Ergod. Th. & Dynam. Sys. (2008), **28**, 791–815
doi: 10.1017/S0143385707000491

On the Aubry–Mather theory for symbolic dynamics

E. GARIBALDI† and A. O. LOPES‡

† *Institut de Mathématiques, Université Bordeaux 1, F-33405 Talence, France*
(e-mail: Eduardo.Garibaldi@math.u-bordeaux1.fr)
‡ *Instituto de Matemática, UFRGS, 91509-900 Porto Alegre, Brazil*
(e-mail: artur.lopes@ufrgs.br)

(*Received* 15 *August* 2006 *and accepted in revised form* 1 *June* 2007)

Abstract. We propose a new model of ergodic optimization for expanding dynamical systems: the holonomic setting. In fact, we introduce an extension of the standard model used in this theory. The formulation we consider here is quite natural if one wants a meaning for possible variations of a real trajectory under the forward shift. In other contexts (for twist maps, for instance), this property appears in a crucial way. A version of the Aubry–Mather theory for symbolic dynamics is introduced. We are mainly interested here in problems related to the properties of maximizing probabilities for the two-sided shift. Under the transitive hypothesis, we show the existence of sub-actions for Hölder potentials also in the holonomic setting. We analyze then connections between calibrated sub-actions and the Mañé potential. A representation formula for calibrated sub-actions is presented, which drives us naturally to a classification theorem for these sub-actions. We also investigate properties of the support of maximizing probabilities.

1. *The holonomic condition*

Consider X a compact metric space. Given a continuous transformation $T : X \to X$, we denote by $\mathcal{M}_T$ the convex set of T-invariant Borel probability measures. As usual, we consider on $\mathcal{M}_T$ the weak* topology.

The triple $(X, T, \mathcal{M}_T)$ is the standard model used in ergodic optimization. Thus, given a potential $A \in C^0(X)$, one of the main objectives is the characterization of maximizing probabilities, that is, the probabilities belonging to

$$\left\{\mu \in \mathcal{M}_T : \int_X A(x)\, d\mu(x) = \max_{\nu \in \mathcal{M}_T} \int_X A(x)\, d\nu(x)\right\}.$$

Several results have been obtained related to this maximizing question, among them [**2–4, 9, 16–19**]. For maximization with constraints, see [**12, 13, 20**]. Naturally, if we change the maximizing notion for the minimizing one, the analogous properties will be true.

Our focus here will be on symbolic dynamics. So let $\sigma : \Sigma \to \Sigma$ be a one-sided subshift of finite type given by an $r \times r$ transition matrix $\mathbf{M}$. More precisely, we have

$$\Sigma = \{\mathbf{x} \in \{1, \ldots, r\}^{\mathbb{N}} : \mathbf{M}(x_j, x_{j+1}) = 1 \text{ for all } j \geq 0\}$$

and σ is the left shift acting on Σ, $\sigma(x_0, x_1, \ldots) = (x_1, x_2, \ldots)$. Recall that, for fixed $\lambda \in (0, 1)$, we consider Σ with the metric $d(\mathbf{x}, \bar{\mathbf{x}}) = \lambda^k$, where $\mathbf{x} = (x_0, x_1, \ldots)$, $\bar{\mathbf{x}} = (\bar{x}_0, \bar{x}_1, \ldots) \in \Sigma$ and $k = \min\{j : x_j \neq \bar{x}_j\}$.

In this particular situation, given a continuous potential $A : \Sigma \to \mathbb{R}$, one should be *a priori* interested in A-maximizing probabilities for the triple $(\Sigma, \sigma, \mathcal{M}_\sigma)$.

Nevertheless, this standard model of ergodic optimization has a major difference to the twist maps theory or to the Lagrangian Aubry–Mather problem: the dynamics of the shift is not defined (via a critical path problem) from the potential to be maximized. In similar terms, in the usual shift standard model, the notion of maximizing segment is not present. One would like to have small variations of an optimal trajectory, by means of a path which is not a true trajectory, but a small variation of a real trajectory of the dynamical system. We will describe a model of ergodic optimization for subshifts of finite type where the concept of maximizing segment can be introduced: the holonomic setting. In Aubry–Mather theory for Lagrangian systems (continuous or discrete time), the set of holonomic probabilities has been considered before by Mañé, Mather, Contreras and Gomes. Main references on these topics are [**1, 7, 11, 15, 21**].

In order to define the holonomic model of ergodic optimization, we introduce the dual subshift $\sigma^* : \Sigma^* \to \Sigma^*$ using as transition matrix the transposed $\mathbf{M}^{\mathrm{T}}$. In clear terms, we consider thus the space

$$\Sigma^* = \{\mathbf{y} \in \{1, \ldots, r\}^{\mathbb{N}} : \mathbf{M}(y_{j+1}, y_j) = 1 \text{ for all } j \geq 0\}$$

and the shift $\sigma^*(\ldots, y_1, y_0) = (\ldots, y_2, y_1)$. It is possible, in this way, to identify the space of the dynamics $(\hat{\Sigma}, \hat{\sigma})$, the natural extension of (Σ, σ), with a subset of $\Sigma^* \times \Sigma$. In fact, if $\mathbf{y} = (\ldots, y_1, y_0) \in \Sigma^*$ and $\mathbf{x} = (x_0, x_1, \ldots) \in \Sigma$, then $\hat{\Sigma}$ will be the set of points $(\mathbf{y}, \mathbf{x}) = (\ldots, y_1, y_0 \mid x_0, x_1, \ldots) \in \Sigma^* \times \Sigma$ such that (y_0, x_0) is an allowed word, namely, such that $\mathbf{M}(y_0, x_0) = 1$.

We define then the transformation $\tau : \hat{\Sigma} \to \Sigma$ by

$$\tau(\mathbf{y}, \mathbf{x}) = \tau_{\mathbf{y}}(\mathbf{x}) = (y_0, x_0, x_1, \ldots).$$

Note that $\hat{\sigma}^{-1}(\mathbf{y}, \mathbf{x}) = (\sigma^*(\mathbf{y}), \tau_{\mathbf{y}}(\mathbf{x}))$.

Let $\mathcal{M}$ be the convex set of probability measures over the Borel sigma-algebra of $\hat{\Sigma}$.

