A(x,y) + \log \bar{\varphi}_\beta(y) - \log \bar{\varphi}_\beta(x) - \log \lambda_\beta) d\nu_\beta\,.$$

Then

$$\int \beta A\, \theta_\beta K_\beta dx dy + S[\theta_\beta K_\beta]$$

$$= \log \lambda_\beta + \int (\log \bar{\varphi}_\beta(x) - \log \bar{\varphi}_\beta(y)) \theta_\beta(x) K_\beta(x,y) dx dy\,,$$

and the last integral is zero because $\nu_\beta = \theta_\beta K_\beta \in \mathbf{M}_{ac}$.

To show that ν_β is maximizing let ν be any measure in $\mathbf{M}_{ac}$ and $0 \leq \tau \leq 1$. We claim that the function

$$I[\tau] := \int \beta A d\nu_\tau + S[\nu_\tau]$$

where $\nu_\tau = (1-\tau)\nu_\beta + \tau\nu$, is concave and $I'(0) = 0$

Indeed, see proof of Theorem 33 of [GV]. We just point out that the entropy term in [GV] has a difference of sign. □

Lemma 3.

(a) *There exists a constant $c > 0$ such that, for all $x \in [0,1]$, we have*

$$e^{-\beta c} \leq \varphi_\beta(x) \leq e^{\beta c} \quad \textit{and} \quad e^{-\beta c} \leq \bar{\varphi}_\beta(x) \leq e^{\beta c}.$$

Also,

$$\beta \mapsto \frac{1}{\beta} \log \pi_\beta \quad \textit{and} \quad \beta \mapsto \frac{1}{\beta} \log \lambda_\beta$$

are bounded functions, defined for $\beta > 0$.

(b) *The sets*

$$\left\{ \frac{1}{\beta} \log(\varphi_\beta) \mid \beta > 1 \right\} \quad \textit{and} \quad \left\{ \frac{1}{\beta} \log(\bar{\varphi}_\beta) \mid \beta > 1 \right\}$$

are equicontinuous, and relatively compact in the supremum norm.

Proof.

(a) Fix $\beta > 0$. Using the normalization $\int \varphi_\beta(z)dz = 1$, we choose x_0 and x_1 in [0, 1] satisfying $\varphi_\beta(x_0) \leq 1$ and $\varphi_\beta(x_1) \geq 1$. Now, if $\|A\|$ is the supremum norm of A, we have

$$\lambda_\beta = \frac{1}{\varphi_\beta(x_1)} \int e^{\beta A(z,x_1)} \varphi_\beta(z)dz \leq e^{\beta\|A\|}$$

and

$$\lambda_\beta = \frac{1}{\varphi_\beta(x_0)} \int e^{\beta A(z,x_0)} \varphi_\beta(z)dz \geq e^{-\beta\|A\|} .$$

Thus, $-\|A\| < \frac{1}{\beta}\log\lambda_\beta < \|A\|$.

Now we use the inequalities above and the fact that

$$\varphi_\beta(x) = \frac{1}{\lambda_\beta} \int e^{\beta A(z,x)} \varphi_\beta(z)dz$$

to prove that

$$\varphi_\beta(x) \leq \frac{1}{\lambda_\beta} \int e^{\beta\|A\|} \varphi_\beta(z)dz \leq e^{2\beta\|A\|} ,$$

and

$$\varphi_\beta(x) \geq \frac{1}{\lambda_\beta} \int e^{-\beta\|A\|} \varphi_\beta(z)dz \geq e^{-2\beta\|A\|} .$$

We define $c = 2\|A\|$. The eigenfunctions $\overline{\varphi}_\beta$ are bounded by an analogous estimative. Now,

$$\pi_\beta = \int \varphi_\beta(x)\overline{\varphi}_\beta(x)dx,$$

and thus $e^{-2\beta c} \leq \pi_\beta \leq e^{2\beta c}$, which implies that $\frac{1}{\beta}\log\pi_\beta$ is a bounded function of β.

(b) We just have to prove the equicontinuity of both sets. Once we have that, and considering the fact that both sets are sets of functions defined in the compact set [0, 1], we use item (a) and Arzela-Ascoli's Theorem to get the relative compactness of these sets.

To have the equicontinuity for the first set, let y be a point in [0, 1], and let $\beta > 1$. Let $\epsilon > 0$. We will use the fact that A is a uniformly continuous map: We know there exists $\delta > 0$, such that $|y - z| < \delta$, implies $|A(x, y) - A(x, z)| <$

ϵ, $\forall x \in [0,1]$. Without any loss of generality, we suppose that $\varphi_\beta(y) \geq \varphi_\beta(z)$. We have:

$$\begin{aligned}
&\left|\frac{1}{\beta}\log(\varphi_\beta(y)) - \frac{1}{\beta}\log(\varphi_\beta(z))\right| \\
&\quad = \frac{1}{\beta}\left(\log\left(\frac{1}{\lambda_\beta}\int e^{\beta A(x,y)}\varphi_\beta(x)dx\right) - \log\left(\frac{1}{\lambda_\beta}\int e^{\beta A(x,z)}\varphi_\beta(x)dx\right)\right) \\
&\quad = \frac{1}{\beta}\log\left(\frac{\int e^{\beta A(x,y)}\varphi_\beta(x)dx}{\int e^{\beta A(x,z)}\varphi_\beta(x)dx}\right) \leq \frac{1}{\beta}\log\left(\frac{\int e^{\beta(A(x,z)+\epsilon)}\varphi_\beta(x)dx}{\int e^{\beta A(x,z)}\varphi_\beta(x)dx}\right) \\
&\quad = \frac{1}{\beta}\log\left(e^{\beta\epsilon}\frac{\int e^{\beta A(x,z)}\varphi_\beta(x)dx}{\int e^{\beta A(x,z)}\varphi_\beta(x)dx}\right) = \epsilon\,.
\end{aligned}$$

We prove the equicontinuity for the second set in the same way. □

From the above, we can find $\beta_n \to \infty$ which defines convergent subsequences $\dfrac{1}{\beta_n}\log\varphi_{\beta_n}$.

Let us fix a subsequence β_n such that $\beta_n \to \infty$ and all the three following limits exist:

$$\begin{aligned}
V(x) &:= \lim_{n\to\infty}\frac{1}{\beta_n}\log\varphi_{\beta_n}(x), \\
\bar{V}(x) &:= \lim_{n\to\infty}\frac{1}{\beta_n}\log\bar{\varphi}_{\beta_n}(x), \\
\tilde{m} &:= \lim_{n\to\infty}\frac{1}{\beta_n}\log\lambda_{\beta_n}
\end{aligned}$$

Note that the limits defining V and $\bar{V}$ converge uniformly. In principle, the function V depends on the sequence β_n we choose.

Proposition 3 [Laplace's Method]. *Let $f_k\colon [0,1] \to \mathbb{R}$ be a sequence of functions that converge uniformly, as k goes to ∞, to a function $f\colon [0,1] \to \mathbb{R}$. Then*

$$\lim_k \frac{1}{k}\log\int_0^1 e^{kf_k(x)}dx = \sup_{x\in[0,1]} f(x)$$

Lemma 4.

$$\lim_{n\to\infty}\frac{1}{\beta_n}\log\pi_{\beta_n} = \max_{x\in[0,1]}(V(x) + \bar{V}(x))$$

Proof.

$$\pi_{\beta_n} = \int_0^1 \varphi_{\beta_n}(x)\bar{\varphi}_{\beta_n}(x)dx = \int_0^1 e^{\beta_n\left(\frac{1}{\beta}\log\varphi_{\beta_n}(x)+\frac{1}{\beta}\log\bar{\varphi}_{\beta_n}(x)\right)}dx$$

And note that $\frac{1}{\beta_n}\log\varphi_{\beta_n}(x) \to V(x)$, $\frac{1}{\beta_n}\log\bar{\varphi}_{\beta_n}(x) \to \bar{V}(x)$ uniformly. Hence it follows by Laplace's Method. □

Also by Laplace's method we have the following lemma:

Lemma 5.

$$V(y) = \max_{x\in[0,1]}\big(V(x) + A(x,y) - \tilde{m}\big)$$

and

$$\bar{V}(x) = \max_{y\in[0,1]}\big(\bar{V}(y) + A(x,y) - \tilde{m}\big).$$

For some subsequence (of the subsequence $\{\beta_n\}$ fixed after the proof of Lemma 3, which we will also denote by $\{\beta_n\}$), the measures ν_{β_n} defined in (16) weakly converge to a measure $\nu_\infty \in \mathcal{M}_{[0,1]^2}$. Then

$$\lim_{n\to\infty}\int A d\nu_{\beta_n} = \int A d\nu_\infty .$$

Lemma 6. *The measure* $\nu_\infty \in \mathbf{M}$.

Proof. As $\nu_{\beta_n} \in \mathbf{M}_{ac} \subset \mathbf{M}$, by item (b) of Lemma 1 we have that $\nu_\infty \in \mathbf{M}$. □

Theorem 4.

$$\int A(x,y)d\nu_\infty(x,y) = m$$

i.e., ν_∞ *is a maximizing measure on* $\mathbf{M}$.

In order to prove Theorem 4 we need first some new results.

Proposition 4. *Given a potential* $A \in C([0,1]^2)$ *, we have that*

$$\sup_{\nu\in\,\mathbf{M}}\int A d\nu = \inf_{f\in\,C([0,1])}\ \max_{(x,y)}\,(A(x,y) + f(x) - f(y))$$

This proposition will be a consequence of the Fenchel-Rockafellar duality theorem (see [Roc]). Let us fix the setting we consider in order to apply this theorem.

Let $C([0,1]^2)$ be the set of continuous functions in $[0,1]^2$ with the supremum norm and S the set of signed measures over the Borel $\sigma-$algebra of $[0,1]^2$.

Consider the convex correspondence $H\colon C([0,1]^2) \to \mathbb{R}$ given by $H(\phi) = \max(A+\phi)$ and

$$C := \left\{\phi \in C([0,1]^2)\colon \phi(x,y) = f(x) - f(y),\ \text{for some } f \in C([0,1])\right\}$$

We define a concave correspondence $G\colon C([0,1]^2) \to \mathbb{R}\cup\{-\infty\}$ by $G(\phi) = 0$ if $\phi \in C$ and $G(\phi) = -\infty$ otherwise.

Then the corresponding Fenchel transforms, $H^*\colon S \to \mathbb{R}\cup\{+\infty\}$, $G^*\colon S \to \mathbb{R}\cup\{-\infty\}$, are given by

$$H^*(\nu) = \sup_{\phi\in C([0,1]^2)} \left[\int \phi(x,y)d\nu(x,y) - H(\phi)\right]$$

and

$$G^*(\nu) = \inf_{\phi\in C([0,1]^2)} \left[\int \phi(x,y)d\nu(x,y) - G(\phi)\right]$$

We define $S_0 := \{\nu \in S\colon \int f(x) - f(y)d\nu(x,y) = 0\,\forall f \in C[0,1]\}$, and we note that $S_0 \cap \mathcal{M}_{[0,1]^2} = \mathbf{M}$.

