Functions for relative maximization

E. Garibaldi and A. O. Lopes

July 12, 2007

Abstract

We introduce functions for relative maximization in a general context: the beta and alpha applications. After a systematic study of different kinds of regularities, we investigate how to approximate certain values of these functions using periodic orbits. We also show that the differential of an alpha application determines the asymptotic behavior of the optimal trajectories.

1. Introduction

Let (X, d) be a compact metric space. If $T : X \to X$ is a continuous function, consider $\mathcal{M}_T$ the set of T-invariant Borel probability measures. Remind that $\mathcal{M}_T$ is convex and weak* compact.

Given a continuous function $A : X \to \mathbb{R}$, we denote

$$\beta_A = \max_{\mu \in \mathcal{M}_T} \int A \, d\mu.$$

In ergodic optimization on compact spaces, the characterization of the invariant probability measures whose integral of A reaches the maximum value β_A is one of the main goals. We call any of these probabilities an A-maximizing probability. References, general definitions and problems consider in this theory can be found, for instance, in Jenkinson's notes (see [19]).

In the present work, we will look to the problem of the characterization of maximizing probabilities in a slight different formulation. Given A as above, we introduce also a continuous application $\varphi : X \to \mathbb{R}^n$ (which plays the role of a *constraint*) and we extend the concept of β_A to a real function defined in a convex subset of $\mathbb{R}^n$, the *rotation set.* We use here a terminology borrowed from the Aubry-Mather theory [6]. This concave application will be called the *beta function* $\beta_{A,\varphi}$ (associated to A and φ) and its Fenchel transform, the *alpha function.*

A motivation for analyzing such kind of problem is furnished by [27]. In that paper, for the Lagrangian $L(x, v) = \|v\|_x^2/2$ obtained from the Riemannian metric in a compact constant negative curvature surface of genus 2, the authors consider Mather measures for L subject to a certain homological condition. Via the Bowen-Series transformation $T : S^1 \to S^1$, it is shown that this situation can be translated into a relative maximization problem for the potential $A = -\log T'$.

A simple example of the theory studied here is the following. Consider $X = \{1, 2, 3\}^{\mathbb{N}}$ and $A(x_0, x_1, x_2, \ldots) = A(x_0, x_1)$ depending only on the two first coordinates. Suppose $A(2, 3) = A(3, 2)$ and $A(2, 3) > A(i, j)$ for the other possibilities. The maximizing shift-invariant probability for A without constraint is the one supported on the periodic orbit $(2, 3, 2, 3, \ldots)$. If we require $\displaystyle\int \varphi \, d\mu = 1$, where φ is the indicator function of the cylinder $\bar{1}$, then the maximization is given by the fixed point $(1, 1, \ldots)$.

In the setting presented here, we could change maximization for minimization and the analogous statements would be immediately verify. In the first section of the present paper, all the definitions will be carefully discussed. We will then derive some basic properties associated to these concepts.

In other topic, the behavior of alpha and beta applications when the constraint is modified will be analyzed. We will show, for instance, the Lipschitz character of the correspondence associating constraints to alpha functions values. Concerning the influence over a beta function, we will verify typically continuity of the respective correspondence.

We will also investigate the possibility of approximating alpha and beta functions values using probability measures supported on periodic orbits. Nevertheless, we will deal with this problem specifically for the symbolic dynamics case. Under the hypothesis of joint recurrence (to be discussed later), we will show the existence of periodic orbits carrying out the task.

Finally we will present a theorem that points out an interesting connection between the differential of an alpha application and the asymptotic behavior of certain trajectories (see [13] for a different setting). In the proof of this result, the sub-action concept will appear. This notion has been largely studied [2, 3, 4, 7, 8, 18, 25, 32, 33].

This paper is part of the first author's PhD thesis [10]. It can be seen as an analysis of the properties of the beta function $\beta_{A,\varphi}$ – a generalization of the maximal constant β_A – and of its Fenchel transform. On the other hand, once the graph of a beta application is part of the boundary of a rotation set contained in $\mathbb{R}^{n+1}$, this study also brings some information on such set. However, evidently it does not make it in a so explicit way as, for instance, Kwapisz [20, 21, 22] for certain rotation sets arising from two-torus maps homotopic to the identity, Bousch [3] and Jenkinson [17] when analyzing the set of the barycentres of invariant measures for circle maps, or Ziemian [35] in symbolic dynamics.

We would like to point out that one can just cite few relative ergodic optimization results for very special situations, mostly for Sturm probabilities (see, for instance, pages 104-105 in [5]).

The results presented here should be seen as the general abstract setting for ergodic optimization. We consider with great detail connections between *periodic probabilities* and the case of *rational rotation* (see Corollaries 11, 12 and Proposition 14) and still some properties obtained from the differentiability of the alpha function (see Theorem 17). We also analyze joint recurrence (see Proposition 9). In the future, we will analyze similar problems with more stringent hypothesis (maybe a concept similar to convexity), where the notion of non-crossing trajectories shall be present. We believe it will be possible to obtain stronger results in this case.

2. First Definitions

Let $\varphi : X \to \mathbb{R}^n$ be a continuous application with coordinate functions $\varphi_1, \ldots, \varphi_n$. We have then an induced map $\varphi_* : \mathcal{M}_T \to \mathbb{R}^n$ given by $\varphi_*(\mu) = \left(\int \varphi_1 \, d\mu, \ldots, \int \varphi_n \, d\mu \right)$. Clearly, φ_* is a continuous and affine function.

We call $\varphi_*(\mu)$ the rotation vector of the measure $\mu \in \mathcal{M}_T$. (When $n = 1$, we will prefer the expression rotation number of the measure.) Note that the image $\varphi_*(\mathcal{M}_T) \subset \mathbb{R}^n$ is a convex compact set, inheriting it of $\mathcal{M}_T$. We call $\varphi_*(\mathcal{M}_T)$ a rotation set. For $h \in \varphi_*(\mathcal{M}_T)$, the fiber $\varphi_*^{-1}(h)$ is called the rotation class of h. Also $\varphi_*^{-1}(h) \subset \mathcal{M}_T$ is a convex compact set.

Proposition 1. *For the induced map $\varphi_* : \mathcal{M}_T \to \varphi_*(\mathcal{M}_T)$, we verify:*

(i) if the fiber $\varphi_*^{-1}(h)$ is a singleton set containing an ergodic measure, then h is an extremal point of $\varphi_*(\mathcal{M}_T)$;

(ii) if h is an extremal point of $\varphi_*(\mathcal{M}_T)$, then the extremal points of $\varphi_*^{-1}(h)$ are ergodic measures.

Actually, this proposition is just a general version of results presented in Jenkinson's PhD thesis (see lemmas 3.2 and 3.3 of [16]).

For $A \in C^0(X)$, we define the beta function $\beta_{A,\varphi} : \varphi_*(\mathcal{M}_T) \to \mathbb{R}$ by

$$\beta_{A,\varphi}(h) = \sup \left\{ \int A \, d\mu : \mu \in \varphi_*^{-1}(h) \right\}.$$

In this context, we call the function φ a constraint and the function A a potential. Important objects will be the probability measures belonging to the rotation class of h that, on such set, maximize the integral of the potential A. In other words, consider the set

$$m_{A,\varphi}(h) = \left\{ \mu \in \varphi_*^{-1}(h) : \int A \, d\mu = \beta_{A,\varphi}(h) \right\}.$$

If $\mu \in m_{A,\varphi}(h)$, we say that μ is an (A,h)-maximizing probability.

Since the rotation class of h is a compact set, it is easy to prove that $m_{A,\varphi}(h)$ is a nonempty compact set. It follows that $\beta_{A,\varphi} : \varphi_*(\mathcal{M}_T) \to \mathbb{R}$ is a concave application. Moreover, since the correspondence $\mu \mapsto \int A \, d\mu$ is continuous, we have that $\beta_{A,\varphi}$ is continuous on the whole rotation set.