Definition 1. In an analogous way to [**15**], we consider the convex compact subset

$$\mathcal{M}_0 = \left\{ \hat{\mu} \in \mathcal{M} : \int_{\hat{\Sigma}} f(\tau_{\mathbf{y}}(\mathbf{x}))\, d\hat{\mu}(\mathbf{y}, \mathbf{x}) = \int_{\hat{\Sigma}} f(\mathbf{x})\, d\hat{\mu}(\mathbf{y}, \mathbf{x})\ \forall\, f \in C^0(\Sigma) \right\}.$$

A probability $\hat{\mu} \in \mathcal{M}_0$ will be called holonomic.

Note that $\mathcal{M}_{\hat{\sigma}} \subset \mathcal{M}_0$. It is also not difficult to verify that, whenever $\mu^* \times \mu \in \mathcal{M}_0$, we have $\mu \in \mathcal{M}_\sigma$. Moreover, if $\hat{\mu} \in \mathcal{M}_0$, then $\hat{\mu} \circ \pi_1^{-1} \in \mathcal{M}_\sigma$, where $\pi_1 : \hat{\Sigma} \to \Sigma$ is the canonical projection. Indeed, if $f \in C^0(\Sigma)$, then

$$\begin{aligned}
\int_\Sigma f \circ \sigma(\mathbf{x})\, d(\hat{\mu} \circ \pi_1^{-1})(\mathbf{x}) &= \int_{\hat{\Sigma}} f \circ \sigma(\mathbf{x})\, d\hat{\mu}(\mathbf{y}, \mathbf{x}) \\
&= \int_{\hat{\Sigma}} f \circ \sigma(\tau_{\mathbf{y}}(\mathbf{x}))\, d\hat{\mu}(\mathbf{y}, \mathbf{x}) \\
&= \int_{\hat{\Sigma}} f(\mathbf{x})\, d\hat{\mu}(\mathbf{y}, \mathbf{x}) \\
&= \int_\Sigma f(\mathbf{x})\, d(\hat{\mu} \circ \pi_1^{-1})(\mathbf{x}).
\end{aligned}$$

However, $\mathcal{M}_0$ does not contain just $\hat{\sigma}$-invariant probabilities. In fact, if $\mathbf{x} \in \Sigma$ is a periodic point of period M, fix any subset $\{\mathbf{y}^0, \ldots, \mathbf{y}^{M-1}\} \subset \Sigma^*$ with $y_0^j = x_{M-1+j}$ for $0 \le j \le M-1$. It is easy to see that

$$\hat{\mu} = \frac{1}{M} \sum_{j=0}^{M-1} \delta_{\mathbf{y}^j} \times \delta_{\sigma^j(\mathbf{x})} \in \mathcal{M}_0.$$

For the ergodic optimization problem, there is very little difference (from a purely abstract point of view) in relation to which convex compact set of probability measures over the Borel sigma-algebra is made the maximization. In fact, an adaptation of [**9,** Proposition 10] assures that, when considering a convex compact subset $\mathcal{N} \subset \mathcal{M}$, a generic Hölder potential admits a single maximizing probability in $\mathcal{N}$.

Taking a continuous application $A : \hat{\Sigma} \to \mathbb{R}$, a natural situation is then to formulate the maximization problem over the set $\mathcal{M}_0$.

Definition 2. Given a potential $A \in C^0(\hat{\Sigma})$, denote

$$\beta_A = \max_{\hat{\mu} \in \mathcal{M}_0} \int_{\hat{\Sigma}} A(\mathbf{y}, \mathbf{x})\, d\hat{\mu}(\mathbf{y}, \mathbf{x}).$$

We point out that sometimes, even if one is interested just in the problem for a Hölder potential $A : \Sigma \to \mathbb{R}$, one has to go to the dual problem and consider the dual potential $A^* : \Sigma^* \to \mathbb{R}$. This happens, for instance, when one is trying to analyze a large deviation principle for the equilibrium probabilities associated to the family of Hölder potentials $\{tA\}_{t>0}$ (see [**2**]).

Actually, the maximization problem over $\mathcal{M}_{\hat{\sigma}}$ is not so interesting, because any Hölder potential $A : \hat{\Sigma} \to \mathbb{R}$ is cohomologous to a potential that depends just on future coordinates (see, for instance, [**23**]). In this case, the problem can in principle be analyzed in the standard model, that is, over $\mathcal{M}_\sigma$.

Furthermore, in order to analyze maximization of the integral of a potential $A \in C^0(\Sigma)$, no new maximal value will be found, because

$$\max_{\hat{\mu} \in \mathcal{M}_0} \int_{\hat{\Sigma}} A(\mathbf{x})\, d\hat{\mu}(\mathbf{y}, \mathbf{x}) = \max_{\mu \in \mathcal{M}_\sigma} \int_\Sigma A(\mathbf{x})\, d\mu(\mathbf{x}).$$

Indeed, the correspondence $\hat{\mu} \in \mathcal{M}_0 \mapsto \hat{\mu} \circ \pi_1^{-1} \in \mathcal{M}_\sigma$ preserves the integration on $C^0(\Sigma)$ and the same property is verified by the correspondence $\mu \in \mathcal{M}_\sigma \mapsto \mu \circ \pi_1 \circ \hat{\sigma}^{-1} \in \mathcal{M}_0$.

Therefore, we could say that the holonomic model of ergodic optimization $(\hat{\Sigma}, \hat{\sigma}, \mathcal{M}_0)$ is an extension of the standard model $(\Sigma, \sigma, \mathcal{M}_\sigma)$.

This paper is part of the first author's PhD thesis [**12**]. We will be interested here in the maximization question over $\mathcal{M}_0$ and, if possible, in some properties that one can get for the problem over (Σ, σ). In §2, we will show the dual identity

$$\beta_A = \inf_{f \in C^0(\Sigma)} \max_{(\mathbf{y},\mathbf{x}) \in \hat{\Sigma}} [A(\mathbf{y}, \mathbf{x}) + f(\mathbf{x}) - f(\tau_{\mathbf{y}}(\mathbf{x}))].$$

We will then analyze the problem of finding a function $u \in C^0(\Sigma)$ which realizes the infimum of the previous expression, that is, a sub-action for A.