Lemma 7. *Given H and G as above, then*

$$H^*(\nu) = \begin{cases} -\int A(x,y)d\nu(x,y) & \text{if } \nu \in \mathcal{M}_{[0,1]^2} \\ +\infty & \text{otherwise} \end{cases}$$

$$G^*(\nu) = \begin{cases} 0 & \text{if } \nu \in S_0 \\ -\infty & \text{otherwise} \end{cases}$$

This lemma follows from Lemma 2 of [GL].

Proof of Proposition 4. The duality theorem of Fenchel-Rockafellar says that

$$\sup_{\phi\in C([0,1]^2)} \left[G(\phi) - H(\phi)\right] = \inf_{\nu\in S}\left[H^*(\nu) - G^*(\nu)\right].$$

Hence, by Lemma 7 and the uniform convergence we have

$$\sup_{\phi\in C}\left[-\max_{(x,y)}(A+\phi)(x,y)\right] = \inf_{\nu\in\mathbf{M}}\left[-\int A(x,y)d\nu(x,y)\right].$$

Using the definition of C we have that

$$\sup_{\nu\in\mathbf{M}}\int A d\nu = \inf_{f\in C([0,1])}\ \max_{(x,y)}\left(A(x,y) + f(x) - f(y)\right). \qquad \square$$

Lemma 8. $\tilde{m} = m$.

Proof. Note that by Proposition 4 and Lemma 5 we have that $m \leq \tilde{m}$. To show the other inequality remember that

$$\log \lambda_{\beta_n} = \int \beta_n A \, d\nu_{\beta_n} + S[\nu_{\beta_n}].$$

Then

$$\tilde{m} = \lim_{n\to\infty} \int A \, d\nu_{\beta_n} + \frac{1}{\beta_n} S[\nu_{\beta_n}].$$

Note that $\nu_{\beta_n} \in \mathbf{M}$, which implies $\int A \, d\nu_{\beta_n} \leq m$.

As $S[\nu_{\beta_n}] \leq 0$, we have

$$\int A \, d\nu_{\beta_n} + \frac{1}{\beta_n} S[\nu_{\beta_n}] \leq m \,\forall n$$

Then $\tilde{m} \leq m$. □

Proof of Theorem 4. Remember that $\nu_{\beta_n} \rightharpoonup \nu_\infty$, then

$$\lim_{n\to\infty} \int A d\nu_{\beta_n} = \int A d\nu_\infty.$$

By Lemma 8 and the fact that $S[\nu_{\beta_n}] \leq 0$, we obtain

$$m = \lim_{n\to\infty} \int A \, d\nu_{\beta_n} + \frac{1}{\beta_n} S[\nu_{\beta_n}] \leq \int A \, d\nu_\infty.$$

Hence, using Lemma 6, we have that $m = \int A d\nu_\infty$. □

4 Uniqueness of maximizing measures and calibrated subactions

We want to remark here that for the results of this section we were inspired by the works of [Gom, GLM] and [GL]. Hence, jointing all these ideas, and what was proved before, we are able to show that there is a unique maximizing probability for A in $\mathbf{M}$, if A is generic in Mañé's sense, the potential A is C^2 and satisfy the twist property. Similar result is true for the calibrated subaction. The precise definitions will be given in what follows.

The differentiable structure of $[0, 1]$ will help us to get the uniqueness required when we want to show the graph property for the support of the maximizing probability.

We repeat here the important definition (Definition 5) of forward (backward)-calibrated subactions:

Definition 10. *A continuous function* $u\colon [0,1] \to \mathbb{R}$ *is called a*

(a) calibrated forward-subaction *if, for any* y *we have*

$$u(y) = \max_x \big[A(x,y) + u(x) - m\big]. \tag{17}$$

(b) calibrated backward-subaction *if, for any* x *we have*

$$u(x) = \max_y \big[A(x,y) + u(y) - m\big]. \tag{18}$$

Note also that if we add a constant to a calibrated forward-subaction, this will be a new calibrated forward-subaction. When we say here that under some conditions, the calibrated forward-subaction is unique, we say this up to an additive constant.

Note that V and $\bar{V}$ defined in Lemma 5 are, respectively, forward and backward calibrated subactions (remember that $\tilde{m} = m$ by Lemma 8).

Subactions (see also [CLT]) play the role in discrete time dynamics of fixed points of the Lax-Oleinik operators of Mather Theory [Fathi].

Let u be a calibrated backward-subaction, using the fact that $[0,1]$ is compact, there exists $y(x)$ (maybe not unique) such that

$$u(x) = A(x, y(x)) + u(y(x)) - m. \tag{19}$$

Proposition 5. *Let* $\nu \in \mathfrak{M}_0$ *be any maximizing measure, and u be a calibrated backward-subaction. Then for all* $(x,y) \in \operatorname{supp}(\nu)$ *we have*

$$u(x) = A(x,y) + u(y) - m.$$

Proof. Note that

$$u(x) \geq A(x,y) + u(y) - m \quad \text{for all } y \in [0,1].$$

As $\nu \in \mathfrak{M}_0$, we have $\int A d\nu = m$ and $\int (u(x) - u(y))\, d\nu = 0$. This proves that the equality in the statement of the theorem is true ν-almost everywhere, in the other points of the support of ν this holds by continuity. □

We point out that a calibrated-subaction (backward or forward) does not need to be differentiable. We want to show that, in certain points of $[0,1]$, a calibrated-subaction is differentiable. In order to do that we introduce the following generalized differentials.

Definition 11. *Let $u\colon B \to \mathbb{R}$ and $x \in B$, where B is an open set in* $[0, 1]$. *The sets*

$$D^+u(x) = \left\{ p \in \mathbb{R} \mid \limsup_{|v|\to 0} \frac{u(x+v) - u(x) - p\,v}{|v|} \leq 0 \right\},$$

$$D^-u(x) = \left\{ p \in \mathbb{R} \mid \liminf_{|v|\to 0} \frac{u(x+v) - u(x) - p\,v}{|v|} \geq 0 \right\},$$

are called, respectively, the superdifferential *and the* subdifferential of u at x.

The main point here is that the differentiable structure of $[0, 1]$ will help us to get the uniqueness required by what we will call later the graph property.

Proposition 6. *Let $u\colon B \to \mathbb{R}$ and $x \in B$, where B is an open set in* $[0, 1]$. *$D^-u(x)$ and $D^+u(x)$ are both nonempty sets if and only if u is differentiable at x. In this case $D^-u(x) = D^+u(x) = Du(x)$.*

Proof. See [CS, Proposition 3.1.5]. □

Lemma 9. *Let u be a calibrated backward-subaction. We have the following statements:*

(a) *$\forall x \in [0, 1]$, $D^-u(x) \neq \emptyset$, and $\dfrac{\partial A}{\partial x}(x, y(x)) \in D^-u(x)$, where $y(x)$ is such that $(x, y(x))$ satisfies equation* (19)*;*

For $(x, y(x))$ satisfying equation (19)*:*

(b) *$D^+u(y(x)) \neq \emptyset$ and $-\dfrac{\partial A}{\partial y}(x, y(x)) \in D^+u(y(x))$;*

(c) *u is differentiable at $y(x)$.*

Proof.

(a) Let $x \in [0, 1]$, then there exists $y(x)$ such that $(x, y(x))$ satisfies equation (19). For any $w \in [0, 1]$, using equation (18), we have that

$$u(x+w) \geq A(x+w, y(x)) + u(y(x)) - m\,.$$

This inequality and the equality in equation (19), give that

$$u(x+w) - u(x) - A(x+w, y(x)) + A(x, y(x)) \geq 0,$$

and then

$$\liminf_{|w|\to 0} \frac{u(x+w)-u(x)-(\frac{\partial A}{\partial x}(x,y(x))\; w+o(w))}{|w|} \geq 0\,,$$

and this implies $\frac{\partial A}{\partial x}(x,y(x)) \in D^{-}u(x)$.

(b) Also for $(x, y(x))$ satisfying (19) and any $w \in [0,1]$, using equation (18), we have

$$u(x) \geq A(x, y(x)+w) + u(y(x)+w) - m.$$

Using equation (19), we get that

$$u(y(x)+w) - u(y(x)) + A(x, y(x)+w) - A(x, y(x)) \leq 0.$$

Then

$$\limsup_{|w|\to 0} \frac{u(y(x)+w)-u(y(x))-\big(-\frac{\partial A}{\partial y}(x,y(x))\; w+o(w)\big)}{|w|} \leq 0.$$

Hence, $-\frac{\partial A}{\partial y}(x,y(x)) \in D^{+}u(y(x))$.

(c) It is just a consequence of items (a) and (b) and of Proposition 6. □

Lemma 10. *For any measure* $\nu \in \mathbf{M}$*, we have that, for almost every point* $(x,y) \in \mathrm{supp}(\nu)$*, there exists* z *such that* $(z,x) \in \mathrm{supp}(\nu)$.

Proof. Define the set

$$R = \Big\{(x,y) \in \mathrm{supp}(\nu)\colon\; \mathrm{supp}(\nu) \cap ([0,1] \times \{x\}) = \emptyset\Big\}$$

Suppose, by contradiction, that $\nu(R) = \epsilon > 0$.

Let $\pi_j : [0,1]^2 \to [0,1]$ be the projection on the j-th coordinate.

Let ν_2 be the measure on the Borel sets of $[0,1]$ given by $\nu_2(B) = \nu(\pi_2^{-1}(B))$, where B is any Borel set in $[0,1]$.

Consider $R_1 = \pi_1(R)$. We have

$$R_1 = \Big\{x \in \pi_1(\mathrm{supp}(\nu))\colon\; \mathrm{supp}(\nu) \cap ([0,1] \times \{x\}) = \emptyset\Big\}.$$

We claim that

$$\nu_2(R_1) = \int_{\mathrm{supp}(\nu)} \chi_{R_1}(y)\, d\nu_\infty(x,y) = 0.$$

Indeed, the first equality is immediate. To prove the second equality, take $(x, y) \in \text{supp}(\nu)$. We have two possibilities: If $y \notin \pi_1(\text{supp}(\nu))$, then $y \notin R_1$. And if $y \in \pi_1(\text{supp}(\nu))$ we have $(x, y) \in \text{supp}(\nu)$ and then $y \notin R_1$. This shows the claim.