The properties of the beta application legitimate the definiton of a concave function $\alpha_{A,\varphi} : \mathbb{R}^n \to \mathbb{R}$ via Fenchel transform

$$\alpha_{A,\varphi}(c) = \min_{h \in \varphi_*(\mathcal{M}_T)} [\langle c, h\rangle - \beta_{A,\varphi}(h)].$$

Such application is called the alpha function (associated to A and φ).

It is interesting to examine the behaviors of the beta and alpha applications when the parameters that define them are changed. For instance, we can question how a potential modification affects a beta function. Given $h \in \varphi_*(\mathcal{M}_T)$, in a natural way we obtain a function $\beta_{\cdot,\varphi}(h) : C^0(X) \to \mathbb{R}$ that, to each potential A, simply associates the value $\beta_{A,\varphi}(h)$. It is not difficult to verify that this application is Lipschitz, with $\mathrm{Lip}(\beta_{\cdot,\varphi}(h)) \le 1$.

A first consequence of this fact is the Lipschitz regularity of an alpha function, with $\mathrm{Lip}(\alpha_{A,\varphi}) \le \|\varphi\|_0$. Indeed, since

$$\alpha_{A,\varphi}(c) = -\max_{h \in \varphi_*(\mathcal{M}_T)} \beta_{A - \langle c, h\rangle, \varphi}(h),$$

if we take $h' \in \varphi_*(\mathcal{M}_T)$ such that $\alpha_{A,\varphi}(c') = -\beta_{A-\langle c', h'\rangle,\varphi}(h')$, we have

$$\begin{aligned} \alpha_{A,\varphi}(c) - \alpha_{A,\varphi}(c') &\le \beta_{A-\langle c',h'\rangle,\varphi}(h') - \beta_{A-\langle c,h'\rangle,\varphi}(h') \\ &\le |\langle c - c', h'\rangle| \ \le \ \|\varphi\|_0 \|c - c'\|. \end{aligned}$$

(As usual, $\|\cdot\|$ denotes the Euclidean norm on $\mathbb{R}^n$.)

A second immediate consequence is the following version of the Fenchel inequality

$$\begin{aligned} \beta_{A,\varphi}(h) + \alpha_{B,\varphi}(c) &\le \beta_{A,\varphi}(h) + \langle c, h\rangle - \beta_{B,\varphi}(h) \\ &\le \langle c, h\rangle + \|A - B\|_0. \end{aligned}$$

Using this inequality, we see that it is also Lipschitz the application that makes to correspond $\alpha_{A,\varphi}(c)$ to each potential A, namely, the function $\alpha_{\cdot,\varphi}(c) : C^0(X) \to \mathbb{R}$, besides we have $\mathrm{Lip}(\alpha_{\cdot,\varphi}(c)) \le 1$.

Some properties of the applications $\beta_{\cdot,\varphi}(h), \alpha_{\cdot,\varphi}(c) : C^0(X) \to \mathbb{R}$ are summarized in the proposition below. The simple proof will be omitted.

Proposition 2. *If $A, B \in C^0(X)$, $a \in \mathbb{R}$ and $t, t' \in [0,1]$ with $t + t' = 1$, then the functions $\beta_{\cdot,\varphi}(h), \alpha_{\cdot,\varphi}(c) : C^0(X) \to \mathbb{R}$ verify*

(i) $\beta_{aA,\varphi}(h) = |a| \beta_{sgn(a)A,\varphi}(h)$;

(ii) $\beta_{A+B\circ T-B+a,\varphi}(h) = \beta_{A,\varphi}(h) + a$, $\alpha_{A+B\circ T-B+a,\varphi}(c) = \alpha_{A,\varphi}(c) - a$*;*

(iii) $\beta_{A+B,\varphi}(h) \leq \beta_{A,\varphi}(h) + \beta_{B,\varphi}(h)$*;*

(iv) $\beta_{tA+t'B,\varphi}(h) \leq t\beta_{A,\varphi}(h)+t'\beta_{B,\varphi}(h)$, $\alpha_{tA+t'B,\varphi}(c) \geq t\alpha_{A,\varphi}(c)+t'\alpha_{B,\varphi}(c)$*;*

(v) $A \leq B$ *implies* $\beta_{A,\varphi}(h) \leq \beta_{B,\varphi}(h)$, $\alpha_{A,\varphi}(c) \geq \alpha_{B,\varphi}(c)$*.*

Note that complementing expressions for the items (i) and (iii) would be

$$\alpha_{aA,\varphi}(c) = |a|\alpha_{\mathrm{sgn}(a)A,\varphi}(c/|a|) \text{ (for } a \neq 0) \text{ and}$$

$$\alpha_{A+B,\varphi}(c + c') \geq \alpha_{A,\varphi}(c) + \alpha_{B,\varphi}(c'),$$

which are not properties of the application $\alpha_{\cdot,\varphi}(c)$.

Besides redefining a beta function, the modification of the potential also redescribes the set of maximizing probabilities. Though, a particularity prevails typically.

Proposition 3. *Assume* $h \in \varphi_*(\mathcal{M}_T)$*. There is a residual subset* $\mathcal{G} = \mathcal{G}(h) \subset C^0(X)$ *such that, for each potential* $A \in \mathcal{G}$, $m_{A,\varphi}(h)$ *contains an unique probability measure.*

This result can be seen as a particular version of a more general formulation obtained in proposition 10 of [7]. The proof there is also valid for any compact metric space X.

3. The role of the constraint

Our objective now will be to discuss how the changing of the constraint affects the beta and alpha functions.

Of course, one can present several simple properties. For instance, suppose $A, B \in C^0(X)$ and $\varphi, \psi \in C^0(X, \mathbb{R}^n)$. Take yet $a \in \mathbb{R}^*$, $b \in \mathbb{R}^n$ and $t, t' \in [0, 1]$ with $t + t' = 1$. Then, we have

(i) $\beta_{A,a\varphi}(h) = \beta_{A,\varphi}(h/a)$, $\alpha_{A,a\varphi}(c) = \alpha_{A,\varphi}(ac)$;
(ii) $\beta_{A,\varphi+\psi\circ T-\psi+b}(h) = \beta_{A,\varphi}(h - b)$, $\alpha_{A,\varphi+\psi\circ T-\psi+b}(c) = \alpha_{A,\varphi}(c) + \langle c, b\rangle$;
(iii) $\alpha_{A+B,\varphi+\psi}(c) \geq \alpha_{A,\varphi}(c) + \alpha_{B,\psi}(c)$;
(iv) $\alpha_{tA+t'B,t\varphi+t'\psi}(c) \geq t\alpha_{A,\varphi}(c) + t'\alpha_{B,\psi}(c)$;
(v) $m_{A,\varphi}(h) \cap m_{A,\psi}(h') \neq \emptyset \Rightarrow t\beta_{A,\varphi}(h) + t'\beta_{A,\psi}(h') \leq \beta_{A,\varphi+\psi}(h + h')$.

The proof of these items is left to the reader.

In order to make interesting the investigation of the relationship between constraints and the beta and alpha functions, notice that there is an initial difficulty: the constraint also determines the domain of a beta application. Therefore, we first need to establish which effect the change of this parameter produces on the rotation set.

For a complete metric space Y, we will denote by $\mathcal{K}(Y)$ the set of all nonempty compact subsets of Y. With the Hausdorff metric, $\mathcal{K}(Y)$ becomes a complete metric space. For the proof, we need to note the linear operator $*: C^0(X, \mathbb{R}^n) \to C^0(\mathcal{M}_T, \mathbb{R}^n)$ is bounded, with norm smaller or equal to 1.