Definition 3. A sub-action $u \in C^0(\Sigma)$ for the potential $A \in C^0(\hat{\Sigma})$ is a function satisfying, for any $(\mathbf{y}, \mathbf{x}) \in \hat{\Sigma}$,

$$u(\mathbf{x}) \le u(\tau_{\mathbf{y}}(\mathbf{x})) - A(\mathbf{y}, \mathbf{x}) + \beta_A.$$

Assuming the dynamics (Σ, σ) is topologically mixing and the potential A is Hölder, we will show in §3 the existence of a Hölder sub-action of maximal character. Furthermore, under the transitivity hypothesis, for a potential θ-Hölder, we will show that we can always find a calibrated sub-action $u \in C^\theta(\Sigma)$.

Definition 4. A calibrated sub-action $u \in C^0(\Sigma)$ for $A \in C^0(\hat{\Sigma})$ is a function satisfying

$$u(\mathbf{x}) = \min_{\mathbf{y} \in \Sigma^*_{\mathbf{x}}} [u(\tau_{\mathbf{y}}(\mathbf{x})) - A(\mathbf{y}, \mathbf{x}) + \beta_A],$$

where, for each point $\mathbf{x} \in \Sigma$, we denote by $\Sigma^*_{\mathbf{x}}$ the subset of elements $\mathbf{y} \in \Sigma^*$ such that $(\mathbf{y}, \mathbf{x}) \in \hat{\Sigma}$.

In the transitive context, we will introduce in §4 the Mañé potential $S_A : \Sigma \times \Sigma \to \mathbb{R} \cup \{+\infty\}$ (the terminology is borrowed from Aubry–Mather theory). Thus, we will establish a family of Hölder calibrated sub-actions, namely, $\{S_A(\mathbf{x}, \cdot)\}_{\mathbf{x} \in \Omega(A)}$, where $\Omega(A)$ denotes the set of non-wandering points with respect to the potential $A \in C^\theta(\hat{\Sigma})$. All these notions will be precisely defined later. Besides, these concepts already appear in [**9**] for the forward shift setting.

Definition 5. We will denote by

$$m_A = \left\{ \hat{\mu} \in \mathcal{M}_0 : \int_{\hat{\Sigma}} A(\mathbf{y}, \mathbf{x})\, d\hat{\mu}(\mathbf{y}, \mathbf{x}) = \beta_A \right\}$$

the set of the A-maximizing holonomic probabilities.

When we investigate the connections between sub-actions and the supports of holonomic probabilities, the A-maximizing holonomic probability notion is of great importance. One of the main results of §5 is the representation formula for calibrated sub-actions. More specifically, given a calibrated sub-action u for a potential $A \in C^\theta(\hat{\Sigma})$, the following expression holds:

$$u(\bar{\mathbf{x}}) = \inf_{\mathbf{x} \in \Omega(A)} [u(\mathbf{x}) + S_A(\mathbf{x}, \bar{\mathbf{x}})].$$

Such characterization is analogous to the one obtained for weak KAM solutions in Lagrangian systems (see [**6**]). Under the transitivity hypothesis, this representation formula and its reciprocal will describe, by means of an isometric bijection, the set of the calibrated sub-actions for a Hölder potential A. We will show yet that $\hat{\mu} \in m_A$ with $\hat{\mu} \circ \pi_1^{-1}$ ergodic implies $\pi_1(\mathrm{supp}(\hat{\mu})) \subset \Omega(A)$. This property will drive us naturally to other questions like, for instance, the possibility of reducing contact loci.

2. *The dual formulation*

We start by presenting the main goal of this section.

THEOREM 1. *Given a potential $A \in C^0(\hat{\Sigma})$, we have*

$$\beta_A = \inf_{f \in C^0(\Sigma)} \max_{(\mathbf{y},\mathbf{x}) \in \hat{\Sigma}} [A(\mathbf{y}, \mathbf{x}) + f(\mathbf{x}) - f(\tau_{\mathbf{y}}(\mathbf{x}))].$$

One observes that this formula corresponds in Lagrangian Aubry–Mather theory to the characterization of Mañé's critical value (see [**8,** Theorem A]). Theorem 1 is just a consequence of the Fenchel–Rockafellar theorem. For the standard model $(X, T, \mathcal{M}_T)$, a similar result was established before (consult, for instance, [**10, 24**]). We will present, anyway, the complete proof for the holonomic setting.

First, consider the convex correspondence $F : C^0(\hat{\Sigma}) \to \mathbb{R}$ defined by $F(g) = \max(A + g)$. Consider also the subset

$$\mathcal{C} = \{g \in C^0(\hat{\Sigma}) : g(\mathbf{y}, \mathbf{x}) = f(\mathbf{x}) - f(\tau_{\mathbf{y}}(\mathbf{x})), \text{ for some } f \in C^0(\Sigma)\}.$$

We establish then a concave correspondence $G : C^0(\hat{\Sigma}) \to \mathbb{R} \cup \{-\infty\}$ taking $G(g) = 0$ if $g \in \bar{\mathcal{C}}$ and $G(g) = -\infty$ otherwise.

Let $\mathcal{S}$ be the set of the signed measures over the Borel sigma-algebra of $\hat{\Sigma}$. Remember that the corresponding Fenchel transforms, $F^* : \mathcal{S} \to \mathbb{R} \cup \{+\infty\}$ and $G^* : \mathcal{S} \to \mathbb{R} \cup \{-\infty\}$, are given by

$$F^*(\hat{\mu}) = \sup_{g \in C^0(\hat{\Sigma})} \left[\int_{\hat{\Sigma}} g(\mathbf{y}, \mathbf{x})\, d\hat{\mu}(\mathbf{y}, \mathbf{x}) - F(g) \right]$$

and

$$G^*(\hat{\mu}) = \inf_{g \in C^0(\hat{\Sigma})} \left[\int_{\hat{\Sigma}} g(\mathbf{y}, \mathbf{x})\, d\hat{\mu}(\mathbf{y}, \mathbf{x}) - G(g) \right].$$

Denote

$$\mathcal{S}_0 = \left\{ \hat{\mu} \in \mathcal{S} : \int_{\hat{\Sigma}} f(\tau_{\mathbf{y}}(\mathbf{x}))\, d\hat{\mu}(\mathbf{y}, \mathbf{x}) = \int_{\hat{\Sigma}} f(\mathbf{x})\, d\hat{\mu}(\mathbf{y}, \mathbf{x})\ \forall f \in C^0(\Sigma) \right\}.$$