By the other hand, note that $\pi_1^{-1}(R_1) \cap \text{supp}(\nu) = R$, and thus

$$\int_{\text{supp}(\nu)} \chi_{R_1}(x) d\nu(x, y) = \int_{\text{supp}(\nu)} \chi_{\pi_1^{-1}(R_1)}(x, y) d\nu(x, y) = \nu(R) = \epsilon.$$

Now let U be an open set of $[0, 1]$ which contains R_1 and such that $\nu_2(U) < \nu_2(R_1) + \epsilon/2 = \epsilon/2$. Consider a sequence of continuous function f_j such that $f_j \uparrow \chi_U$. Using the monotonous convergence theorem and $\nu \in \mathbf{M}$, we have:

$$\begin{aligned} \epsilon/2 > \nu_2(U) &= \int \chi_U(y) d\nu(x, y) = \lim_j \int f_j(y) d\nu(x, y) \\ &= \lim_j \int f_j(x) d\nu(x, y) = \int \chi_U(x) d\nu(x, y) \\ &\geq \int \chi_{R_1}(x) d\nu(x, y) = \epsilon \end{aligned}$$

which is a contradiction. □

Theorem 5. *Let $\nu \in \mathfrak{M}_0$ be any maximizing measure. If the observable A is C^2, and $\frac{\partial^2 A}{\partial x \partial y} > 0$, then the measure ν is supported on a graph.*

Proof. Let u be any calibrated backward-subaction and $(x_0, y_0) \in \text{supp}(\nu)$, then by Proposition 5 (x_0, y_0) satisfies equation (19).

On the other hand, by Lemma 10, there exists z_0 such that $(z_0, x_0) \in \text{supp}(\nu)$, and this means that $x_0 = y(z_0)$. Thus item (c) of Lemma 9 implies that u is differentiable at x_0.

Now by item (a) of Lemma 9 and Proposition 6, we have that

$$Du(x_0) = \frac{\partial A}{\partial x}(x_0, y_0) \tag{20}$$

Note that, for any fixed p and x, the equation $p = \dfrac{\partial A}{\partial x}(x, y)$ has at most one solution $y(x, p)$ because

$$\frac{\partial^2 A}{\partial x \partial y} > 0.$$

Let $p = Du(x_0)$ and $x = x_0$ then y_0 is the unique point that satisfies the equation (20). □

Remark. Using the same arguments of the proof of Theorem 5, we see that, if u is a calibrated subaction, u is differentiable at x and (x, y) satisfies equation (19) (note that, for each x there exists at least one y with this property), then we get that this y is the unique point that satisfies the equation $Du(x) = \frac{\partial A}{\partial x}(x, y)$. Therefore, y is the unique point that satisfies equation (19).

Lemma 11. *If the observable A is C^2, and $\frac{\partial^2 A}{\partial x \partial y} > 0$, then $\cup_{\nu \in \mathfrak{M}_0} \mathrm{supp}(\nu)$ is contained in a graph.*

Proof. Let ν_1 and ν_2 be two maximizing measures. Suppose there exists $x \in \pi_1(\mathrm{supp}(\nu_1)) \cap \pi_1(\mathrm{supp}(\nu_2))$. Let y_1 and y_2 be the (unique) points such that $(x, y_1) \in \mathrm{supp}(\nu_1)$ and $(x, y_2) \in \mathrm{supp}(\nu_2)$.

Let u be a calibrated backward-subaction, using the same arguments of the proof of Theorem 5 for $(x, y_1) \in \mathrm{supp}(\nu_1)$, and for $(x, y_2) \in \mathrm{supp}(\nu_2)$, we get, respectively

$$Du(x_0) = \frac{\partial A}{\partial x}(x_0, y_1) \quad \text{and} \quad Du(x_0) = \frac{\partial A}{\partial x}(x_0, y_2)\,.$$

But, as before, the equation $p = \frac{\partial A}{\partial x}(x, y)$ has at most one solution $y(x, p)$, then $y_1 = y_2$. $\square$

Definition 12. *Given k and $x, y \in [0, 1]$, we will call a* k-path beginning in x and ending at y *an ordered sequence of points*

$$(x_1, \ldots, x_k) \in [0, 1] \times \cdots \times [0, 1]$$

satisfying $x_1 = x$, $x_k = y$.

We will denote by $\mathcal{P}_k(x, y)$ the set of such k-paths.

Remark.

1) Here we shall note that the results that we will get can not be a particular case of the results obtain in [Gom, GLM] for the theory of Aubry-Mather, because in A-M theory a Lagrangian $L\colon [0, 1] \times \mathbb{R} \to \mathbb{R}$, satisfy the hypothesis that $L(x, v) \to +\infty$ when $|v| \to \infty$.

2) A path in A-M theory (see [GLM]) is an orderer sequence of points $(x_0, \ldots, x_k) \in \mathbb{R}^N \times \cdots \times \mathbb{R}^N$ such that for each x_j we associate a velocity $v_j = x_{j+1} - x_j$, $0 \le j < k$. With those pairs (x_j, v_j) we are able to calculate the action of the path $(x_0, \ldots, x_k)$. In our setting there is no velocity and only the points of the path are used to calculate the action of the path.

Definition 13. *A point $x \in [0,1]$ is called* non-wandering with respect to A *if, for each $\epsilon > 0$, there exists $k \geq 1$ and a k-path $(x_1, \dots, x_k)$ in $\mathcal{P}_k(x,x)$ such that*

$$\left| \sum_{i=1}^{k-1} (A - m)(x_i, x_{i+1}) \right| < \epsilon.$$

We will denote by $\Omega(A)$ the set of non-wandering points with respect to A.

The above definition is analogous (to the case of discrete time dynamics) to the continuous time one in Mather Theory (see [Fathi, CI, GLM]). There, a point x is non wandering for the Lagrangian L, if you can move from x to x by means of connecting paths γ, with action $\int L(\gamma, \gamma')\, dt$ so small as you want.

Lemma 12. *Suppose that the observable A is C^2, and $\frac{\partial^2 A}{\partial x \partial y} > 0$. Let $\nu \in \mathfrak{M}_0$ be any maximizing measure. We claim that $\pi_1(\mathrm{supp}(\nu)) \subset \Omega(A)$.*

Proof. Let u be a backward calibrated subaction, and $\mathrm{dom}(Du)$ be the set of differentiable points of u. Let $Y_0 \colon \mathrm{dom}(Du) \to [0,1]$ be the map defined by $Y_0(x) = y$, where y is the unique point such that (x, y) satisfies (19) (see the remark after Theorem 5). As we will see in Proposition 13, this map is monotonous, hence we can define a measurable map $Y \colon [0,1] \to [0,1]$, by

$$Y(x) = Y_0(x) \quad \text{if } x \in \mathrm{dom}(Du), \quad \text{and}$$
$$Y(x) = \lim_{z \to x^-, z \in \mathrm{dom}(Du)} Y_0(z) \quad \text{if } x \notin \mathrm{dom}(Du).$$

Note that $\nu_\infty \circ \pi_1^{-1}$-a.e. $\pi_1(\mathrm{supp}\, \nu_\infty) \subset \mathrm{dom}(Du)$.

Let us prove that $\nu_\infty \circ \pi_1^{-1}$ is an invariant measure for Y. Indeed, for $f \in C^0(\Omega(A))$, we have that:

$$\int f \circ Y(x)\, d\nu_\infty \circ \pi_1^{-1}(x) = \int f \circ Y(x)\, d\nu_\infty(x, y)$$
$$= \int f(y)\, d\nu_\infty(x, y) = \int f(x)\, d\nu_\infty(x, y) = \int f(x)\, d\nu_\infty \circ \pi_1^{-1}(x),$$

where in the second equality we used the fact that, if $(x, y) \in \mathrm{supp}(\nu_\infty)$, then $y = Y(x)$, and in the third equality we used item (a) of Lemma 1.

Take $(x, y) \in \mathrm{supp}\, \nu_\infty$ and B a ball centered in x. We can see that $\pi_1^{-1}(B)$ is an open set which contains (x, y), and this implies $\nu_\infty \circ \pi_1^{-1}(B) > 0$. Using Poincaré recurrence theorem, there exists $x_1 \in B \cap \mathrm{dom}(Du)$ such that, for infinitely many $j's$, $x_{j+1} := Y^j(x_1)$ is in B.

Note that the points x_j satisfy the following equation:

$$u(x_j) - u(x_{j+1}) = A(x_j, x_{j+1}) - m$$

because, by Lemma 9, u is differentiable in each x_j and then there exists only one $y(x_j)$ (that coincides with x_{j+1}) which satisfies the equation (19).

We fix $\epsilon > 0$ and $x_j \in B$, we can construct the following path: $(\tilde{x}_1, \dots, \tilde{x}_j) = (x, x_2, \dots, x_{j-1}, x)$, and we have that

$$\sum_{i=1}^{j-1} (A - m)(\tilde{x}_i, \tilde{x}_{i+1})$$

$$= u(x_1) - u(x_j) + A(x, x_2) - A(x_1, x_2) + A(x_{j-1}, x) - A(x_{j-1}, x_j) \leq \epsilon$$

if B is small enough, because u is Lipschitz (and A is C^2). □

Definition 14. *Let us define*

$$S_k(x, y) = \inf_{(x_1, \dots, x_k) \in \mathcal{P}_k(x,y)} \left[-\sum_{i=1}^{k-1} (A - m)(x_i, x_{i+1}) \right].$$

We call Mañé potential *the function* $S\colon [0, 1] \times [0, 1] \to \mathbb{R}$ *defined by*

$$S(x, y) = \inf_k S_k(x, y),$$

and Peierls barrier *the function* $h\colon [0, 1] \times [0, 1] \to \mathbb{R} \cup \{+\infty\}$ *defined by*

$$h(x, y) = \liminf_{k \to \infty} S_k(x, y).$$

The value $h(x, y)$ (or, $S(x, y)$) measures, in a certain sense, the cost to move from x to y. This will be a main tool for showing the uniqueness of the calibrated subaction.

It is easy to see that

$$\Omega(A) = \{x \in [0, 1]\colon S(x, x) = h(x, x) = 0\}.$$

The functions S and h have the following properties

(a) if $x, y, z \in [0, 1]$ then $S(x, z) \leq S(x, y) + S(y, z)$.

(b) $S(\cdot, y)$ is a forward-subaction and $S(x, \cdot)$ is a backward-subaction.

(c) $h(\cdot, y)$ is a calibrated forward-subaction and $h(x, \cdot)$ is a calibrated backward-subaction.

Definition 15. *We will say that a property is* generic for A, $A \in C^2([0,1]^2)$, in Mañé's sense*, if the property is true for* $A+f$*, for any* f*,* $f \in C^2([0,1])$*, in a set* G *which is generic (in Baire sense).*

We want to prove that, for A which is generic in Mañé's sense [Man], the functions V and $\bar{V}$ are unique (up to a constant). To do that, first we show that generically the maximizing measure is unique, as we will see in the following proposition.