Proposition 4. *If the application* $\Gamma_T : C^0(X, \mathbb{R}^n) \to \mathcal{K}(\mathbb{R}^n)$ *is given by* $\Gamma_T(\varphi) = \varphi_*(\mathcal{M}_T)$*, then* Γ_T *is Lipschitz, with* $Lip(\Gamma_T) \leq 1$*.*

Proof. Observe that, for any $\varphi, \psi \in C^0(X, \mathbb{R}^n)$ and $\mu \in \mathcal{M}_T$, we have

$$\begin{aligned} d(\varphi_*(\mu), \psi_*(\mathcal{M}_T)) &= \inf_{\nu \in \mathcal{M}_T} \|\varphi_*(\mu) - \psi_*(\nu)\| \\ &\leq \|\varphi_*(\mu) - \psi_*(\mu)\| \\ &\leq \|(\varphi - \psi)_*\|_0 \\ &\leq \|\varphi - \psi\|_0. \end{aligned}$$

However, by the construction of the Hausdorf metric, this argument is enough to establish the proposition. □

Suppose we associate, to each map $T \in C^0(X, X)$, some application $\varphi_T \in C^0(X, \mathbb{R}^n)$. This happens, for instance, in the case of rotation sets arising from n-dimensional torus homeomorphisms homotopic to the identity or when one wants to analyze the spectrum of the Lyapunov exponents of a differential application. Motivated by the proposition above, we could question which regularity the map $T \mapsto (\varphi_T)_*(\mathcal{M}_T)$ presents.

Proposition 5. *Consider a subset* $\mathcal{U} \subset C^0(X, X)$ *with the induced topology. Let* $T \in \mathcal{U} \mapsto \varphi_T \in C^0(X, \mathbb{R}^n)$ *be a continuous map. Then the application* $\Gamma_{\mathcal{U}} : \mathcal{U} \to \mathcal{K}(\mathbb{R}^n)$ *defined as* $\Gamma_{\mathcal{U}}(T) = (\varphi_T)_*(\mathcal{M}_T)$ *is upper semi-continuous.*

Proof. If the upper semi-continuity of $\Gamma_{\mathcal{U}}$ was not verified, this would mean the existence of a map $T \in \mathcal{U}$ and some $\epsilon > 0$ for which we could determine a sequence $\{T_j\} \subset \mathcal{U}$ convergent to T and a sequence of Borel probability measures $\{\mu_j\}$ satisfying both $\mu_j \in \mathcal{M}_{T_j}$ and $d((\varphi_{T_j})_*(\mu_j), \Gamma_{\mathcal{U}}(T)) \geq \epsilon$. Therefore, we would have a subsequence $\{\mu_{j_k}\}$ convergent to some Borel probability measure μ. Now, given any function $f \in C^0(X)$, and then passing to the limit in

$$\int f \circ T_{j_k} \, d\mu_{j_k} = \int f \, d\mu_{j_k},$$

we would reach a contradiction: $\mu \in \mathcal{M}_T$ and $d((\varphi_T)_*(\mu), (\varphi_T)_*(\mathcal{M}_T)) \geq \epsilon$. □

One could still remark that the map $T \mapsto \mathcal{M}_T$ is upper semi-continuous. For rotation sets arising from n-dimensional torus continuous maps homotopic to the identity, a result in the spirit of the previous proposition is already known (see [14, 29]).

In the analysis of the influences of the constraint on the beta and alpha functions, the next lemma will be useful.

Lemma 6. *Given a constraint $\varphi \in C^0(X, \mathbb{R}^n)$, take $h \in \varphi_*(\mathcal{M}_T)$. Consider a sequence of constraints $\{\varphi_j\}$ converging to φ. It follows*

(i) $\lim_{j\to\infty} d(h, (\varphi_j)_*(m_{A,\varphi}(h))) = 0$;

(ii) $\lim_{j\to\infty} d(h, \varphi_*(m_{A,\varphi_j}(h))) = 0$ *when* $h \in (\varphi_j)_*(\mathcal{M}_T)$;

(iii) $\lim_{j\to\infty} d(h, \varphi_*(m_{A,\varphi_j}(h_j))) = 0$ *when* $h_j \in (\varphi_j)_*(\varphi_*^{-1}(h))$.

Proof. By a similar reasoning to the one used in proposition 4, we obtain $d(h', \psi_*(m_{A,\psi'}(h'))) \leq \|\psi - \psi'\|_0$, which clearly gives the items (i) and (ii).

Besides, we have $d(h, \varphi_*(m_{A,\varphi_j}(h_j))) \leq \|h - h_j\| + d(h_j, \varphi_*(m_{A,\varphi_j}(h_j)))$. Consequently, for the second term, just as in the previous paragraph, we have $d(h_j, \varphi_*(m_{A,\varphi_j}(h_j))) \leq \|\varphi_j - \varphi\|_0$. And for the first one, choosing $\mu \in \varphi_*^{-1}(h) \cap (\varphi_j)_*^{-1}(h_j)$, we get $\|h - h_j\| = \|\varphi_*(\mu) - (\varphi_j)_*(\mu)\| \leq \|\varphi - \varphi_j\|_0$, which concludes the proof of item (iii). □

Using the lemma above, we can show a kind of *holographic continuity* for the beta application as function of the constraint.

Proposition 7. *About the behavior of the beta and alpha functions when the constraint is modified, we have the following results.*

(I) Given a constraint $\varphi \in C^0(X, \mathbb{R}^n)$, take $h \in \varphi_*(\mathcal{M}_T)$. Let $\{\varphi_j\}$ be a sequence of constraints convergent to φ. Assume that $\{h_j\} \subset \mathbb{R}^n$ is a sequence satisfying $h_j \in (\varphi_j)_*(m_{A,\varphi}(h))$. Then $\lim \beta_{A,\varphi_j}(h_j) = \beta_{A,\varphi}(h)$.

(II) For $c \in \mathbb{R}^n$, the map $\varphi \mapsto \alpha_{A,\varphi}(c)$ is Lipschitz, with Lip$(\alpha_{A,\cdot}(c)) \leq \|c\|$.

Proof. (I) Initially, note that, from the choice of h_j, it happens $\beta_{A,\varphi}(h) \leq \beta_{A,\varphi_j}(h_j)$. Define a sequence $\{\eta_j\} \subset \mathbb{R}^n$ such that, for each integer j, the vector $\eta_j \in \varphi_*(m_{A,\varphi_j}(h_j))$ satisfies $\|h - \eta_j\| = d(h, \varphi_*(m_{A,\varphi_j}(h_j)))$. Therefore, we obtain $\beta_{A,\varphi}(h) \leq \beta_{A,\varphi_j}(h_j) \leq \beta_{A,\varphi}(\eta_j)$. Besides, by the item (iii) of the lemma above, we use $\lim \eta_j = h$ to legitimate $\lim \beta_{A,\varphi}(\eta_j) = \beta_{A,\varphi}(h)$.

(II) Given $\epsilon > 0$, consider $h \in \varphi_*(\mathcal{M}_T)$ accomplishing $\langle c, h\rangle - \beta_{A,\varphi}(h) < \alpha_{A,\varphi}(c) + \epsilon/2$. Afterwards, take a probability measure $\mu \in \varphi_*^{-1}(h)$ satisfying $\int A\, d\mu > \beta_{A,\varphi}(h) - \epsilon/2$. Therefore, $\langle c, \varphi_*(\mu)\rangle - \int A\, d\mu < \alpha_{A,\varphi}(c) + \epsilon$. Besides, if $\psi \in C^0(X, \mathbb{R}^n)$ is a constraint, the Fenchel inequality gives $\alpha_{A,\psi}(c) + \int A\, d\mu \leq \alpha_{A,\psi}(c) + \beta_{A,\psi}(\psi_*(\mu)) \leq \langle c, \psi_*(\mu)\rangle$. Thus, we verify $\alpha_{A,\psi}(c) - \alpha_{A,\varphi}(c) < \langle c, (\psi - \varphi)_*(\mu)\rangle + \epsilon \leq \|c\|\|\psi - \varphi\|_0 + \epsilon$. And the result follows from the symmetrical role carried out by φ and ψ and from the arbitrariness of ϵ. □

The need of the sequence $\{h_j\}$ in the hypothesis of proposition 7.I is a little bit disappointing. Because of the item (i) of lemma 6 we can choose it converging to h, but we should ask: when $\lim \beta_{A,\varphi_j}(h) = \beta_{A,\varphi}(h)$? The first aspect to be noted is the requirement $h \in (\varphi_j)_*(\mathcal{M}_T)$. However, if $h \in \mathrm{int}(\varphi_*(\mathcal{M}_T))$, the proposition 4 assures that, for a constraint ψ sufficiently close to φ, we have $h \in \mathrm{int}(\psi_*(\mathcal{M}_T))$.