LEMMA 2. *Given F and G as above, we verify*

$$F^*(\hat{\mu}) = \begin{cases} -\displaystyle\int_{\hat{\Sigma}} A(\mathbf{y}, \mathbf{x})\, d\hat{\mu}(\mathbf{y}, \mathbf{x}) & \textit{if } \hat{\mu} \in \mathcal{M}, \\ +\infty & \textit{otherwise,} \end{cases}$$

and

$$G^*(\hat{\mu}) = \begin{cases} 0 & \textit{if } \hat{\mu} \in \mathcal{S}_0, \\ -\infty & \textit{otherwise.} \end{cases}$$

Proof. Assume first that $\hat{\mu} \in \mathcal{S}$ is not positive, that is, $\hat{\mu}$ gives a negative value for some Borel set. Therefore, we can find a sequence of functions $\{g_j\} \subset C^0(\hat{\Sigma}, \mathbb{R}^-)$ such that $\lim \int_{\hat{\Sigma}} g_j(\mathbf{y}, \mathbf{x})\, d\hat{\mu}(\mathbf{y}, \mathbf{x}) = +\infty$. Once $F(g_j) \le F(0) < +\infty$, we have $F^*(\hat{\mu}) = +\infty$.

Suppose that $\hat{\mu} \in \mathcal{S}$ is such that $\hat{\mu} \ge 0$ and $\hat{\mu}(\hat{\Sigma}) \ne 1$. In this case, we observe that

$$\sup_{g \in C^0(\hat{\Sigma})} \left[\int_{\hat{\Sigma}} g(\mathbf{y}, \mathbf{x})\, d\hat{\mu}(\mathbf{y}, \mathbf{x}) - F(g) \right] \ge \sup_{a \in \mathbb{R}} \left[\int_{\hat{\Sigma}} a\, d\hat{\mu}(\mathbf{y}, \mathbf{x}) - F(a) \right]$$
$$= \sup_{a \in \mathbb{R}} [a(\hat{\mu}(\hat{\Sigma}) - 1) - F(0)] = +\infty.$$

On the other hand, when we consider $\hat{\mu} \in \mathcal{M}$, directly from the inequality $\int_{\hat{\Sigma}} A(\mathbf{y}, \mathbf{x})\, d\hat{\mu}(\mathbf{y}, \mathbf{x}) + \int_{\hat{\Sigma}} g(\mathbf{y}, \mathbf{x})\, d\hat{\mu}(\mathbf{y}, \mathbf{x}) \le F(g)$, we have

$$- \int_{\hat{\Sigma}} A(\mathbf{y}, \mathbf{x})\, d\hat{\mu}(\mathbf{y}, \mathbf{x}) \ge \sup_{g \in C^0(\hat{\Sigma})} \left[\int_{\hat{\Sigma}} g(\mathbf{y}, \mathbf{x})\, d\hat{\mu}(\mathbf{y}, \mathbf{x}) - F(g) \right].$$

Once $F(-A) = 0$, we get the characterization of F^*.

Now we will consider G^*. If $\hat{\mu} \notin \mathcal{S}_0$, there exists a function $f \in C^0(\Sigma)$ such that $\int_{\hat{\Sigma}} f(\tau_{\mathbf{y}}(\mathbf{x}))\, d\hat{\mu}(\mathbf{y}, \mathbf{x}) \ne \int_{\hat{\Sigma}} f(\mathbf{x})\, d\hat{\mu}(\mathbf{y}, \mathbf{x})$. Therefore, we verify

$$G^*(\hat{\mu}) = \inf_{g \in \mathcal{C}} \int_{\hat{\Sigma}} g(\mathbf{y}, \mathbf{x})\, d\hat{\mu}(\mathbf{y}, \mathbf{x})$$
$$\le \inf_{a \in \mathbb{R}} a \int_{\hat{\Sigma}} [f(\tau_{\mathbf{y}}(\mathbf{x})) - f(\mathbf{x})]\, d\hat{\mu}(\mathbf{y}, \mathbf{x}) = -\infty.$$

Besides, for $\hat{\mu} \in \mathcal{S}_0$, clearly $G^*(\hat{\mu}) = 0$. □

Using this lemma, we can show the dual expression of the beta constant $\beta_A = \max_{\hat{\mu} \in \mathcal{M}_0} \int_{\hat{\Sigma}} A(\mathbf{y}, \mathbf{x})\, d\hat{\mu}(\mathbf{y}, \mathbf{x})$.

Proof of Theorem 1. Once the correspondence F is Lipschitz, the Fenchel–Rockafellar duality theorem assures

$$\sup_{g \in C^0(\hat{\Sigma})} [G(g) - F(g)] = \inf_{\hat{\mu} \in \mathcal{S}} [F^*(\hat{\mu}) - G^*(\hat{\mu})].$$

Thus, by Lemma 2,

$$\sup_{g \in \mathcal{C}} \left[- \max_{(\mathbf{y}, \mathbf{x}) \in \hat{\Sigma}} (A + g)(\mathbf{y}, \mathbf{x}) \right] = \inf_{\hat{\mu} \in \mathcal{M}_0} \left[- \int_{\hat{\Sigma}} A(\mathbf{y}, \mathbf{x})\, d\hat{\mu}(\mathbf{y}, \mathbf{x}) \right].$$

Finally, from the definition of $\mathcal{C}$, we get the statement of the theorem. □

Relative maximization is studied in [**13**]. In this case, the dual formula is also true. More specifically, if we introduce a constraint $\varphi \in C^0(\hat{\Sigma}, \mathbb{R}^n)$ with coordinate functions $\varphi_1, \ldots, \varphi_n$, we can then consider an induced map $\varphi_* \in C^0(\mathcal{M}_0, \mathbb{R}^n)$ given by

$$\varphi_*(\hat{\mu}) = \left(\int_{\hat{\Sigma}} \varphi_1(\mathbf{y}, \mathbf{x})\, d\hat{\mu}(\mathbf{y}, \mathbf{x}), \ldots, \int_{\hat{\Sigma}} \varphi_n(\mathbf{y}, \mathbf{x})\, d\hat{\mu}(\mathbf{y}, \mathbf{x}) \right).$$