Proposition 7. *Suppose that the observable* A *is* C^2*, and* $\dfrac{\partial^2 A}{\partial x \partial y} > 0$*. Then the set*

$$G_2 = \left\{ f \in C^2([0,1]) \,|\, \mathfrak{M}_0(A+f) = \{\nu\} \ \textit{and} \ \pi_1(\mathrm{supp}(\nu)) = \Omega(A+f) \right\}$$

is generic (in Baire sense) in $C^2([0,1])$.

We will use a result of [BC] in order to prove Proposition 7. First we will show that

$$G_1 = \left\{ f \in C^2([0,1]) \,|\, \mathfrak{M}_0(A+f) = \{\nu\} \right\} \tag{21}$$

is generic (in Baire sense).

Remark. We point out that if one considers above, in the definition of G_2, potentials of the form $A(x,y) + l\,x$, where l is constant, instead of $A(x,y) + f(x)$, the same result is true for a generic $l \in \mathbb{R}$. This new statement is natural (and means something interesting) once it is common to consider a magnetization as a function of this form. In this way, for example, considering fixed the term $\frac{1}{2}(x-y)^2$, for a dense set of $l \in \mathbb{R}$, we have that the zero-temperature state for $A(x,y) = \frac{1}{2}(x-y)^2 + l\,x$ is unique.

Let us fix some notation: C is the set of continuous functions in $[0,1]^2$, $F = C^*$ the vector space of continuous functionals $\nu : C \to \mathbb{R}$, $E = C^2([0,1])$ provided with the C^2 topology, and G is the vector space of finite Borel signed measures on $[0,1]$. $K \subset G$ is the set of Borel probability measure on $[0,1]$, and note that $\mathbf{M} \subset F$. We denote by $F_A : \mathbf{M} \to \mathbb{R}$ the linear functional defined by $F_A(\nu) = \int A\, d\nu$. Note that $\mathfrak{M}_0(A)$ is the set of points of $\nu \in \mathbf{M}$ which maximize $F_A|_{\mathbf{M}}$. Finally, let $\pi : F \to G$ be the projection induced by $\pi_1 : [0,1]^2 \to [0,1]$.

Lemma 13. *There exists a generic subset* $\mathcal{O} \subset E$ *(in Baire sense) such that, for all* $f \in \mathcal{O}$*, we have*

$$\#\pi\left(\mathfrak{M}_0(A+f)\right) = 1.$$

Proof. We just note that F_A is a affine subspace of dimension 0 of $\mathbf{M}^*$, then proposition follows by Theorem 5 of [BC]. $\square$

Note that, in order to have (21), we need to prove that $\#\mathfrak{M}_0(A+f)=1$.

Lemma 14. *If the observable A is C^2, and $\frac{\partial^2 A}{\partial x \partial y}>0$, then we have $\#\mathfrak{M}_0(A)=\#\pi(\mathfrak{M}_0(A))$.*

Proof. By Lemma 11 we know that the restriction to $\cup_{\nu\in\mathfrak{M}_0}\operatorname{supp}(\nu)$ of the projection $[0,1]^2\to[0,1]$ is a injective map. Hence the linear map $\pi:\mathfrak{M}_0(A)\to G$ is injective, and $\#\pi(\mathfrak{M}_0(A))=\#\mathfrak{M}_0(A)$. $\square$

Proof of Proposition 7. Note that, by Lemmas 13 and 14, we have that the set G_1 given in (21) is generic.

Let $f_0\in G_1$, and $f_1\in C^2([0,1])$ such that $f_1\geq 0$ and $\{x: f_1(x)=0\}=\pi_1(\operatorname{supp}(\nu))$. Then $\pi_1(\operatorname{supp}(\nu))\subset\Omega(A+f_0+f_1)$.

Claim. If $x_1\notin\pi_1(\operatorname{supp}(\nu))$ then $x_1\notin\Omega(A+f_0+f_1)$.

Indeed, $f_1(x_1)>0$, and

$$\begin{aligned} h^{(A+f_0+f_1)}(x_1,x_1) &= \liminf_{k\to\infty}\left(\inf_{\mathcal{P}_k(x_1,x_1)}\sum_{i=1}^{k-1}(A+f_0+f_1-m)(x_i,x_{i+1})\right)\\ &\geq \liminf_{k\to\infty}\left(\inf_{\mathcal{P}_k(x_1,x_1)}\sum_{i=1}^{k-1}(A+f_0-m)(x_i,x_{i+1})+f_1(x_1)\right)\\ &= h^{(A+f_0)}(x_1,x_1)+f_1(x_1)>0. \end{aligned}$$

Hence $\pi_1(\operatorname{supp}(\nu))=\Omega(A+f_0+f_1)$. $\square$

Proposition 8. *If u is a calibrated backward-subaction, then for any x we have*

$$u(x)=\sup_{p\in\Omega(A)}\{u(p)-h(p,x)\}.$$

Proof. For $(x_1,\ldots,x_k)\in\mathcal{P}_k(x,\bar{x})$, we have

$$u(x_i)-u(x_{i+1})\geq A(x_i,x_{i+1})-m$$

and

$$u(x_k) - u(x_1) \leq -\sum_{i=1}^{k-1} A(x_i, x_{i+1}) - m\,.$$

Hence, $u(\bar{x}) - u(x) \leq h(x, \bar{x})$, and therefore

$$u(x) \geq \sup_{p \in \Omega(A)} \{u(p) - h(x, p)\}.$$

Now we show the other inequality. We denote by $x_1 = x$. The fact that u is a backward calibrated subaction implies the existence of x_2 such that $u(x_1) = u(x_2) + A(x_1, x_2) - m$. Thus, recursively, we can construct $(x_1, x_2, \ldots, x_n, \ldots)$ such that $u(x_n) = u(x_{n+1}) + A(x_n, x_{n+1}) - m$.

Let p be an accumulation point of the sequence $\{x_n\}$. We claim that $p \in \Omega(A)$. Indeed, if $x_{n_j} \to p$, we fix $j > i$, and then we construct $(\tilde{x}_1, \ldots, \tilde{x}_{n_j - n_i}) = (p, x_{n_{i+1}}, \ldots, x_{n_{j-i}}, p)$. Hence, we have

$$\begin{aligned}
\sum_{i=1}^{n_j - n_i - 1} (A - m)(\tilde{x}_i, \tilde{x}_{i+1}) &= \sum_{k=n_i}^{n_{j-1}} (A - m)(x_k, x_{k+1}) + A(p, x_{n_{i+1}}) \\
&\quad - A(x_{n_i}, x_{n_{i+1}}) + A(x_{n_{j-1}}, p) - A(x_{n_{j-1}}, x_{n_j}) \\
&= u(x_{n_j}) - u(x_{n_i}) + A(p, x_{n_{i+1}}) - A(x_{n_i}, x_{n_{i+1}}) \\
&\quad + A(x_{n_{j-1}}, p) - A(x_{n_{j-1}}, x_{n_j})
\end{aligned}$$

Then for $\epsilon > 0$ fixed and i large enough we have that

$$\left| \sum_{i=1}^{n_j - n_i - 1} (A - m)(\tilde{x}_i, \tilde{x}_{i+1}) \right| \leq \epsilon\,.$$

Therefore $p \in \Omega(A)$.

Now take $(\tilde{x}_1, \ldots, \tilde{x}_{n_j}) = (x_1, x_2 \ldots, x_{n_j - 1}, p)$. We have

$$\begin{aligned}
&- \sum_{i=1}^{n_j - 1} (A - m)(\tilde{x}_i, \tilde{x}_{i+1}) + u(x) - u(p) \\
&= - \sum_{i=1}^{n_j - 1} (A - m)(x_i, x_{i+1}) + A(x_{n_j - 1}, x_{n_j}) \\
&\quad - A(x_{n_j - 1}, p) + u(x) - u(p) \\
&= u(x_{n_j}) - u(p) + A(x_{n_j - 1}, x_{n_j}) - A(x_{n_j - 1}, p)\,.
\end{aligned}$$

Given $k > 0$ there exists n_k such that

$$-\sum_{i=1}^{n_k-1}(A-m)(\tilde{x}_i, \tilde{x}_{i+1}) \leq u(p) - u(x) + \frac{1}{k}\,.$$

Making $k \to \infty$ we obtain $h(x, p) \leq u(p) - u(x)$. Then

$$u(x) = \sup_{p\in\Omega(A)} \{u(p) - h(x, p)\}. \qquad \square$$

Proposition 9. *There exists a bijective correspondence between the set of calibrated backward-subactions and the set of functions $f \in C^0(\Omega(A))$ satisfying $f(y) - f(x) \leq h(x, y)$, for all points x, y in $\Omega(A)$.*

Proof. Let us suppose that f satisfies $f(y) - f(x) \leq h(x, y)$. We define the following map

$$f \mapsto u_f(x) := \sup_{p\in\Omega(A)} \{f(p) - h(x, p)\}.$$

We will just show that this map is a bijection. The proof of the fact that u_f is a calibrated backward-subaction is similar to the proof of Theorem 13 in [GL].

We will prove that the map is injective: let $f \in C^0(\Omega(A))$ satisfying $f(y) - f(x) \leq h(x, y)$. For $x \in \Omega(A)$, we have that $h(x, x) = 0$, and hence

$$f(p) - h(x, p) \leq f(x) \leq \sup_{p\in\Omega(A)} \{f(p) - h(x, p)\} = u_f(x)\,.$$

Then $u_f(x) = f(x), \forall x \in \Omega(A)$. Therefore $f \neq \tilde{f}$ implies $u_f \neq u_{\tilde{f}}$.

Now, we will prove that the map is surjective: let u be a calibrated subaction. Define $f = u|_{\Omega(A)}$. By Proposition 8, we have that f satisfies

$$f(y) - f(x) \leq h(x, y)$$

and

$$u(x) = \sup_{p\in\Omega(A)} \left\{u(p) - h(x, p)\right\} = \sup_{p\in\Omega(A)} \left\{f(p) - h(x, p)\right\} = u_f(x). \qquad \square$$

Now suppose that A has a unique maximizing measure ν_∞ and also that $\pi_1(\text{supp}(\nu_\infty)) = \Omega(A)$. As we have explained in the proof of Lemma 12 above, we can define a measurable map $Y\colon \Omega(A) \to \Omega(A)$. Indeed, when x is such that there is unique y satisfying $(x, y) \in \text{supp}(\nu_\infty)$, then $y = Y(x)$. In the other case, we define Y via the limit coming from the left side.