Proposition 8. *Let $\varphi \in C^0(X, \mathbb{R}^n)$ be a constraint. Take $h \in int(\varphi_*(\mathcal{M}_T))$. If $\{\varphi_j\}$ is a sequence convergent to φ, then $\lim \beta_{A,\varphi_j}(h) = \beta_{A,\varphi}(h)$.*

Proof. Without loss of generality, we can suppose that $h \in \mathrm{int}((\varphi_j)_*(\mathcal{M}_T))$. However, we will need a stronger version of this hypothesis. Fortunately, the proposition 4 also allows to assume that $D_\epsilon[h] \subset \mathrm{int}((\varphi_j)_*(\mathcal{M}_T))$, where $D_\epsilon[h]$ is a closed ball of center h and radius $\epsilon > 0$ contained in $\mathrm{int}(\varphi_*(\mathcal{M}_T))$.

Define a sequence of probability measures $\{\mu_j\} \subset \mathcal{m}_{A,\varphi}(h)$ such that, for each integer j, we have $\|h - (\varphi_j)_*(\mu_j)\| = d(h, (\varphi_j)_*(\mathcal{m}_{A,\varphi}(h)))$. Putting $h_j = (\varphi_j)_*(\mu_j)$, we set

$$\epsilon_j = \frac{\|h - h_j\|}{\|h - h_j\| + \dfrac{\epsilon}{3}}.$$

Write, then, $h'_j = h_j + \epsilon_j^{-1}(h - h_j)$. Note that, in reason of the item (i) of the lemma 6, for an integer j sufficiently large, it happens $\|h - h_j\| \le \epsilon/3$. Hence, for such indexes, we verify $h'_j \in D_\epsilon[h] \subset \mathrm{int}((\varphi_j)_*(\mathcal{M}_T))$, that is, we obtain $h'_j = (\varphi_j)_*(\mu'_j)$ for some T-invariant probability measure μ'_j.

Put, for j sufficiently large, $\mu''_j = \epsilon_j \mu'_j + (1 - \epsilon_j)\mu_j$. Note that $(\varphi_j)_*(\mu''_j) = \epsilon_j h'_j + (1 - \epsilon_j)h_j = h$. Therefore, if the vector $\eta_j \in \varphi_*(\mathcal{m}_{A,\varphi_j}(h))$ accomplishes $\|h - \eta_j\| = d(h, \varphi_*(\mathcal{m}_{A,\varphi_j}(h)))$, we have

$$\epsilon_j \int A \, d\mu'_j + (1 - \epsilon_j)\beta_{A,\varphi}(h) = \int A \, d\mu''_j \le \beta_{A,\varphi_j}(h) \le \beta_{A,\varphi}(\eta_j).$$

And the result follows directly of the items (i) and (ii) of the lemma 6. □

The proposition above could have a more direct proof, but maybe less instructive. Actually, it would be enough to apply the conclusion of the proposition 4 to the functions $\Phi = (\varphi, A)$ and $\Phi_j = (\varphi_j, A)$. This argument will be explored ahead in the text[1].

4. Approximation by Periodic Orbits

Although, in this section, we will restrict the class of dynamical systems to be examined, limiting us to study the approximation problem for periodic orbits in the context of the symbolic dynamics, we draw a general itinerary in certain aspects. This itinerary describes how, when the purpose is to estimate certain values of a beta application or of an alpha function, we can

[1]See, for instance, the proof of the proposition 14.

find probability measures supported on periodic orbits accomplishing such task.

Some comments on definitions and notations will be useful. Note that, in any probability space $(Y, \mathcal{B}, \nu)$, given an integrable application $f : Y \to \mathbb{R}^n$, we still have the natural concept of rotation vector of the measure ν. The integrability, in fact, plays the main role when we write

$$f_*(\nu) = \left(\int f_1 \, d\nu, \ldots, \int f_n \, d\nu \right).$$

Given an ergodic function $F : Y \to Y$, we set $b(f)$ to indicate the set of the elements of Y that, for the application $f \in L^1(Y, \mathcal{B}, \nu)$, satisfy the Birkhoff's ergodic theorem. For the characteristic function of a measurable set D, however, we will prefer to denote it by $b(D)$. Besides, just looking at the mensurability of F, we put $S_k f = \sum_{j=0}^{k-1} f \circ F^j$ for $k > 0$ and $S_0 f = 0$.

We consider $\Xi(D)$ the set of the elements z of D such that, for any $\epsilon > 0$, it exists a positive integer L accomplishing both $F^L(z) \in D$ and $\|S_L f(z) - L f_*(\nu)\| < \epsilon$. Thus, we say that the integrable function f is joint recurrent (with respect to the probability measure ν) if, for each $D \in \mathcal{B}$, it happens $\nu(\Xi(D)) = \nu(D)$. (When $n = 1$, we will simply say that f is recurrent.) If we want to identify functions verifying such property, the following proposition describes a sufficient condition.

Proposition 9. *Let $(Y, \mathcal{B}, \nu)$ be a probability space. Consider an ergodic transformation $F : Y \to Y$ and an integrable function $f : Y \to \mathbb{R}^n$ satisfying*

$$\lim_{k \to \infty} \frac{1}{k^{1/n}} \|S_k f(y) - k f_*(\nu)\| = 0$$

for ν-almost every point $y \in Y$. Then f is joint recurrent.

Note that, if $n = 1$, for every integrable application, we have immediately the required limit by the Birkhoff's ergodic theorem. In simple terms, the proposition 9 shows that any integrable function $f : Y \to \mathbb{R}$ is recurrent. This result when $n = 1$ was used by Mañé in one of his works on minimizing measures of Lagrangian systems (see lemma 2.2 of [28]). Nevertheless, two decades before, a theorem containing the one-dimensional version of proposition 9 had been obtained by Atkinson in [1]. The proof that we will present for the general case $n \geq 1$ is a generalization of a proof for the particular situation when $n = 1$, more specifically, of the proof given for the lemma 3-6.4 of [6].

Proof. Without loss of generality, we can take $f_*(\nu) = 0$. Suppose $D \in \mathcal{B}$ with $\nu(D) > 0$. Assuming $\epsilon > 0$, let $\Xi_\epsilon(D)$ denote the set of points $z \in D$ for which there is a positive integer L such that $F^L(z) \in D$ and $\|S_L f(z)\| < \epsilon$. Since $\Xi(D) = \bigcap \Xi_{1/j}(D)$, it is enough to show that $\nu(\Xi_\epsilon(D)) = \nu(D)$.

Take $y \in D \cap b(f) \cap b(D) \cap b(\Xi_\epsilon(D))$ such that $\lim k^{-1/n} \|S_k f(y)\| = 0$. Consequently, let $L_1 < L_2 < \ldots L_k < \ldots$ be the positive integers such that $F^{L_k}(y) \in D$. Defining $a_k = S_{L_k} f(y)$, consider yet

$$R = \{k : \forall \;\; m > k, \;\; \|a_m - a_k\| \geq \epsilon\} \;\; \text{and} \;\; R_k = R \cap \{1, \ldots k\}.$$

Note that, for each $l \in \{1, \ldots k\} - R_k$, there exists $m > l$ such that $\|S_{L_m - L_l} f(F^{L_l}(y))\| = \|a_m - a_l\| < \epsilon$. In other words, $l \in \{1, \ldots k\} - R_k$ implicates $F^{L_l}(y) \in \Xi_\epsilon(D)$. Therefore, we verify

$$\begin{aligned} 1 + \#R_k &\geq 1 + \#\{1 \leq l < k : F^{L_l}(y) \notin \Xi_\epsilon(D)\} \\ &\geq \#\{0 \leq j < L_k : F^j(y) \in D - \Xi_\epsilon(D)\}. \end{aligned}$$

Hence, since

$$\nu(D - \Xi_\epsilon(D)) = \lim_{k \to \infty} \frac{1}{L_k} \sum_{j=0}^{L_k - 1} \chi_{D - \Xi_\epsilon(D)}(F^j(y)),$$

the proposition will be proved when we obtain a subsequence of $\{(\#R_k)/L_k\}$ converging to zero.