Thus, if $A \in C^0(\hat{\Sigma})$, we can immediately define a concave and continuous function $\beta_{A,\varphi} : \varphi_*(\mathcal{M}_0) \to \mathbb{R}$ by

$$\beta_{A,\varphi}(h) = \max_{\hat{\mu} \in \varphi_*^{-1}(h)} \int_{\hat{\Sigma}} A(\mathbf{y}, \mathbf{x})\, d\hat{\mu}(\mathbf{y}, \mathbf{x}).$$

Using a slightly more refined argument as [**24**], we could demonstrate the dual formula for a beta function,

$$\beta_{A,\varphi}(h) = \inf_{(f,c)\in C^0(\Sigma)\times\mathbb{R}^n} \max_{(\mathbf{y},\mathbf{x})\in\hat{\Sigma}} (A + f\circ\pi_1 - f\circ\pi_1\circ\hat{\sigma}^{-1} - \langle c, \varphi - h\rangle)(\mathbf{y},\mathbf{x}).$$

Nevertheless, the unconstrained dual formula raises a natural question: Can we find functions accomplishing the infimum of the dual expression? In an equivalent way, is there a function $u \in C^0(\Sigma)$ such that

$$A + u\circ\pi_1 - u\circ\pi_1\circ\hat{\sigma}^{-1} \le \beta_A \ ?$$

As we mentioned in the first section, we call any function u as above a sub-action for A. This terminology is motivated by the inequality

$$A + u\circ\sigma - u \le \beta_A,$$

which is present in the usual definition of a sub-action u for the forward shift setting (see [**9**] for instance). The next sections are mainly dedicated to showing the existence of sub-actions in the holonomic setting.

3. *Sub-actions: maximality and calibration*

We start by showing not only the existence of sub-actions but also, in fact, the existence of a maximal sub-action. To that end, remember that a dynamical system (X, T) is topologically mixing if, for any pair of non-empty open sets $D, E \subset X$, there is an integer $K > 0$ such that $T^k(D)\cap E \neq \emptyset$ for all $k > K$.

PROPOSITION 3. *Consider any topologically mixing subshift of finite type $\sigma : \Sigma \to \Sigma$ and a potential $A \in C^\theta(\hat{\Sigma})$. Then, there exists a sub-action $u_A \in C^\theta(\Sigma, \mathbb{R}^-)$ such that, for any other sub-action $u \in C^0(\Sigma, \mathbb{R}^-)$, we have $u_A \ge u$.*

A sub-action like this one (not necessarily Hölder) will be called maximal.

Proof. Without loss of generality, we can assume $\beta_A = 0$. Then, for each $\mathbf{x} \in \Sigma$, set

$$u_A(\mathbf{x}) = \inf\left\{-\sum_{j=0}^{k-1} A(\mathbf{y}^j, \mathbf{x}^j) : k \ge 0, \mathbf{x}^0 = \mathbf{x}, \mathbf{y}^j \in \Sigma^*_{\mathbf{x}^j}, \mathbf{x}^{j+1} = \tau_{\mathbf{y}^j}(\mathbf{x}^j)\right\}.$$

By convention, we assume that the sum is zero when $k = 0$.

Suppose for a moment that u_A is a well-defined Hölder application. Note that, if $\mathbf{y}^0 = \mathbf{y}$ and $\mathbf{x}^0 = \mathbf{x}$, then

$$\begin{aligned} A(\mathbf{y},\mathbf{x}) &= \sum_{j=0}^{k} A(\mathbf{y}^j, \mathbf{x}^j) - \sum_{j=0}^{k-1} A(\mathbf{y}^{j+1}, \mathbf{x}^{j+1}) \\ &\le -\sum_{j=0}^{k-1} A(\mathbf{y}^{j+1}, \mathbf{x}^{j+1}) - u_A(\mathbf{x}). \end{aligned}$$

Clearly $\mathbf{x}^1 = \tau_{\mathbf{y}^0}(\mathbf{x}^0) = \tau_{\mathbf{y}}(\mathbf{x})$. Thus, since the inequality is true for all $k \ge 0$ and any points $(\mathbf{y}^1, \mathbf{x}^1), \ldots, (\mathbf{y}^k, \mathbf{x}^k) \in \hat{\Sigma}$ such that $\mathbf{x}^{j+1} = \tau_{\mathbf{y}^j}(\mathbf{x}^j)$, it follows that $A(\mathbf{y},\mathbf{x}) \le u_A(\tau_{\mathbf{y}}(\mathbf{x})) - u_A(\mathbf{x})$, that is, u_A is a sub-action for the potential A.

So let us prove that the function u_A is well defined. Remember that, when $\bar{\mathbf{x}} \in \Sigma$ is a periodic point of period k, if we choose any points $\bar{\mathbf{y}}^j \in \Sigma^*$ satisfying $\bar{y}_0^j = \bar{x}_{k-(j+1)}$, we obtain $\hat{\mu} = (1/k) \sum_{j=0}^{k-1} \delta_{\bar{\mathbf{y}}^j} \times \delta_{\sigma^{k-j}(\bar{\mathbf{x}})} \in \mathcal{M}_0$. Hence, we immediately verify

$$-\sum_{j=0}^{k-1} A(\bar{\mathbf{y}}^j, \sigma^{k-j}(\bar{\mathbf{x}})) = -k \int_{\hat{\Sigma}} A(\mathbf{y}, \mathbf{x})\, d\hat{\mu}(\mathbf{y}, \mathbf{x}) \geq 0.$$

Given $\mathbf{x} \in \Sigma$, we choose then points $(\mathbf{y}^0, \mathbf{x}^0), \ldots, (\mathbf{y}^{k-1}, \mathbf{x}^{k-1}) \in \hat{\Sigma}$ satisfying $\mathbf{x}^0 = \mathbf{x}$ and $\mathbf{x}^{j+1} = \tau_{\mathbf{y}^j}(\mathbf{x}^j)$. As (Σ, σ) is topologically mixing, there exists an integer $K > 0$ such that, for any $k > K$, we can find a periodic point $\bar{\mathbf{x}}$ of period k satisfying $d(\mathbf{x}^k, \bar{\mathbf{x}}) < \lambda^{k-K}$, where $\mathbf{x}^k = \tau_{\mathbf{y}^{k-1}}(\mathbf{x}^{k-1})$. Thus, when we put $\bar{\mathbf{y}}^j = \mathbf{y}^j$ for $K \leq j \leq k-1$, it follows that

$$\left| \sum_{j=0}^{k-1} A(\mathbf{y}^j, \mathbf{x}^j) - \sum_{j=0}^{k-1} A(\bar{\mathbf{y}}^j, \sigma^{k-j}(\bar{\mathbf{x}})) \right| \leq \frac{\text{Höld}_\theta(A)}{1-\lambda^\theta} + 2K\|A\|_0,$$

which assures that u_A is well defined.