Lemma 15. *If A is generic in the Mañé sense, then the measure $\nu_\infty \circ \pi_1^{-1}$ is an invariant ergodic measure for Y.*

Proof. First we prove the invariance: Let $f \in C^0(\Omega(A))$. We have:

$$\begin{aligned}\int f \circ Y(x)\, d\nu_\infty \circ \pi_1^{-1}(x) &= \int f \circ Y(x)\, d\nu_\infty(x,y) = \int f(y)\, d\nu_\infty(x,y) \\ &= \int f(x)\, d\nu_\infty(x,y) = \int f(x)\, d\nu_\infty \circ \pi_1^{-1}(x)\,.\end{aligned}$$

Now we will prove that Y is uniquely ergodic: let η be a measure in the Borel sets of $\Omega(A)$ which is invariant for Y. If we define, for each Borel set B of $[0,1]^2$, $\nu(B) = \eta(\pi_1(B \cap \text{supp}(\nu_\infty)))$, we have that ν is a measure probability in $[0,1]^2$ such that

(1) $\text{supp}(\nu) \subset \text{supp}(\nu_\infty)$,

(2) $\pi_1(\nu) = \eta$,

(3) $\nu \in \mathbf{M}$.

In order to prove (3), consider $f \in C([0,1])$. We have

$$\begin{aligned}\int f(y) d\nu(x,y) &= \int f(Y(x)) d\nu(x,y) = \int f(Y(x)) d\nu(x) \\ &= \int f(x) d\nu(x) = \int f(x) d\nu(x,y)\,,\end{aligned}$$

where we used, in sequence: (1); (2); η is Y-invariant; (2).

Note that for any calibrated backward-subaction u we have

$$\int A(x,y) d\nu(x,y) = \int (u(x) - u(y) + m)\, d_\nu(x,y) = m\,,$$

where in the first equality we used (1) and Proposition 5, and, in the second equality we used (3). Thus we have that ν is a maximizing measure, and by uniqueness $\nu = \nu_\infty$. This implies $\eta = \pi_1(\nu_\infty)$, which shows that there exists an unique invariant measure for Y, which is a ergodic measure. □

Proposition 10. *If $\nu \circ \pi_1^{-1}$ is an ergodic measure in $[0,1]$, and u, u' are two calibrated backward-subactions for A, then $u - u'$ is constant in $\pi_1(\text{supp}(\nu))$.*

For the proof of this proposition see Theorem 17 of [GL].

Theorem 6. *If A is generic in the Mañé sense, then the set of calibrated backward-subactions has an unique element.*

Proof. By the hypothesis ν_∞ is the unique maximizing measure, hence $\nu_\infty \circ \pi_1^{-1}$ is ergodic, and $\pi_1(\text{supp}(\nu_\infty)) = \Omega(A)$.

Let $f, f' \colon \Omega(A) \to \mathbb{R}$ be continuous functions satisfying the hypothesis of Proposition 9. In the proof of Proposition 9 we see that we can get two calibrated subactions u_f, $u_{f'}$ such that $f - f' = u_f - u_{f'}$ in $\Omega(A)$, and hence, by Proposition 10 $u_f - u_{f'}$ is constant in $\Omega(A)$. Again, from Proposition 9, we show that the set of calibrated backward-subactions has an unique element. □

If we consider V and $\bar{V}$ given in Lemma 5, Theorem 6 proves that $\bar{V}$ is unique. The proof that V is unique uses similar arguments.

5 The shift in the Bernoulli space $[0,1]^{\mathbb{N}}$, and a Large Deviation Principle

Let us come back to the maximization problem, over $\mathcal{M}_0$, of

$$\int A d\mu \,. \tag{22}$$

We get in this section (and from what we proved before) a family of absolutely continuous Markov measures μ_β, indexed by a real parameter β, and this family of measures weakly converges, when $\beta \to \infty$, to the maximizing measure μ_∞. A natural question is to know the speed (in logarithm scale) of convergence of the probability $\mu_\beta(C) \to 0$, of a μ_∞-null set C, when $\beta \to \infty$. In this direction we will present a Large Deviation Principle. This is our main goal in this section.

The following proposition allows us to conclude that, generically in Mañé's sense, all such maximizing measures, after projection in the first two coordinates, are unique.

Proposition 11. *Suppose that ν_∞ is a maximizing measure in* **M**.

(i) *If A has an unique maximizing measure in* **M**, *then any maximizing measure in $\mathcal{M}_0$ is projected by Π in ν_∞, where $\Pi \colon [0,1]^{\mathbb{N}} \to [0,1]^2$ is the projection in the first two coordinates.*

(ii) *ν_∞ can be extended to a maximizing measure $\mu_\infty \in \mathcal{M}_0$ which is a stationary Markov measure.*

(iii) *If ν_β is the family of measures given by* (16)*, then this measures can be extended to absolutely continuous Markov measures μ_β, and this sequence of measures weakly converge to the maximizing measure μ_∞.*

Proof. Item (i) follows by items (b) and (c) of Proposition 1 and by Proposition 7. Item (ii) follows by item (a) of Proposition 1. Item (iii) is a consequence of the remark after the proof of Proposition 1. □

From now on, until the end of this section, we will suppose that the maximizing measure ν_∞, and the functions V and $\bar{V}$ are unique. This is a generic property in Mañé sense.

Thus, for the maximization problem in the Bernoulli shift, we have shown the existence of a maximizing measure μ_∞ which can be approximated by absolutely continuous stationary Markov measures μ_β, which were explicitly calculated.

Now we will show a Large Deviation Principle for the family of measures $\{\mu_\beta\}$. We will also exhibit a Large Deviation Principle for the bidimensional measures ν_β which, by the earlier sections, converge to ν_∞.

Lemma 16. *Suppose $k \geq 2$. Let $F_k \colon [0,1]^k \to \mathbb{R}$ be the function given by*

$$F_k(x_1, \ldots, x_k) := \max(V + \bar{V}) - V(x_1) - \bar{V}(x_k) - \sum_{i=1}^{k-1} (A - m)(x_i, x_{i+1}).$$

Let $D_k = A_1 \ldots A_k$ be a cylinder of size k. Then, there exists the limit

$$\lim_{\beta\to\infty} \frac{1}{\beta} \log \mu_\beta(D_k) = - \inf_{(x_1,\ldots,x_k)\in D_k} F_k(x_1, \ldots, x_k).$$

Proof. Let us define

$$f_{k,\beta}(x_1, \ldots, x_k) := \frac{1}{\beta} \log \pi_\beta + \frac{k-1}{\beta} \log \lambda_\beta - \sum_{i=1}^{k-1} A(x_i, x_{i+1}) - \frac{1}{\beta} \log \varphi_\beta(x_1) - \frac{1}{\beta} \log \bar{\varphi}_\beta(x_k).$$

We have that $f_{k,\beta} \to F_k$ uniformly when $\beta \to \infty$. This is a consequence of the uniqueness of V and $\tilde{V}$.

We begin by proving the

Claim. Let $C_k = A_1 \ldots A_k$ be a cylinder of size k. We have

$$\limsup_{\beta\to\infty} \frac{1}{\beta} \log \mu_\beta(C_k) \leq - \inf_{(x_1,\ldots,x_k)\in C_k} F_k(x_1, \ldots, x_k)\,.$$

To prove the Claim, note that we have

$$\begin{aligned}\mu_\beta(C_k) &= \int_{A_1}\dots\int_{A_k}\frac{e^{\beta A(x_{k-1},x_k)}\bar{\varphi}_\beta(x_k)}{\bar{\varphi}_\beta(x_{k-1})\lambda_\beta}\\ &\qquad \dots\frac{e^{\beta A(x_1,x_2)}\bar{\varphi}_\beta(x_2)}{\bar{\varphi}_\beta(x_1)\lambda_\beta}\;\frac{\varphi_{\beta(x_1)}\bar{\varphi}_{\beta(x_1)}}{\pi_\beta}dx_k\dots dx_1 \\ &= \int_{A_1}\dots\int_{A_k} e^{-\beta f_{k,\beta}(x_1,\dots x_k)}dx_k\dots dx_1 \\ &\le e^{-\beta\inf_{C_k} f_{k,\beta}(x_1,\dots x_k)}|C_k|\,,\end{aligned}\tag{23}$$

where $|C_k|$ denotes the Lebesgue measure of C_k. Hence

$$\frac{1}{\beta}\log\mu_\beta(C_k)\le -\inf_{C_k} f_{k,\beta}(x_1,\dots x_k)+\frac{1}{\beta}\log|C_k|,$$

and then, by the uniform convergence, we have:

$$\limsup_{\beta\to\infty}\frac{1}{\beta}\log\mu_\beta(C_k)\le -\inf_{C_k} F_k(x_1,\dots x_k)\,,$$

which finishes the proof of the Claim.

Now we will prove the lemma: if we fix $\delta>0$, using the continuity of F_k we can find a point $(x_1,\dots,x_k)\in D_k^0$ (the interior of D_k) such that

$$\inf_{D_k} F_k\le F_k(x_1,\dots,x_k)<\inf_{D_k} F_k+\delta.\tag{24}$$

Now, let D_δ be a cylinder of size k, such that $(x_1,\dots,x_k)\in D_\delta\subset D_k^0$, and

$$\inf_{D_k} F_k\le F_k(y_1,\dots,y_k)<\inf_{D_k} F_k+2\delta\,\forall(y_1,\dots,y_k)\in D_\delta\,.\tag{25}$$

We have that

$$\mu_\beta(D_k)\ge\mu_\beta(D_\delta)\ge e^{-\beta\sup_{D_\delta} f_{k,\beta}(y_1,\dots,y_k)}|D_\delta|,$$

where the last inequality cames from (23). Now we use again the uniform convergence of $f_{k,\beta}$ to F_k in order to get

$$\liminf_{\beta\to\infty}\frac{1}{\beta}\log\mu_\beta(D_k)\ge -\sup_{D_\delta} F_k\,.$$

By (25), we get

$$\liminf_{\beta\to\infty}\frac{1}{\beta}\log\mu_\beta(D_k)\ge -\inf_{D_k} F_k-2\delta\tag{26}$$

Sending $\delta\to 0$, and using the Claim, we finish the proof of the lemma. □

Note that if we set $k=2$ above, we get a LDP for the family $\nu_\beta\to\nu_\infty$.

Theorem 7. *Let $I : [0,1]^{\mathbb{N}} \to [0,+\infty]$ be the function defined by*

$$I(\mathbf{x}) := \sum_{i \geq 1} V(x_{i+1}) - V(x_i) - (A-m)(x_i, x_{i+1}) .$$

Let $D = A_1 \dots A_k$ be a cylinder of any size k. Then, there exists the limit

$$\lim_{\beta \to \infty} \frac{1}{\beta} \log \mu_\beta(D) = - \inf_{\mathbf{x} \in D} I(\mathbf{x}) .$$

Note that, by Lemma 5,

$$V(x_{i+1}) - V(x_i) - A(x_i, x_{i+1}) + m \geq 0,$$

therefore the sequence of partial sums of the series in the definition of $I(\mathbf{x})$ is a non-decreasing sequence. This shows that $I(\mathbf{x})$ is well defined (note that $I(x)$ can be $+\infty$).