If R is a finite set, there is nothing to argue. Suppose, otherwise, R is an infinite set. Then, by construction, $\{a_k : k \in R\}$ is unbounded. In such case, choose an infinite sequence $S \subset R$ such that, for every $k \in S$,

$$\|a_k\| = \max_{l \in R_k} \|a_l\|.$$

When denoting by $D_\rho(\gamma)$ the open ball of center $\gamma \in \mathbb{R}^n$ and radius $\rho > 0$, we observe that, given $k \in S$, $D_{\epsilon/2}(a_l) \subset D_{\|a_k\| + \epsilon/2}(0)$ for each $l \in R_k$. Besides, for the definition of R, these balls $D_{\epsilon/2}(a_l)$, $l \in R_k$, are all disjoint. Consequently,

$$\#R_k \leq \frac{\left(\|a_k\| + \frac{\epsilon}{2}\right)^n}{\left(\frac{\epsilon}{2}\right)^n} = \sum_{j=0}^{n} \binom{n}{j} \|S_{L_k} f(y)\|^j \left(\frac{2}{\epsilon}\right)^j.$$

Reminding that $\lim k^{-1/n} \|S_k f(y)\| = 0$, to verify

$$\lim_{k \in S} \frac{\#R_k}{L_k} = 0$$

is an easy task. $\square$

Let f be a joint recurrent function with respect to a probability measure ν. If D is a measurable set of positive measure, write $\Xi^j(D) = \Xi(\Xi^{j-1}(D))$. Then, observe that $\nu(\bigcap \Xi^j(D)) = \nu(D) > 0$. In particular, if we have $E \in \mathcal{B}$ with $\nu(E) = 1$, then $\bigcap \Xi^j(D) \cap E \neq \emptyset$. This simple fact will play a crucial role in the proof of the next result. We will need, however, more structure to obtain the statement of next theorem. Thus, our study will be driven towards the symbolic dynamics setting.

Let us begin, nevertheless, reminding concepts which are not restricted to this dynamics. Given a periodic point $x \in X$ of period M, naturally we have a T-invariant probability measure defined by

$$\mu = \frac{1}{\#\mathrm{orb}(x)} \sum_{y\, \in \mathrm{orb}(x)} \delta_y = \frac{1}{M} \sum_{j=0}^{M-1} \delta_{T^j(x)}.$$

A way to refer to a such measure μ will be calling it a periodic probability measure. When taking $x, y \in X$ and any positive interger k, other item to be remembered is the synthesis between the metric and the dynamics indicated by

$$d_k(x, y) = \max_{0 \leq j < k} d(T^j(x), T^j(y)).$$

A special collection of potentials will be the focus of our work: the Walters potentials[2]. A function $f \in C^0(X)$ is a Walters function if it admits a Walters module, that is, if there exists a function $H : \mathbb{R}^+ \to \mathbb{R}^+ \cup \{+\infty\}$ increasing, null and continuous in zero, such that

$$\forall\, s \in \mathbb{R}^+,\ \forall\, k > 0,\ \forall\, x, y \in X,\ d_k(x, y) \leq s \Rightarrow |S_k f(x) - S_k f(y)| \leq H(s).$$

For hyperbolic dynamical systems, the set of the Walters functions includes (see the definition-proposition 2 of [4]) all the functions of summable variation, in particular the Hölder functions are then examples of Walters functions.

Finally, let $\sigma : \Sigma \to \Sigma$ be a subshift of finite type. Given a constant $\lambda \in (0, 1)$, we consider in Σ the metric $d(\mathbf{x}, \mathbf{y}) = \lambda^k$, where $\mathbf{x}, \mathbf{y} \in \Sigma$, $\mathbf{x} = (x_0, x_1, \ldots)$, $\mathbf{y} = (y_0, y_1, \ldots)$ and $k = \min\{j : x_j \neq y_j\}$. We will say that a continuous function $g : \Sigma \to \mathbb{R}^n$ is locally constant if there exists an integer $j \geq 0$ such that $g(\mathbf{x}) = g(\mathbf{y})$ whenever $x_0 = y_0, \ldots, x_j = y_j$. We could also say that this application depends on $j + 1$ coordinates.

Theorem 10. *Suppose $\varphi \in C^0(\Sigma, \mathbb{Q}^n)$ is a locally constant constraint and A is a Walters potential. Let φ be a joint recurrent application with respect to an ergodic probability measure $\nu \in \varphi_*^{-1}(r)$, where $r \in \varphi_*(\mathcal{M}_\sigma) \cap \mathbb{Q}^n$. Then, for each $\epsilon > 0$, there exists a periodic probability measure $\mu \in \varphi_*^{-1}(r)$ such that* $\left| \int A \, d\nu - \int A \, d\mu \right| < \epsilon$.

[2]The Walters condition was introduced in the article [34], where its convenience for the thermodynamic formalism was explored.

Proof. Take $\mathbf{x} \in \operatorname{supp}(\nu)$. For any integer $l \geq 0$, we denote the open ball centered in $\mathbf{x}$ of radius λ^l by $D_l = \{\mathbf{y} \in \Sigma : y_j = x_j \ \ \forall\, 0 \leq j < l\}$. Let H be a Walters module for the potential A. Given $\epsilon > 0$, we chose l sufficiently large (taking it larger than the number of coordinates on which depends φ) in such way that $H(\lambda^l) < \epsilon/2$.

As the constraint $\varphi : \Sigma \to \mathbb{Q}^n$ is locally constant, its image is reduced to a finite set of vectors with rational coordinates. Suppose these numbers are written in irreducible fractions and let $Q > 0$ be the product of their denominators. In the same way, let us consider $q > 0$ the product of the denominators of the coordinates of r.

The joint recurrence of φ assures there is a point $\mathbf{y} \in \bigcap \Xi^j(D_l) \cap \mathfrak{b}(A)$. Then, we obtain a positive integer M_0 such that, for $M \geq M_0$, we have

$$\left| \frac{1}{M} S_M A(\mathbf{y}) - \int A \, d\nu \right| < \frac{\epsilon}{2}.$$

Besides, since in particular $\mathbf{y} \in \Xi^{M_0}(D_l)$, a simple inductive argument gives positive integers $L_1, \ldots, L_{M_0}$ satisfying both $\sigma^{L_1+\ldots+L_k}(\mathbf{y}) \in \Xi^{M_0-k}(D_l)$ and

$$\|S_{L_1+\ldots+L_k}\varphi(\mathbf{y}) - (L_1 + \ldots + L_k)\varphi_*(\nu)\| < \frac{1}{qQ} \sum_{j=1}^{k} \frac{1}{2^{l+j}}$$

for every $k \in \{1, \ldots, M_0\}$.

Put $M = L_1 + \ldots + L_{M_0} \geq M_0$. Take, then, the periodic point $\mathbf{z} \in \Sigma$ given by the repetition of the word $(y_0, \ldots, y_{M-1})$. Finally, let μ be the σ-invariant probability measure by $\mathbf{z}$ defined. We only need to verify that such periodic probability measure accomplishes what is required.