The application u_A is θ-Hölder. Indeed, fix $\mathbf{x}, \bar{\mathbf{x}} \in \Sigma$ with $d(\mathbf{x}, \bar{\mathbf{x}}) \leq \lambda$ and consider once more points $(\mathbf{y}^0, \mathbf{x}^0), \ldots, (\mathbf{y}^{k-1}, \mathbf{x}^{k-1}) \in \hat{\Sigma}$ satisfying $\mathbf{x}^0 = \mathbf{x}$ and $\mathbf{x}^{j+1} = \tau_{\mathbf{y}^j}(\mathbf{x}^j)$. Putting $\bar{\mathbf{x}}^0 = \bar{\mathbf{x}}$ and $\bar{\mathbf{x}}^{j+1} = \tau_{\mathbf{y}^j}(\bar{\mathbf{x}}^j)$, we obtain

$$\left| \sum_{j=0}^{k-1} A(\mathbf{y}^j, \mathbf{x}^j) - \sum_{j=0}^{k-1} A(\mathbf{y}^j, \bar{\mathbf{x}}^j) \right| \leq \frac{\text{Höld}_\theta(A)}{1-\lambda^\theta} d(\mathbf{x}, \bar{\mathbf{x}})^\theta.$$

As the collection of points $\{(\mathbf{y}^j, \mathbf{x}^j)\}$ was chosen arbitrarily, it follows that

$$|u_A(\mathbf{x}) - u_A(\bar{\mathbf{x}})| \leq \frac{\text{Höld}_\theta(A)}{1-\lambda^\theta} d(\mathbf{x}, \bar{\mathbf{x}})^\theta.$$

To prove the maximal character of u_A, just observe that, for any sub-action $u \in C^0(\Sigma, \mathbb{R}^-)$, we have

$$u(\mathbf{x}) \leq u(\tau_{\mathbf{y}^{k-1}}(\mathbf{x}^{k-1})) - \sum_{j=0}^{k-1} A(\mathbf{y}^j, \mathbf{x}^j) \leq -\sum_{j=0}^{k-1} A(\mathbf{y}^j, \mathbf{x}^j)$$

when $k \geq 0$, $\mathbf{x}^0 = \mathbf{x}$, $\mathbf{y}^j \in \Sigma^*_{\mathbf{x}^j}$ and $\mathbf{x}^{j+1} = \tau_{\mathbf{y}^j}(\mathbf{x}^j)$. □

An interesting question is the existence of a sub-action of minimal character. Given a potential $A \in C^\theta(\hat{\Sigma})$, a possible approach to this demand is to introduce the function $U_A^{K,\theta} \in C^\theta(\Sigma)$ defined by

$$U_A^{K,\theta} = \inf\{u \in C^\theta(\Sigma) : u \text{ sub-action for } A, \text{Höld}_\theta(u) \leq K, \max u = 0\}.$$

The sub-action $U_A^{K,\theta}$ is in some sense minimal.

In the final section, instead of imposing $\max u = 0$, we will consider a suitable normalization of sub-actions in order to present a maximal calibrated one. However, we will need several results before we can discuss this special situation. For instance, the following theorem assures the existence of calibrated sub-actions for any θ-Hölder potential.

THEOREM 4. *Let $\sigma : \Sigma \to \Sigma$ be a transitive subshift of finite type. For each potential $A \in C^\theta(\hat{\Sigma})$, there exists a function $u \in C^\theta(\Sigma)$ such that*

$$u(\mathbf{x}) = \min_{\mathbf{y}\in\Sigma^*_{\mathbf{x}}}[u(\tau_{\mathbf{y}}(\mathbf{x})) - A(\mathbf{y}, \mathbf{x}) + \beta_A].$$

Proof. The idea is to obtain a fixed point of a weak contraction as a limit of fixed points of strong contractions (see [**3, 4**]).

Given $\rho \in (0, 1]$, we define the transformation $\mathcal{L}_\rho : C^0(\Sigma) \to C^0(\Sigma)$ by

$$\mathcal{L}_\rho(f)(\mathbf{x}) = \rho \min_{\mathbf{y}\in\Sigma^*_{\mathbf{x}}}[f(\tau_{\mathbf{y}}(\mathbf{x})) - A(\mathbf{y}, \mathbf{x})].$$

Once $\mathcal{L}_\rho$ is ρ-Lipschitz, consider, when $0 < \rho < 1$, its fixed point $u_\rho \in C^0(\Sigma)$.

The first fact to be noted is the equicontinuity of the family $\{u_\rho\}$. Indeed, note that $\Sigma^*_{\mathbf{x}^0} = \Sigma^*_{\bar{\mathbf{x}}^0}$ when $d(\mathbf{x}^0, \bar{\mathbf{x}}^0) \le \lambda$. Hence, if $\mathbf{y}^0 \in \Sigma^*_{\mathbf{x}^0}$ satisfies

$$u_\rho(\mathbf{x}^0) = \rho[u_\rho(\tau_{\mathbf{y}^0}(\mathbf{x}^0)) - A(\mathbf{y}^0, \mathbf{x}^0)],$$

we obtain

$$u_\rho(\bar{\mathbf{x}}^0) \le \rho[u_\rho(\tau_{\mathbf{y}^0}(\bar{\mathbf{x}}^0)) - A(\mathbf{y}^0, \bar{\mathbf{x}}^0)].$$