In order to prove Theorem 7 we will need some new results and definitions.

For each $N \geq 2$, let us extend the function F_N to the space $[0,1]^{\mathbb{N}}$:

$$\begin{aligned} F_N(\mathbf{z}) &:= F_N(z_1, \dots, z_n) \\ &= \sup(V + \bar{V}) - V(z_1) - \bar{V}(z_N) - \sum_{i=1}^{N-1} (A-m)(z_i, z_{i+1}). \end{aligned}$$

Lemma 17. $\forall\, \mathbf{z} \in [0,1]^{\mathbb{N}}$*, we have*

$$F_N(\mathbf{z}) \geq \max(V + \bar{V}) - (V(z_1) + \bar{V}(z_1)) \geq 0 .$$

Proof. By Lemma 5

$$\bar{V}(x) - \bar{V}(y) \geq A(x,y) - m, \ \forall x, y,$$

then

$$- \sum_{i=1}^{N-1} (A-m)(z_i, z_{i+1}) \geq \bar{V}(z_N) - \bar{V}(z_1) .$$

Hence, by definition of F_N

$$F_N(\mathbf{z}) \geq \max(V + \bar{V}) - (V(z_1) + \bar{V}(z_1)) . \qquad \square$$

Lemma 18.

(a) *for a fixed* $\mathbf{x} \in [0, 1]^{\mathbb{N}}$*, we have that*

$$V(x_1) + \sum_{i=1}^{k-1} (A - m)(x_i, x_{i+1}) + \bar{V}(x_k)$$

is decreasing with respect to k.

(b) *If* $I(\mathbf{x}) < +\infty$*, then there exists the limit*

$$L(\mathbf{x}) = \lim_{k \to +\infty} V\big(\sigma^k(\mathbf{x})\big) + \bar{V}\big(\sigma^k(\mathbf{x})\big).$$

Proof.

(a)

$$\begin{aligned} V(x_1) &+ \sum_{i=1}^{k} (A - m)(x_i, x_{i+1}) + \bar{V}(x_{k+1}) \\ &= V(x_1) + \sum_{i=1}^{k-1} (A - m)(x_i, x_{i+1}) + \bar{V}(x_k) + A(x_k, x_{k+1}) - m \\ &\quad + \bar{V}(x_{k+1}) - \bar{V}(x_k), \end{aligned}$$

by Lemma 5 (remember that $\tilde{m} = m$) $A(x_k, x_{k+1}) - m + \bar{V}(x_{k+1}) - \bar{V}(x_k) \leq 0$, and we have (a).

(b) We have

$$\begin{aligned} I(\mathbf{x}) &= \sum_{i \geq 1} V(x_{i+1}) - V(x_i) - (A - m)(x_i, x_{i+1}) = \lim_{k \to +\infty} V(x_k) \\ &\quad + \bar{V}(x_k) - \lim_{k \to \infty} \left(V(x_1) + \sum_{i=1}^{k-1} (A - m)(x_i, x_{i+1}) + \bar{V}(x_k) \right) \end{aligned} \tag{27}$$

Hence, if $I(\mathbf{x}) < +\infty$, it follows, thanks to item (a), that

$$V(x_k) + \bar{V}(x_k) = V\big(\sigma^k(\mathbf{x})\big) + \bar{V}\big(\sigma^k(\mathbf{x})\big)$$

must converge. □

Lemma 19. *Suppose $I(\mathbf{x}) < +\infty$. Then, if we define, for each $M \in \mathbb{N}$, the probability measure*

$$\mu_M = \frac{1}{M} \sum_{j=1}^{M-1} \delta_{\sigma^j(\mathbf{x})} ,$$

we have that $\Pi(\mu_M) \to \nu_\infty$ in the weak-$\star$ topology (where Π is the projection in the two first coordinates).

Proof. Given $\epsilon > 0$, there exists $N_\epsilon \in \mathbb{N}$ such that , for all $N \geq N_\epsilon$, and all $M > N$,

$$\sum_{i=N}^{M-1} V(x_{i+1}) - V(x_i) - (A - m)(x_i, x_{i+1}) < \epsilon .$$

Thus

$$V(x_M) - V(x_N) + (M - N)m < \sum_{i=N}^{M-1} A\big(\sigma^i(\mathbf{x})\big) + \epsilon ,$$

and

$$\frac{1}{M - N} \sum_{i=N}^{M-1} A\big(\sigma^i(\mathbf{x})\big) > m + \frac{V(x_M) - V(x_N)}{M - N} - \frac{\epsilon}{M - N} ,$$

and then we get that

$$\liminf_{M \to +\infty} \frac{1}{M} \sum_{i=1}^{M-1} A\big(\sigma^i(\mathbf{x})\big) \geq m .$$

Now we remember that

$$\frac{1}{M} \sum_{i=1}^{M-1} A\big(\sigma^i(\mathbf{x})\big) = \int A d\mu_M \leq m ,$$

and finally we get

$$\lim_{M \to +\infty} \int A d\mu_M = \lim_{M \to +\infty} \frac{1}{M} \sum_{i=1}^{M-1} A\big(\sigma^i(\mathbf{x})\big) = m .$$

If we use the compactness of the closed ball of radius 1 in the weak-$\star$ topology, we get that $\{\mu_M\}$ has convergent subsequences. Any limit of a convergent subsequence is a stationary measure (a σ-invariant measure) and must be a maximizing measure, by the last equality. As any maximizing measure is projected in μ_∞ by Π, we get the lemma. □

Proposition 12. *If $I(\mathbf{x}) < +\infty$, then*

$$\lim_{k\to+\infty} V(\sigma^k(\mathbf{x})) + \bar{V}(\sigma^k(\mathbf{x})) = \max(V + \bar{V}).$$

Proof. Let

$$\mathbf{z} = (z_1, z_2, z_3, \ldots) \in \operatorname{supp}(\mu_\infty).$$

We have that $(z_1, z_2) \in \operatorname{supp}(\nu_\infty)$. Thus, by Lemma 19 there exists a subsequence such that $\Pi(\sigma^{k_l}(\mathbf{x})) \to (z_1, z_2)$.

Fix $\epsilon > 0$. Let

$$B_{k_l,\epsilon}(\mathbf{x}) := \left\{\mathbf{y} \in [0,1]^{\mathbb{N}} : |y_j - x_{j+k_l}| \le \epsilon,\ \forall\, 1 \le j \le 2\right\}$$

be the closed cylinder of size 2 'centered' at $\sigma^{k_l}(\mathbf{x})$.

If l is big enough, we have that

$$B_{k_l,\epsilon}(\mathbf{x}) \subset \left\{\mathbf{y} \in [0,1]^{\mathbb{N}} : |y_j - z_j| \le 2\epsilon,\ \forall\, 1 \le j \le 2\right\}.$$

Note that $\mu_\infty(B_{k_l,\epsilon}(\mathbf{x})) = \nu_\infty(B_{k_l,\epsilon}(\mathbf{x})) > 0$, and thus using Lemma 16 with $k = 2$, it follows that there exists a point $(z_{1,\epsilon}, z_{2,\epsilon}, z_{3,\epsilon}, z_{4,\epsilon}, \ldots) \in B_{k_l,\epsilon}(\mathbf{x})$, such that $F_2((z_{1,\epsilon}, z_{2,\epsilon})) = 0$.

Then, we can use the fact that F_2 depends only on its first 2 coordinates in order to obtain that $F_2(\mathbf{w}_\epsilon) = 0$, where $\mathbf{w}_\epsilon = (z_{1,\epsilon}, z_{2,\epsilon}, z_3, z_4, \ldots)$ is defined by the point of $[0,1]^{\mathbb{N}}$ whose first 2 coordinates are equal to those of $(z_{1,\epsilon}, z_{2,\epsilon})$, while the other coordinates are equal to those of $\mathbf{z}$.

Now, if we send $\epsilon \to 0$, we have that $\mathbf{w}_\epsilon \to \mathbf{z}$. Thus we can use the continuity of F_N to get that $F_2(\mathbf{z}) = 0$.

Using again the continuity of F_2, we have that $F_2(\sigma^{k_l}(\mathbf{x})) \to 0$.

Lemma 17 shows that

$$\lim_{l\to+\infty} V(\sigma^{k_l}(\mathbf{x})) + \bar{V}(\sigma^{k_l}(\mathbf{x})) = \max(V + \bar{V}),$$

and finally using Lemma 18(b) we prove Proposition 12. □

Proof of Theorem 7. First we need to prove the following claim.

Claim.

$$I(\mathbf{x}) = \max(V + \bar{V}) - \lim_{k\to\infty}\left(V(x_1) + \sum_{i=1}^{k-1}(A - m)(x_i, x_{i+1}) + \bar{V}(x_k)\right).$$

In order to prove the Claim, we have to consider two possibilities: if $I(\mathbf{x}) < +\infty$, then (27) can be combined with Proposition 12 to give the Claim. If $I(\mathbf{x}) = +\infty$, we just have to use the expression

$$I(x) = \lim_{k\to\infty} \left(V(x_k) - V(x_1) - \sum_{i=1}^{k-1} (A-m)(x_i, x_{i+1}) \right).$$

Thanks to Lemma 16, we just have to show that

$$-\inf_{(x_1,\dots,x_k)\in D} F_k(x_1,\dots,x_k) = -\inf_{\mathbf{x}\in D} I(\mathbf{x})\,.$$

We begin by proving that

$$-\inf_{(x_1,\dots,x_k)\in D} F_k(x_1,\dots,x_k) \le -\inf_{\mathbf{x}\in D} I(\mathbf{x})\,.$$

Given $\delta > 0$, there exists a point $(y_1,\dots,y_k) \in D$ such that

$$F_k(y_1,\dots,y_k) < \inf_{(x_1,\dots,x_k)\in C} F_k(x_1,\dots,x_k) + \delta\,.$$

By the definition of F_k,

$$F_k(y_1,\dots,y_k) = \max(V+\bar{V}) - V(y_1) - \bar{V}(y_k) - \sum_{i=1}^{k-1}(A-m)(y_i, y_{i+1}).$$

For each $j \ge k$ we choose a y_{j+1} that satisfies $\bar{V}(y_j) = \bar{V}(y_{j+1}) + A(y_j, y_{j+1}) - m$. Then we define $\mathbf{y} := (y_1, \dots y_k, y_{k+1}, \dots)$.