Due to the fact we have taken l larger than the number of coordinates on which depends the constraint φ, we have $\varphi(\sigma^j(\mathbf{y})) = \varphi(\sigma^j(\mathbf{z}))$ when $j \in \{0, \ldots, M-1\}$. Therefore,

$$M\,\|\varphi_*(\mu) - r\| = \|S_M\varphi(\mathbf{y}) - M\varphi_*(\nu)\| < \frac{1}{qQ} \sum_{j=1}^{M_0} \frac{1}{2^{l+j}} < \frac{1}{qQ} \cdot \frac{1}{2^l}.$$

Once $QM\varphi_*(\mu) = QS_M\varphi(\mathbf{z}) \in \mathbb{Z}^n$, the inequality above assures $\varphi_*(\mu) = r$. Besides, observe that $d_M(\mathbf{y}, \mathbf{z}) \leq \lambda^l$ implies

$$\left| \int A \, d\mu - \frac{1}{M} S_M A(\mathbf{y}) \right| = \frac{1}{M} \left| S_M A(\mathbf{z}) - S_M A(\mathbf{y}) \right| \leq \frac{1}{M} H(\lambda^l) < \frac{\epsilon}{2}.$$

This ends the proof. $\square$

There are two points of view to understand the conclusion of the theorem above. The first is suggested by the well known fact according to which a circle homeomorphism of rational rotation number has a periodic point, whose period is equal to the denominator of the rational number. Such point of view follows the same spirit, for instance, of Franks theorem for certain rotation sets arising from two-torus homeomorphisms homotopic to the identity (see [9]). In the context of the symbolic dynamics, a result of this kind was obtained by Ziemian (see theorem 4.2 of [35]). The difference between the result of Ziemian and the one obtained here is the transitivity hypothesis. We give up this condition, but we introduce the hypothesis of joint recurrence. Thus, we have the immediate corollary.

Corollary 11. *Suppose $\varphi \in C^0(\Sigma, \mathbb{Q}^n)$ is a locally constant function. Given $r \in \varphi_*(\mathcal{M}_\sigma) \cap \mathbb{Q}^n$, if there is in the fiber $\varphi_*^{-1}(r)$ an ergodic probability measure with respect to which φ is joint recurrent, then in this fiber also exists a periodic probability measure.*

The second consequence of the theorem 9 is in the possibility of supplying a special description to a beta function. We will postpone the statement of the second corollary so that we can stop shortly at this point.

In general, for an alpha function, we can indicate the characterizations:

$$\begin{aligned}
\alpha_{A,\varphi}(c) &= \min_{\mu \in \mathcal{M}_T} \int (\langle c, \varphi \rangle - A)\, d\mu \\
&= \sup_{f \in C^0(X)} \min_{x \in X} (\langle c, \varphi \rangle - A + f - f \circ T)(x) \\
&= \inf_{x \in \operatorname{Reg}(\langle c, \varphi \rangle - A, T)} \lim_{k \to \infty} \frac{1}{k} S_k(\langle c, \varphi \rangle - A)(x) \\
&= \inf_{x \in X} \liminf_{k \to \infty} \frac{1}{k} S_k(\langle c, \varphi \rangle - A)(x),
\end{aligned}$$

where $\operatorname{Reg}(f, T)$ simply denotes the set of the points $x \in X$ for which is assured the existence of the limit of $k^{-1} S_k f(x)$ when k tends to infinite. The first of the equalities above, the reader will notice, comes directly from the definition of the alpha function. The second expression is the dual version of the previous one (see, for instance, [8, 31]). Starting from the first, the last two identities can be assured via Birkhoff's ergodic theorem. As a reference, it is possible to obtain these identities adapting lemmas contained in the work of Hunt and Yuan (see the lemmas 2.3 and 2.4 of [15]).

With respect to the representation of a beta function, we always verify the dual formula

$$\beta_{A,\varphi}(h) = \inf_{(f,c) \in C^0(X) \times \mathbb{R}^n} \max_{x \in X} (A + f - f \circ T - \langle c, \varphi - h \rangle)(x).$$

As a consequence of Fenchel-Rockafellar duality theorem, Radu established such equality in [31]. Starting from this equality and using the other identities above, we get the characterizations:

$$\begin{aligned}\beta_{A,\varphi}(h) &= \inf_{c\in\mathbb{R}^n} \beta_{A-\langle c,\varphi-h\rangle} \\ &= -\sup_{c\in\mathbb{R}^n} \alpha_{A,\varphi-h}(c) \\ &= \inf_{c\in\mathbb{R}^n} \sup_{x\in\mathrm{Reg}(A-\langle c,\varphi-h\rangle,T)} \lim_{k\to\infty} \frac{1}{k} S_k(A-\langle c,\varphi-h\rangle)(x) \\ &= \inf_{c\in\mathbb{R}^n} \sup_{x\in X} \limsup_{k\to\infty} \frac{1}{k} S_k(A-\langle c,\varphi-h\rangle)(x).\end{aligned}$$

Theorem 9 assures the following for subshifts of finite type.

Corollary 12. *Let $\varphi \in C^0(\Sigma,\mathbb{Q}^n)$ be a locally constant constraint and A be a Walters potential. Taking $r \in \varphi_*(\mathcal{M}_\sigma) \cap \mathbb{Q}^n$, assume the existence of an ergodic (A,r)-maximizing probability with respect to which φ is joint recurrent. Then*

$$\beta_{A,\varphi}(r) = \sup\left\{\int A\,d\mu : \mu \in \varphi_*^{-1}(r), \mu \textit{ periodic probability measure}\right\}.$$

The applicability of the corollary above is limited because the need of finding a maximizing probability with respect to which the constraint is joint recurrent. However, if we just concentrate on constraints taking values in $\mathbb{Q}$, the proposition 9, as we saw, assures the recurrence.

The ergodic requirement also circumscribes the applicability of corollary 12. Though, if we assume from now on in this section that our subshift of finite type $\sigma : \Sigma \to \Sigma$ is transitive, we have the following result.

Theorem 13. *Let $g \in C^0(\Sigma,\mathbb{R}^n)$ be a locally constant function. Every point of the interior of the rotation set $g_*(\mathcal{M}_\sigma)$ is a rotation vector of an ergodic probability measure.*

This theorem was obtained (see theorem 4.6 of [35]) for Ziemian when the function g depends on two coordinates. By passing to a higher block presentation of Σ, the reader familiarized with this argument (see otherwise the initial chapter of [23]) will notice that the general case is reduced to the situation treated by Ziemian.

We point out here the crucial importance of the locally constant assumption. Let $D \subset \mathbb{R}^n$ be a dense subset. Consider $\mathcal{F}_j(D)$ the collection of the functions of $C^0(\Sigma, D)$ that depend on $j+1$ coordinates. It is easily verified that $\bigcup \mathcal{F}_j(D)$ is dense in $C^0(\Sigma,\mathbb{R}^n)$. Given a function $g \in C^0(\Sigma,\mathbb{R}^n)$, we can without difficulty find a sequence $\{g_j\}$ convergent to g such that $g_j \in \mathcal{F}_j(D)$ for each index $j \geq 0$.

Proposition 14. *Suppose $\varphi \in C^0(\Sigma,\mathbb{Q})$ is a locally constant constraint and A is a Walters potential. Consider a rational number $r \in int(\varphi_*(\mathcal{M}_\sigma))$.*

Then

$$\beta_{A,\varphi}(r) = \sup\left\{\int A\, d\mu : \mu \in \varphi_*^{-1}(r), \mu \textit{ periodic probability measure}\right\}.$$

Proof. Taking into account theorem 10, fixed $\epsilon > 0$, it is enough to assure the existence of an ergodic probability measure ν with rotation number r satisfying $\beta_{A,\varphi}(r) - \epsilon < \int A\, d\nu$. Considering $\Phi = (\varphi, A)$, the strategy is to use the fact that the graph of the application $\beta_{A,\varphi}$ is part of the boundary of the rotation set $\Phi_*(\mathcal{M}_\sigma)$. Thus, if this rotation set consists of a segment, the existence of an ergodic probability measure as required follows from theorem 13.