Therefore, taking $\mathbf{x}^1 = \tau_{\mathbf{y}^0}(\mathbf{x}^0)$ and $\bar{\mathbf{x}}^1 = \tau_{\mathbf{y}^0}(\bar{\mathbf{x}}^0)$, we have the inequality

$$u_\rho(\bar{\mathbf{x}}^0) - u_\rho(\mathbf{x}^0) \le \rho[A(\mathbf{y}^0, \mathbf{x}^0) - A(\mathbf{y}^0, \bar{\mathbf{x}}^0)] + \rho[u_\rho(\bar{\mathbf{x}}^1) - u_\rho(\mathbf{x}^1)].$$

In this way, defining $\mathbf{x}^j = \tau_{\mathbf{y}^{j-1}}(\mathbf{x}^{j-1})$ and $\bar{\mathbf{x}}^j = \tau_{\mathbf{y}^{j-1}}(\bar{\mathbf{x}}^{j-1})$, we continue inductively obtaining $\mathbf{y}^j \in \Sigma^*_{\mathbf{x}^j}$ such that $u_\rho(\mathbf{x}^j) = \rho[u_\rho(\tau_{\mathbf{y}^j}(\mathbf{x}^j)) - A(\mathbf{y}^j, \mathbf{x}^j)]$. As a consequence of this construction, it follows that

$$u_\rho(\bar{\mathbf{x}}^0) - u_\rho(\mathbf{x}^0) \le \sum_{j=0}^{k-1} \rho^{j+1}[A(\mathbf{y}^j, \mathbf{x}^j) - A(\mathbf{y}^j, \bar{\mathbf{x}}^j)] + \rho^k[u_\rho(\bar{\mathbf{x}}^k) - u_\rho(\mathbf{x}^k)].$$

Thus, we verify

$$\begin{aligned} u_\rho(\bar{\mathbf{x}}^0) - u_\rho(\mathbf{x}^0) &\le \sum_{j=0}^{\infty} \rho^{j+1}[A(\mathbf{y}^j, \mathbf{x}^j) - A(\mathbf{y}^j, \bar{\mathbf{x}}^j)] \\ &\le \mathrm{H\ddot{o}ld}_\theta(A) \sum_{j=0}^{\infty} \rho^{j+1} d(\mathbf{x}^j, \bar{\mathbf{x}}^j)^\theta \\ &\le \mathrm{H\ddot{o}ld}_\theta(A) d(\mathbf{x}^0, \bar{\mathbf{x}}^0)^\theta \sum_{j=0}^{\infty} \rho^{j+1}\lambda^{j\theta} \\ &= \frac{\rho \mathrm{H\ddot{o}ld}_\theta(A)}{1 - \rho\lambda^\theta} d(\mathbf{x}^0, \bar{\mathbf{x}}^0)^\theta. \end{aligned}$$

We have proved that the family $\{u_\rho\}$ is uniformly θ-Hölder; in particular, it is an equicontinuous family of functions.

The family $\{u_\rho\}$ presents also uniformly bounded oscillation. Indeed, given a point $(\mathbf{y}, \mathbf{x}) \in \hat{\Sigma}$, note that

$$\begin{aligned} u_\rho(\mathbf{x}) - \min u_\rho &\le \rho[u_\rho(\tau_{\mathbf{y}}(\mathbf{x})) - A(\mathbf{y}, \mathbf{x})] - \min \rho[u_\rho \circ \pi_1 \circ \hat{\sigma}^{-1} - A] \\ &\le \rho[\max A - A(\mathbf{y}, \mathbf{x})] + \rho[u_\rho(\tau_{\mathbf{y}}(\mathbf{x})) - \min u_\rho] \\ &\le \mathrm{H\ddot{o}ld}_\theta(A) + u_\rho(\tau_{\mathbf{y}}(\mathbf{x})) - \min u_\rho. \end{aligned}$$

Since (Σ, σ) is transitive, we can define a finite set $\{(\mathbf{y}^j, k_j)\} \subset \Sigma^* \times \mathbb{N}$ by choosing, for each pair of symbols $s, s' \in \{1, \ldots, r\}$, an allowed word $(y^j_{k_j-1}, \ldots, y^j_0)$ such that $y^j_{k_j-1} = s'$ and the word (y^j_0, s) is allowed. Consequently, given $\mathbf{x} \in \Sigma$ with $x_0 = s$, the inequality

$$u_\rho(\mathbf{x}) - \min u_\rho \leq k_j \text{Höld}_\theta(A) + u_\rho(\tau^{k_j}_{\mathbf{y}^j}(\mathbf{x})) - \min u_\rho$$

assures that

$$\max_{x_0=s,\ \bar{x}_0=s'} [u_\rho(\mathbf{x}) - u_\rho(\bar{\mathbf{x}})] \leq k_j \text{Höld}_\theta(A) + 2\frac{\text{Höld}_\theta(A)}{1-\lambda^\theta}\lambda^\theta.$$

Hence, when $K = \max k_j$, it follows that

$$\max_{\mathbf{x},\bar{\mathbf{x}}\in\Sigma} [u_\rho(\mathbf{x}) - u_\rho(\bar{\mathbf{x}})] \leq \left(K + \frac{2\lambda^\theta}{1-\lambda^\theta}\right)\text{Höld}_\theta(A),$$

that is, the family $\{u_\rho\}$ has uniformly bounded oscillation.

From the properties demonstrated above, we immediately obtain that the family $\{u_\rho - \max u_\rho\}$ is equicontinuous and uniformly bounded. Note also that $u_\rho - \max u_\rho = (\rho - 1)\max u_\rho + \mathcal{L}_\rho(u_\rho - \max u_\rho)$. Then, if the function u (necessarily θ-Hölder) is an accumulation point of $\{u_\rho - \max u_\rho\}$ when ρ tends to 1, we have $u = a + \mathcal{L}_1(u)$ for some constant $a \in \mathbb{R}$.