Second Claim. $I(\mathbf{y}) = F_k(y_1,\dots y_k)$. Indeed,

$$\begin{aligned} F_k(y_1,\dots y_k) &= \max(V+\bar{V}) - \left(V(y_1) + \bar{V}(y_k) + \sum_{i=1}^{k-1}(A-m)(y_i,y_{i+1}) \right) \\ &= \max(V+\bar{V}) - \left(V(y_1) + \bar{V}(y_j) + \sum_{i=1}^{j-1}(A-m)(y_i,y_{i+1}) \right), \\ &\quad \forall j \ge k. \end{aligned}$$

Then, from the reasoning above and the way we choose $\mathbf{y}$, we get that $F_k(y_1,\dots y_k)$ is equal to

$$\max(V+\bar{V}) - \lim_{j\to\infty} \left(V(y_1) + \bar{V}(y_j) + \sum_{i=1}^{j-1}(A-m)(y_i,y_{i+1}) \right) = I(\mathbf{y})\,.$$

This implies that

$$-\inf_{(x_1,\ldots,x_k)\in D} F_k(x_1,\ldots,x_k) < -I(\mathbf{y}) + \delta \leq -\inf_{\mathbf{x}\in D} I(\mathbf{x}) + \delta\,.$$

Making $\delta \to 0$, we have the first inequality.

Now, we will prove the second inequality:

$$-\inf_{\mathbf{x}\in D} I(\mathbf{x}) \leq -\inf_{(x_1,\ldots,x_k)\in D} F_k(x_1,\ldots,x_k)\,.$$

We can use Lemma 18(a), and then we get, by the Claim,

$$\begin{aligned} I(\mathbf{x}) &= \max(V+\bar{V}) - \lim_{j\to\infty}\left(V(x_1) + \sum_{i=1}^{j-1}(A-m)(x_i,x_{i+1}) + \bar{V}(x_j)\right) \\ &\geq \max(V+\bar{V}) - \left(V(x_1) + \sum_{i=1}^{k-1}(A-m)(x_i,x_{i+1}) + \bar{V}(x_k)\right) \\ &= F_k(x_1,\ldots,x_k). \end{aligned}$$

□

Here, finally, we can give the proofs of Theorems 1 and 2:

Proof of Theorem 1.

(a) It follows by Proposition 7 and item (i) of Proposition 11.

(b) Theorem 6 shows that, generically, the set of backward calibrated subactions has an unique element. The proof that the set of forward calibrated subactions has an unique element is similar.

Proof of Theorem 2.

(a) It follows by items (ii) and (iii) of Proposition 11 and Theorem 4.

(b) This is Theorem 7, note that the hypothesis are fulfilled when Theorem 1 is true.

We will finish this section showing the monotonicity of the graph under the twist condition.

Suppose A is C^2 and satisfies

$$\frac{\partial^2 A}{\partial x\partial y}(x,y) > 0\,.$$

Then, for all $x < x'$, $y < y'$ we have that

$$A(x, y) + A(x', y') > A(x, y') + A(x', y). \tag{28}$$

Let $\bar{V}$ be the calibrated backward-subaction define above.

As a consequence of A being C^2, we have that $\bar{V}$ is Lipschitz, hence $\bar{V}$ is differentiable λ-a.e., where λ is the Lebesgue measure. Let $\mathrm{dom}(D\bar{V})$ be the set of points where $\bar{V}$ is differentiable.

Following the proof of Theorem 5, we have that, for $x \in \mathrm{dom}(D\bar{V})$, there exists only one $y(x)$ such that

$$\bar{V}(x) = A(x, y(x)) + \bar{V}(y(x)) - m. \tag{29}$$

Proposition 13. *The function* $Y\colon \mathrm{dom}(D\bar{V}) \to [0, 1]$*, defined by* $Y(x) = y(x)$*,* $y(x)$ *satisfying* (29)*, is monotone nondecreasing.*

Proof. Let $x < x'$. Let us call $z = Y(x)$, $z' = Y(x')$, and suppose that $z > z'$. We know that

$$\bar{V}(x) = A(x, z) + \bar{V}(z) - m,\ \bar{V}(x') = A(x', z') + \bar{V}(z') - m,$$

and

$$\bar{V}(x) \geq A(x, z') + \bar{V}(z') - m,\ \bar{V}(x') \geq A(x', z) + \bar{V}(z) - m.$$

Adding the first two equation and comparing with the summation of the last two, we get that

$$A(x, z) + A(x', z') \geq A(x, z') + A(x', z),$$

for $x < x'$, $z' < z$, which is a contradiction with (28). □

If we assume that $\dfrac{\partial^2 A}{\partial x \partial y}(x, y) < 0$, then a function $Y(x)$ as above can be defined, and it will be monotone non-increasing.

6 Separating subactions

There exist subactions which are not calibrated but that are also special. One can ask about the ones which are minimal in a certain sense: the subcohomological inequality is an equality in the smallest possible set. This subactions are called separeted subactions.

The main goal of this section is to show the existence of a separating subaction (see [GLT, GLM] for related results). The idea is: given a potential A, we can find a subaction u such that, in the cohomological equation, the equality just holds in points x that are on $\Omega(A)$ (where it has to hold, anyway). In this way, we have a criteria to separate points of $\Omega(A)$ from the other ones. We can then consider a new potential $\tilde{A} = A(x, y) + u(x) - u(y)$ where the maximum of $\tilde{A}$ is exactly attained in $\Omega(\tilde{A})$.

Definition 16. *A continuous function $u\colon [0, 1] \to \mathbb{R}$ is called a*

(a) forward-subaction *if, for any $x, y \in [0, 1]$ we have*

$$u(y) \geq A(x, y) + u(x) - m. \tag{30}$$

(b) backward-subaction *if, for any $x, y \in [0, 1]$ we have*

$$u(x) \geq A(x, y) + u(y) - m. \tag{31}$$

Definition 17. *We say that a forward subaction u is* separating *if*

$$\max_x \big[A(x, y) + u(x) - u(y)\big] = m \iff x \in \Omega(A),$$

and a backward subaction u is separating *if*

$$\max_y \big[A(x, y) + u(y) - u(x)\big] = m \iff y \in \Omega(A).$$

We will show the existence of a separating backward-subaction.

Lemma 20. *If $x \in \Omega(A)$ there exists $\mathbf{x} = (x_1, \ldots, x_k, \ldots) \in [0, 1]^{\mathbb{N}}$ such that $x_1 = x$ and*

$$h(x_k, x_1) \leq \sum_{i=1}^{k-1} (A - m)(x_i, x_{i+1}).$$

Proof. If $x \in \Omega(A)$, then there exists a sequence of paths $\{(x_1^n, \ldots, x_{j_n}^n)\}_{n\in\mathbb{N}}$ such that $x_1^n = x_{j_n}^n = x$ and $j_n \to \infty$ satisfying

$$\sum_{j=1}^{j_n-1} (A - m)\big(x_j^n, x_{j+1}^n\big) \to 0. \tag{32}$$

Because $|x_j^n| \leq 1$, there exists a ray $(x_1, \ldots, x_k, \ldots)$ which is the limit of the paths above, the convergence being uniform in each compact part.

Fixed $k \in \mathbb{N}$. For $j_n > k$, we have that

$$S^{j_n - k}\big(x_k, x_1\big) \leq -A\big(x_k, x_{k+1}^n\big) + m - \sum_{j=k+1}^{j_n-1} (A-m)\big(x_j^n, x_{j+1}^n\big),$$

and

$$\begin{aligned} S^{j_n - k}\big(x_k, x_1\big) + \sum_{j=1}^{j_n-1}(A-m)\big(x_j^n, x_{j+1}^n\big) \\ \leq -A\big(x_k, x_{k+1}^n\big) + m + \sum_{j=1}^{k}(A-m)\big(x_j^n, x_{j+1}^n\big). \end{aligned}$$

Hence taking the $\liminf_{n\to\infty}$ and using (32) we obtain

$$h\big(x_k, x_1\big) \leq \sum_{j=1}^{k-1}(A-m)\big(x_j, x_{j+1}\big). \qquad \square$$

Lemma 21. *Let u be any backward-subaction, then for all $x \in \Omega(A)$ we have*

$$\max_y \big\{u(y) - u(x) + A(x,y)\big\} = m.$$

Proof. Using the fact that u satisfies equation (31), for any $(x_1, \ldots, x_k) \in \mathcal{P}_k(x,y)$, we have that

$$u(y) - u(x) \leq -\sum_{j=1}^{k-1}(A-m)\big(x_i, x_{i+1}\big).$$

Hence $u(y) - u(x) \leq h(x,y)$.

Let $x \in \Omega(A)$ and let $\mathbf{x} = (x_1, \ldots, x_k, \ldots)$ be the point in $[0,1]^{\mathbb{N}}$ which exists by Lemma 20.

By Lemma 20 we have that

$$u(x_1) - u(x_k) \leq h\big(x_k, x_1\big) \leq \sum_{j=1}^{k-1}(A-m)\big(x_j, x_{j+1}\big),$$

and, as it is a backward-subaction,

$$u(x_k) - u(x_1) \leq -\sum_{j=0}^{k-1} (A - m)\big(x_j, x_{j+1}\big).$$

In particular, for $k = 1$,

$$u(x_2) - u(x_1) = -A(x_1, x_2) + m.$$

This implies

$$\max_y \big\{u(y) - u(x) + A(x, y)\big\} = m. \qquad \square$$

Lemma 22. *If the observable A is Hölder continuous, then the function $S_x(\cdot) := S(x, \cdot)$ is uniformly Hölder and has the same Hölder constant of A.*

Proof. Let us fix $x, \epsilon > 0$ and $y, z \in [0, 1]$, then there exists $(x_1, \ldots, x_k) \in \mathcal{P}_k(x, y)$ such that

$$|-\sum_{i=1}^{k-1} (A - m)(x_i, x_{i+1})| \leq S(x, y) + \epsilon.$$

Consider now the following path:

$$(\tilde{x}_1, \ldots, \tilde{x}_k) = (x_1, \ldots, x_{k-1}, z) \in \mathcal{P}_k(x, z),$$

then

$$-\sum_{i=1}^{k-1} (A - m)(\tilde{x}_i, \tilde{x}_{i+1}) = -\sum_{i=1}^{k-1} (A - m)(x_i, x_{i+1}) + A(x_{k-1}, y) - A(x_{k-1}, z)\,.$$

Therefore,

$$S(x, z) \leq -\sum_{i=1}^{k-1} (A - m)(\tilde{x}_i, \tilde{x}_{i+1}) \leq S(x, y) + \epsilon + \mathrm{Hol}_\alpha(A)|z - y|^\alpha,\ \forall \epsilon,$$

i.e., $S(x, y) - S(x, z) \leq \mathrm{Hol}_\alpha(A)|z - y|^\alpha$. Changing the role of y and z we obtain $|S(x, y) - S(x, z)| \leq \mathrm{Hol}_\alpha(A)|z - y|^\alpha$, which give us the Hölder continuity of S_x, independently of x. $\square$

Theorem 8. *If the observable A is Hölder continuous, there exists a separating backward-subaction.*

Proof. By definition,

$$S(x,y) \le -A(x,y) + m \;\forall\; y \in [0,1]\,.$$

If $x \notin \Omega(A)$, then $S(x,x) > 0$. Hence

$$S_x(y) - S_x(x) < -A(x,y) + m \;\forall\; y \in [0,1]\,.$$

$\Omega(A)$ is a closed set, and thus for each $x \notin \Omega(A)$ we can find a neighborhood $V_x \subset [0,1]\backslash\Omega(A)$ of x such that

$$S_x(y) - S_x(z) < -A(z,y) + m, \;\forall\; y \in [0,1], \forall\; z \in V_x\,.$$

We can extract, from the family of these neighborhoods $\{V_x\}_{x\notin\Omega(A)}$, a countable family $\{V_{x_j}\}_{j=1}^{\infty}$ which is a covering of $[0,1]\backslash\Omega(A)$.