It remains, therefore, to examine the other possibility: $\mathrm{int}(\Phi_*(\mathcal{M}_\sigma)) \neq \emptyset$. First, let $\{A_j\} \subset C^0(\Sigma)$ be a sequence convergent to A such that each function A_j depends on $j+1$ coordinates. Take any $\eta > \beta_{A,\varphi}(r) - \epsilon/2$ with $(r, \eta) \in \mathrm{int}(\Phi_*(\mathcal{M}_\sigma))$. When we put $\Phi_j = (\varphi, A_j)$, from the proposition 4 it results $(r, \eta) \in \mathrm{int}((\Phi_j)_*(\mathcal{M}_\sigma))$ for an index j sufficiently large, which can be supposed accomplishing besides $\|A_j - A\|_0 < \epsilon/2$. By the theorem 13, there exists an ergodic probability measure $\nu \in \mathcal{M}_\sigma$ satisfying $(\Phi_j)_*(\nu) = (r, \eta)$, or better, such that $\varphi_*(\nu) = r$ and $\int A_j\, d\nu = \eta > \beta_{A,\varphi}(r) - \epsilon/2$. However, once

$$\left|\int A_j\, d\nu - \int A\, d\nu\right| \leq \|A_j - A\|_0 < \frac{\epsilon}{2},$$

it happens $\int A\, d\nu > \beta_{A,\varphi}(r) - \epsilon$. □

A natural question is: when $\mathrm{int}(\varphi_*(\mathcal{M}_\sigma)) = \emptyset$?

In the case we consider here, there is a satisfactory answer. To present it, though, it is convenient to describe a few more properties of the general setting. A function $g \in C^0(X)$ is a (topological) coboundary when there exists a function $f \in C^0(X)$ such that $g = f \circ T - f$. Note that trivially every coboundary is a Walters function. Besides, two applications belonging to $C^0(X)$ are said cohomologous if their difference is a coboundary.

From results obtained by Bousch (in [4], consider theorem 4 using theorem 1), it follows a particularly interesting version of Livšic's theorem: an application $f \in C^0(\Sigma)$ is cohomologous to a constant if, and only if, f is a Walters function and $\mathrm{int}(f_*(\mathcal{M}_\sigma)) = \emptyset$. A function locally constant, it is important to point out, is a special example of Walters function.

Corollary 15. *Let $\varphi \in C^0(\Sigma, \mathbb{Q})$ be a locally constant constraint, not cohomologous to a constant. Assume A is a Walters potential. For each $c \in \mathbb{R}$, given $\epsilon > 0$, there exist a rational number $r \in int(\varphi_*(\mathcal{M}_\sigma))$ and a periodic probability measure $\mu \in \varphi_*^{-1}(r)$ satisfying $cr - \int A\, d\mu < \alpha_{A,\varphi}(c) + \epsilon$.*

5. Sub-actions and Differentiability of Alpha Functions

We will obtain in the present section a result relating the asymptotic behavior of optimal trajectories of certain sub-actions and the differential of an alpha function. Let us recall that, given a potential A, an application $u \in C^0(X)$ is a sub-action (for A) if

$$A + u - u \circ T \leq \beta_A.$$

General properties of sub-actions in different settings can be found, for instance, in [2, 3, 4, 7, 8, 11, 12, 18, 24, 25, 26, 30, 32, 33].

We denote
$$m_A = \left\{ \mu \in \mathcal{M}_T : \int A \, d\mu = \beta_A \right\},$$

the set of A-maximizing probabilities. Given a sub-action u, consider $A^u = A + u - u \circ T$. It is easy to see that

$$m_A = \left\{ \mu \in \mathcal{M}_T : \operatorname{supp}(\mu) \subset (A^u)^{-1}(\beta_A) \right\}.$$

Therefore, sub-actions help to locate the support of maximizing probabilities. The compact set $\mathbb{M}_A(u) = (A^u)^{-1}(\beta_A)$ will be called the contact locus[3] of the sub-action u. This is the set of points where the above sub-action inequality turns out to be an equality.

We will consider (X, T) a transitive, expansive dynamical system with a locally constant number of pre-images. Remind that expansiveness means there exist $\zeta > 0$ and $\kappa > 1$ such that, if $d(x, y) < \zeta$, then $d(x, y)\kappa \leq d(T(x), T(y))$. Besides, since the number of pre-images is assumed locally constant, there is $\xi > 0$ such that, whenever $d(x', y') < \xi$ and $x \in T^{-1}(x')$, we can find $y \in T^{-1}(y')$ accomplishing $d(x, y) < \zeta$.

For a function θ-Hölder f, the Hölder constant is

$$\text{Höld}_\theta(f) = \sup_{d(x,y)>0} \frac{|f(x) - f(y)|}{d(x, y)^\theta}.$$

As usual, we denote by $C^\theta(X)$ the Banach space of θ-Hölder functions with the norm $\| \cdot \|_\theta = \text{Höld}_\theta(\cdot) + \| \cdot \|_0$.

We can now present a result that indicates how the variation of the potential affects the sub-actions.

Proposition 16. *Consider (X, T) a transitive, expansive dynamical system with a locally constant number of pre-images. Let $\{B_j\}$ be a sequence of θ-Hölder functions converging in $C^\theta(X)$ to a potential A. Then, for each index j, we can find a sub-action v_j for the potential B_j, so that any accumulation point in $C^0(X)$ of the sequence $\{v_j\}$ is a sub-action for A.*

[3] In [7], it was suggested to call this set a Mañé set.

Proof. We have to show the existence of an equicontinuous and uniformly bounded sequence $\{v_j\}$. Indeed, as

$$\beta_{B_j} \geq B_j + v_j - v_j \circ T,$$

if $u \in C^0(X)$ is any accumulation point of $\{v_j\}$, taking limit in j, we immediately see that the function u is a sub-action for A.

Given a θ-Hölder potential B, it is possible to obtain a sub-action v for B that satisfies

$$|v(x) - v(y)| \leq \frac{\text{Höld}_\theta(B)}{\kappa^\theta - 1}\, d(x,y)^\theta, \text{ if } d(x,y) < \xi, \quad \text{and also}$$

$$\|v\|_0 \leq \text{Höld}_\theta(B) \left(\frac{2\xi^\theta}{\kappa^\theta - 1} + K \text{diam}(X)^\theta \right),$$

being the positive integer K depending just of ξ. For a proof of this statement, see the reasoning of theorem 4.7 in [19].

As we are considering for the potentials the convergence in $C^\theta(X)$, we obtain a sequence of sub-actions $\{v_j\}$ which is equicontinuous and uniformly bounded[4]. □

A sub-action u for a potential θ-Hölder A satisfies

$$u(x) - 2u(T(x)) + u(T^2(x)) \geq -\text{Höld}_\theta(A)\, d(x, T(x))^\theta$$

for every $x \in \mathbb{M}_A(u)$. Indeed, as

$$(A + u - u \circ T)(x) = \beta_A$$

and

$$(A + u - u \circ T)(T(x)) \leq \beta_A,$$

we show the claim by simple subtraction. Besides, for a point x belonging to the support of an A-maximizing probability, we have

$$\left| u(x) - 2u(T(x)) + u(T^2(x)) \right| \leq \text{Höld}_\theta(A)\, d(x, T(x))^\theta.$$

Given a Walters potential A, it is known the existence of a sub-action u such that

$$u(y) = \max_{T(x)=y} (A + u - \beta_A)(x).$$

This application u is called a calibrated sub-action for A.

We will suppose now a weaker assumption. We will consider a transitive dynamical system (X,T) verifying the property of weak expansion, that is, $T^{-1} : \mathcal{K}(X) \to \mathcal{K}(X)$ is 1-Lipschitz with respect to the Hausdorff metric.

[4]One should note that the sequence $\{v_j\}$ admits convergent subsequence in the C^0 topology, but of course not in general in the Hölder topology.

We can assure the existence of calibrated sub-actions also in this context (see [4]).

For a calibrated sub-action u, we will say that a sequence $\{x_j\} \subset X$ is an optimal trajectory (associated to the potential A) when $T(x_{j+1}) = x_j$ and

$$u(x_j) = A(x_{j+1}) + u(x_{j+1}) - \beta_A.$$

As remarked in the previous section, we have $\alpha_{A,\varphi}(c) = -\beta_{A-\langle c,\varphi\rangle}$. This is the last requirement for the formulation of the next theorem.