We define

$$\tilde{S}_{x_j}(z) = S_{x_j}(z) - S_{x_j}(0)\,.$$

S_{x_j} is uniformly Hölder, which implies that $|\tilde{S}_{x_j}(z)| \le \mathrm{Hol}_\alpha(A) z^\alpha$, $\forall\; x_j$, therefore the series

$$u(z) = \sum_{j=1}^{\infty} \frac{\tilde{S}_{x_j}(z)}{2^j}$$

is well defined and uniformly convergent, because $[0,1]$ is compact. Note that u is a infinite convex combination of backward-subactions $\tilde{S}_{x_j}$, then u is also a backward-subaction.

Fix $x \in [0,1]\backslash\Omega(A)$, there exists $k \ge 1$ such that $x \in V_{x_k}$. Now, $\forall y \in [0,1]$ we have

$$\begin{aligned} u(y) - u(x) &= \sum_{j=1}^{\infty} \frac{S_{x_j}(y) - S_{x_j}(x)}{2^j} = \frac{S_{x_k}(y) - S_{x_k}(x)}{2^k} \\ &\quad + \sum_{j\ne k} \frac{S_{x_j}(y) - S_{x_j}(x)}{2^j} < \frac{-A(x,y)+m}{2^k} \\ &\quad + \sum_{j\ne k} \frac{-A(x,y)+m}{2^j} < -A(x,y) + m. \end{aligned}$$

Hence,

$$\max_y \left\{u(y) - u(x) + A(x,y)\right\} < m, \quad \text{if } \; x \notin \Omega(A),$$

and, as u is a backward-subaction, we have by Lemma 21 that

$$\max_y \left\{u(y) - u(x) + A(x,y)\right\} = m, \quad \text{if } \; x \in \Omega(A). \qquad \square$$

Acknowledgments. The authors are partially supported by CNPq, PRONEX – Sistemas Dinâmicos, Instituto do Milênio, and beneficiary of CAPES financial support. J. Mohr was partially supported by CNPQ PhD and also Pos-doc scholarship.

References

[A] S. Adams. *Mathematical Statistical Mechanics*. Max-Plank-Institut fur Math., (2006).

[AL] K. Athreya and S. Lahiri. *Measure Theory and Probability Theory*. Springer Verlag, (2006).

[Ban] V. Bangert. *Mather sets for twist maps and geodesics on tori*. Dynamics Reported, **1** (1998), 1–56.

[BBNg] I. Brevik, J-M. Borven and S. Ng. *Viscous Brane Cohomology with a Brane-Bulk energy interchange term*. General Relativity and Gravitation, **38**(5) (2006), 907–915(9).

[BLT] A. Baraviera, A.O. Lopes and Ph. Thieullen. *A Large Deviation Principle for equilibrium states of Hölder potentials: the zero temperature case*. Stoch. and Dyn., **6** (2006), 77–96.

[Ba] P. Bhattacharya and M. Majumdar. *Random Dynamical Systems*. Cambridge Univ. Press, (2007).

[CG] W. Chou and R. Griffiths. *Ground states of one-dimensional systems using effetive potentials*. Physical Review B, **34**(9) (1986), 6219–6234.

[Cv] M. Cveti, S. Nojiri and S.D. Odintsov. *Black hole thermodynamics and negative entropy in de Sitter and anti-de Sitter Einstein-Gauss-Bonnet gravity*. Nuclear Physics B, **628**, Issues **1-2**, (2002), 295–330.

[BC] P. Bernard and G. Contreras. *A Generic Property of Families of Lagrangian Systems*. Annals of Math., **167**(3) (2008), 1099–1108.

[CI] G. Contreras and R. Iturriaga. *Global minimizers of autonomous Lagrangians*. 22° Colóquio Brasileiro de Matemática, IMPA, (1999).

[CLT] G. Contreras, A.O. Lopes and Ph. Thieullen. *Lyapunov minimizing measures for expanding maps of the circle*. Ergodic Theory and Dynamical Systems, **21** (2001), 1379–1409.

[CG] J.P. Conze and Y. Guivarc'h. *Croissance des sommes ergodiques et principe variationnel*. Manuscript circa (1993).

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

[De] K. Deimling. *Nonlinear Functional Analysis*. Springer Verlag, (1985).

[Dellach] C. Dellacherie. *Probabilities and potential*. North-Holland, (1978).

[DZ] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications*. Springer Verlag, (1998).

[Ev] L.C. Evans. *Weak Convergence Methods for Nonlinear Partial Differential Equations*. Published for the Conference Board of the Mathematical Sciences, Washington, DC, (1990).

[Fathi] A. Fathi. *Théorème KAM faible et théorie de Mather sur les systèmes lagrangiens*. Comptes Rendus de l'Académie des Sciences, Série I, Mathématique, **324** (1997), 1043–1046.

[FS] A. Fathi and A. Siconolfi. *Existence of C^1 critical subsolutions of the Hamilton-Jacobi equations*. Inv. Math., **155** (2004), 363–388.

[GL] E. Garibaldi and A.O. Lopes. *On Aubry-Mather theory for symbolic dynamics*. Ergodic Theory and Dynamical Systems, **28**, Issue **3** (2008), 791–815.

[GLT] E. Garibaldi, A.O. Lopes and Ph. Thieullen. *On separating sub-actions*. Preprint (2006). To appear.

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

[Go] C. Gole. *Sympletic twist maps*. World Sci. Pub Co Inc., (1998).

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

[Gom1] D.A. Gomes. *Calculus of Variations*. IST – Lisboa, (2006).

[GV] D.A. Gomes and E. Valdinoci. *Entropy Penalization Methods for Hamilton-Jacobi Equations*. Adv. Math., **215**(1) (2007), 94–152.

[Hop] E. Hopf. *An inequality for Positive Linear Integral Operators*. Journal of Mathematics and Mechanics, **12**(5) (1963), 683–692.

[Jen1] O. Jenkinson. *Ergodic optimization*. Discrete and Continuous Dynamical Systems, Series A, **15** (2006), 197–224.

[Ju] G. Jumarie. *Relative Information*. Springer Verlag, (1990).

[Ka] S. Karlin. *Total Positivity*. Standford Univ. Press, (1968).

[Lu] E. Lubkin. *Negative entropy, energy, and heat capacity in connection with surface tension: Artifact of a model or real?*. Inter. Journal of Theoretical Physics, **26**(5) (1987), 455–481.

[Man] R. Mañé. *Generic properties and problems of minimizing measures of Lagrangian systems*. Nonlinearity, **9** (1996), 273–310.

[Mat] J. Mather. *Action minimizing invariant measures for positive definite Lagrangian Systems*. Math. Z., **207**(2) (1991), 169–207.

[Mo] I.D. Morris. *A sufficient condition for the subordination principle in ergodic optimization*. Bull. Lond. Math. Soc., **39**(2) (2007), 214–220.

[Mi] I. Mitra. *Introduction to dynamic optimization theory*. Optimization and Chaos. Editors M. Majumdar, T. Mitra and K. Nishimura. Springer Verlag, (2000), 31–108.

[Ni] R.K. Niven. *Cost of s-fold Decisions in Exact Maxwell-Boltzmann, Bose-Einstein and Fermi-Dirac Statistics*. Physica A, **365**, Issue **1** (2006), 142–149.

[Os] A. Ostrowski. *On positive matrices*. Math. Annalen, **150** (1963), 276–284.

[Pe] M. Pettini. *Geometry and topology in Hamiltonian dynamics and statistical mechanics*. Springer Verlag, (2007).

[PP] W. Parry and M. Pollicott. *Zeta functions and the periodic orbit structure of hyperbolic dynamics*. Astérisque, **187-188** (1990).

[Ra] S. Rachev and L. Ruschendorf. *Mass transportation problems, Vol. I and II*. Springer Verlag, (1998).

[Roc] R.T. Rockafellar. *Extention of Fenchel's duality theorem for convex functions*. Duke Math. J., **33** (1966), 81–89.

[RRS] S. Risau-Gusman A.C. Ribeiro-Teixeira and D.A. Stariolo. *Topology and Phase Transitions: The Case of the Short Range Spherical Model*. Journ. of Statist. Physics, **124**(5) (2006), 1231–1253.

[Sch] H.H. Schaefer. *Banach Lattices and Positive Operators*. Springer Verlag, (1974).

[Sp] F. Spitzer. *A Variational characterization of finite Markov chains*. The Annals of Mathematical Statistics, **43**(1) (1972), 303–307.

[Ta] M. Takahashi. *Thermodynamics of one-dimensional solvable models*. Cambridge Press, (2005).

[Th] C. Thompson. *Infinite-Spin Ising Model in one dimension*. Journal of Mathematical Physics, **9**(2) (1968), 241–245.

[V] A. van Enter, S. Romano and V. Zagrebnov. *First-order transitions for some generalized XY models*. J. Phys. A., **39**(26) (2006), 439–445.

[W] W.F. Wrezinski and E. Abdalla. *A precise formulation of the third law of thermodynamics with applications to statistical physics and black holes*. Preprint USP (2007).

A.O. Lopes, **J. Mohr** and **R.R. Souza**
Instituto de Matemática
UFRGS
91509-900 Porto Alegre
BRAZIL

E-mail: arturoscar.lopes@gmail.com

Ph. Thieullen
Institut de Mathématiques
Université Bordeaux 1
F-33405 Talence
FRANCE

E-mail: Philippethieullen@math.u-bordeaux1.fr