Theorem 17. *Let (X, T) be a transitive dynamical system satisfying the property of weak expansion. Consider a Walters potential $A \in C^0(X)$, as well as a Walters constraint $\varphi \in C^0(X, \mathbb{R}^n)$. Given an optimal trajectory $\{x_j\} \subset X$ associated to the potential $A - \langle c, \varphi\rangle$, if $\alpha_{A,\varphi}$ is differentiable at $c \in \mathbb{R}^n$, we verify*

$$\lim_{k\to\infty} \frac{1}{k} \sum_{j=0}^{k-1} \varphi(x_j) = D\alpha_{A,\varphi}(c).$$

Proof. Let $u_c \in C^0(X)$ be the calibrated sub-action used in the defitinion of the optimal trajectory $\{x_j\}$. So we have

$$u_c(x_0) = u_c(x_k) + \sum_{j=0}^{k-1} \left[A(x_j) - \langle c, \varphi(x_j)\rangle + \alpha_{A,\varphi}(c)\right].$$

Consider $\rho > 0$ and $\gamma \in \mathbb{R}^n$ with $\|\gamma\| = 1$. Taking any calibrated sub-action $u_{c+\rho\gamma} \in C^0(X)$ for the potential $A - \langle c + \rho\gamma, \varphi\rangle$, we obtain

$$u_{c+\rho\gamma}(x_0) \geq u_{c+\rho\gamma}(x_k) + \sum_{j=0}^{k-1} \left[A(x_j) - \langle c + \rho\gamma, \varphi(x_j)\rangle + \alpha_{A,\varphi}(c + \rho\gamma)\right].$$

From a simple subtraction, we get

$$\begin{aligned} -2\|u_c - u_{c+\rho\gamma}\|_0 &\leq \sum_{j=0}^{k-1} \left[\langle \rho\gamma, \varphi(x_j)\rangle + \alpha_{A,\varphi}(c) - \alpha_{A,\varphi}(c + \rho\gamma)\right] \\ &= \rho \left\langle \sum_{j=0}^{k-1} \varphi(x_j) - kD\alpha_{A,\varphi}(c), \gamma \right\rangle + o(k\rho), \end{aligned}$$

therefore

$$\rho \left\langle \frac{1}{k} \sum_{j=0}^{k-1} \varphi(x_j) - D\alpha_{A,\varphi}(c), \gamma \right\rangle = O\left(\frac{1}{k}\right) + o(\rho).$$

Now taking limsup when k tends to infinite and using the fact that ρ can be arbitrarily small, we obtain

$$\left\langle \limsup_{k\to\infty} \frac{1}{k}\sum_{j=0}^{k-1} \varphi(x_j) - D\alpha_{A,\varphi}(c), \gamma \right\rangle = 0$$

for all $\gamma \in \mathbb{R}^n$ with $\|\gamma\| = 1$, in other words,

$$\limsup_{k\to\infty} \frac{1}{k}\sum_{j=0}^{k-1} \varphi(x_j) = D\alpha_{A,\varphi}(c).$$

An analogous argument can be applied for the liminf and this proves the theorem. $\square$

A similar result for the discrete Aubry-Mather problem is presented in theorem 6.2 of [13].

References

[1] G. Atkinson, Recurrence of co-cycles and random walks, *The Journal of the London Mathematical Society* **13** (1976), 486-488.

[2] F. M. Branco, *Subação para transformações unidimensionais*, PhD thesis, Universidade Federal do Rio Grande do Sul, 2003.

[3] T. Bousch, Le poisson n'a pas d'arêtes, *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques* **36** (2000), 489-508.

[4] T. Bousch, La condition de Walters, *Annales Scientifiques de l'École Normale Supérieure* **34** (2001), 287-311.

[5] T. Bousch, J. Mairesse, Asymptotic height optimization for topical IFS, Tetris heaps, and the finiteness conjecture, *Journal of the American Mathematical Society* **15** (2002), 77-111.

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

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

[8] J. P. Conze, Y. Guivarc'h, *Croissance des sommes ergodiques et principe variationnel*, manuscript.

[9] J. Franks, Realizing rotation vectors for torus homeomorphisms, *Transactions of the American Mathematical Society* **311** (1989), 107-115.

[10] E. Garibaldi, *Otimização ergódica: da maximização relativa aos homeomorfismos expansivos*, PhD thesis, Universidade Federal do Rio Grande do Sul, 2006.

[11] E. Garibaldi, A. O. Lopes, On the Aubry-Mather theory for symbolic dynamics, to appear in *Ergodic Theory and Dynamical Systems.*

[12] E. Garibaldi, A. O. Lopes, P. Thieullen, On separating sub-actions, *preprint.*

[13] D. A. Gomes, Viscosity solution method and the discrete Aubry-Mather problem, *Discrete and Continuous Dynamical Systems, Series A* **13** (2005), 103-116.

[14] M. R. Herman, Inégalités a priori pour des tores lagrangiens invariants par des difféomorphismes symplectiques, *Publications Mathématiques de l'Institut des Hautes Études Scientifiques* **70** (1989), 47-101.

[15] B. R. Hunt, G. C. Yuan, Optimal orbits of hyperbolic systems, *Nonlinearity* **12** (1999), 1207-1224.

[16] O. Jenkinson, *Conjugacy rigidity, cohomological triviality and barycentres of invariant measures*, PhD thesis, Warwick University, 1996.

[17] O. Jenkinson, Geometric barycentres of invariant measures for circle maps, *Ergodic Theory and Dynamical Systems* **21** (2001), 511-532.

[18] O. Jenkinson, Rotation, entropy, and equilibrium states, *Transactions of the American Mathematical Society* **353** (2001), 3713-3739.

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

[20] J. Kwapisz, Every convex polygon with rational vertices is a rotation set, *Ergodic Theory and Dynamical Systems* **12** (1992), 333-339.

[21] J. Kwapisz, A toral diffeomorphism with a nonpolygonal rotation set, *Nonlinearity* **8** (1995), 461-476.

[22] J. Kwapisz, A priori degeneracy of one-dimensional rotation sets for periodic point free torus maps, *Transactions of the American Mathematical Society* **354** (2002), 2865-2895.

[23] D. Lind, B. Marcus, *An introduction to symbolic dynamics and coding*, Cambridge University Press, 1995.

[24] A. O. Lopes, V. Rosas, R. O. Ruggiero, Cohomology and subcohomology for expansive geodesic flows, *Discrete and Continuous Dynamical Systems, Series A* **17** (2007), 403-422.

[25] A. O. Lopes, P. Thieullen, Sub-actions for Anosov diffeomorfisms, *Astérisque* **287** (2003), 135-146.

[26] A. O. Lopes, P. Thieullen, Sub-actions for Anosov flows, *Ergodic Theory and Dynamical Systems* **25** (2005), 605-628.

[27] A. O. Lopes, P. Thieullen, Mather measures and the Bowen-Series transformation, *Annales de l'Institut Henri Poincaré, Analyse non Linéaire* **23** (2006), 663-682.

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

[29] M. Misiurewicz, K. Ziemian, Rotation sets for maps of tori, *The Journal of the London Mathematical Society* **40** (1989), 490-506.

[30] M. Pollicott, R. Sharp, Livsic theorems, maximizing measures and the stable norm, *Dynamical Systems* **19** (2004), 75-88.

[31] L. Radu, Duality in thermodynamic formalism, *preprint*.

[32] S. V. Savchenko, Cohomological inequalities for finite Markov chains, *Functional Analysis and Its Applications* **33** (1999), 236-238.

[33] R. R. Souza, Sub-actions for weakly hyperbolic one-dimensional systems, *Dynamical Systems* **18** (2003), 165-179.

[34] P. Walters, Invariant measures and equilibrium states for some mappings which expand distances, *Transactions of the American Mathematical Society* **236** (1978), 121-153.

[35] K. Ziemian, Rotation sets for subshifts of finite type, *Fundamenta Mathematicae* **146** (1995), 189-201.