It remains to show that $a = \beta_A$. Put $\widetilde{A} = A + u \circ \pi_1 - u \circ \pi_1 \circ \hat{\sigma}^{-1}$. Since $\widetilde{A} \leq a$, for all $\hat{\mu} \in \mathcal{M}_0$, we verify

$$\int_{\hat{\Sigma}} A(\mathbf{y}, \mathbf{x})\, d\hat{\mu}(\mathbf{y}, \mathbf{x}) = \int_{\hat{\Sigma}} \widetilde{A}(\mathbf{y}, \mathbf{x})\, d\hat{\mu}(\mathbf{y}, \mathbf{x}) \leq a,$$

hence $\beta_A \leq a$. Besides, observe that

$$a = \max_{\mathbf{y}\in\Sigma^*_{\mathbf{x}}} \widetilde{A}(\mathbf{y}, \mathbf{x}) \quad \text{for all } \mathbf{x} \in \Sigma.$$

Thus, given $\mathbf{x}^0 \in \Sigma$, take $\mathbf{y}^0 \in \Sigma^*_{\mathbf{x}^0}$ satisfying $\widetilde{A}(\mathbf{y}^0, \mathbf{x}^0) = a$. Putting $\mathbf{x}^j = \tau_{\mathbf{y}^{j-1}}(\mathbf{x}^{j-1})$, inductively consider $\mathbf{y}^j \in \Sigma^*_{\mathbf{x}^j}$ such that $\widetilde{A}(\mathbf{y}^j, \mathbf{x}^j) = a$. Let $\hat{\mu} \in \mathcal{M}$ be an accumulation point of the sequence of probabilities

$$\hat{\mu}_k = \frac{1}{k}\sum_{j=0}^{k-1} \delta_{(\mathbf{y}^j, \mathbf{x}^j)}.$$

Clearly it is true that $\int_{\hat{\Sigma}} \widetilde{A}(\mathbf{y}, \mathbf{x})\, d\hat{\mu}(\mathbf{y}, \mathbf{x}) = a$. Therefore, if we prove that $\hat{\mu} \in \mathcal{M}_0$, we will obtain $a \leq \beta_A$. For any $f \in C^0(\Sigma)$, note then that

$$\begin{aligned}\left|\int_{\hat{\Sigma}} [f(\tau_{\mathbf{y}}(\mathbf{x})) - f(\mathbf{x})]\, d\hat{\mu}_k(\mathbf{y}, \mathbf{x})\right| &= \frac{1}{k}\left|\sum_{j=0}^{k-1}[f(\tau_{\mathbf{y}^j}(\mathbf{x}^j)) - f(\mathbf{x}^j)]\right| \\ &= \frac{1}{k}|f(\mathbf{x}^k) - f(\mathbf{x}^0)| \leq \frac{2}{k}\|f\|_0.\end{aligned}$$

Now taking the limit when k tends to infinity, we assure $\hat{\mu} \in \mathcal{M}_0$ and this finishes the proof. □

The previous result implies the existence of a calibrated sub-action u for the forward shift setting [**3, 9, 17**]. Indeed, supposing $A \in C^\theta(\Sigma)$, observe that we have $A \circ \tau \in C^\theta(\hat{\Sigma})$. Hence, under the transitivity hypothesis, there exists a function $u \in C^\theta(\Sigma)$ satisfying

$$u(\mathbf{x}) = \min_{\mathbf{y} \in \Sigma^*_{\mathbf{x}}} [u(\tau_{\mathbf{y}}(\mathbf{x})) - A \circ \tau(\mathbf{y}, \mathbf{x}) + \beta_{A \circ \tau}].$$

Once $\beta_{A \circ \tau} = \beta_A = \max_{\mu \in \mathcal{M}_\sigma} \int_\Sigma A(\mathbf{x})\, d\mu(\mathbf{x})$, taking $\mathbf{z} = \tau_{\mathbf{y}}(\mathbf{x})$, we obtain the usual expression (see for instance [**9**])

$$u(\mathbf{x}) = \min_{\sigma(\mathbf{z}) = \mathbf{x}} (u - A + \beta_A)(\mathbf{z}).$$

The notion of calibrated sub-action is an important concept also in relative maximization. In particular, Theorem 4 assures a version for the holonomic setting of [**13,** Theorem 17]. Such a version will point out that the differential of an alpha application dictates the asymptotic behavior of the optimal trajectories. We will state the precise result.

We start by considering the Fenchel transform of the previous beta function $\beta_{A,\varphi}$. Called an alpha application, such a function $\alpha_{A,\varphi} : \mathbb{R}^n \to \mathbb{R}$ is defined simply by

$$\alpha_{A,\varphi}(c) = \min_{h \in \varphi_*(\mathcal{M}_0)} [\langle c, h \rangle - \beta_{A,\varphi}(h)].$$

If $u \in C^0(\Sigma)$ is a calibrated sub-action, we say that a sequence $\{\mathbf{y}^j, \mathbf{x}^j\} \subset \hat{\Sigma}$ is an optimal trajectory (associated to the potential A) in the case $\mathbf{x}^j = \tau_{\mathbf{y}^{j-1}}(\mathbf{x}^{j-1})$ and $u(\mathbf{x}^j) = u(\mathbf{x}^{j+1}) - A(\mathbf{y}^j, \mathbf{x}^j) + \beta_A$. Since the equality $\alpha_{A,\varphi}(c) = -\beta_{A - \langle c, \varphi \rangle}$ is true, we can adapt [**13,** Proof of Theorem 17] to the present case. Therefore, under the transitivity hypothesis, if the potential A and the constraint φ are Hölder, every optimal trajectory $\{\mathbf{y}^j, \mathbf{x}^j\}$ associated to $A - \langle c, \varphi \rangle$ satisfies

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

in the case when the function $\alpha_{A,\varphi}$ is differentiable at the point $c \in \mathbb{R}^n$.

Concluding this section, we would like to say a few words about a version of Livšic's theorem for the model $(\hat{\Sigma}, \hat{\sigma}, \mathcal{M}_0)$. We will say that a function $A \in C^0(\hat{\Sigma})$ is cohomologous to a constant $a \in \mathbb{R}$ if there exists a function $u \in C^0(\Sigma)$ such that

$$A + u \circ \pi_1 - u \circ \pi_1 \circ \hat{\sigma}^{-1} = a.$$

PROPOSITION 5. *Assume that $\sigma : \Sigma \to \Sigma$ is a transitive subshift of finite type and suppose that A is a θ-Hölder function. Then, $m_A = \mathcal{M}_0$ if, and only if, A is cohomologous to β_A.*

Proof. The sufficiency is obvious. Reciprocally, as $m_A = \mathcal{M}_0$ implies $\beta_A = -\beta_{-A}$, consider functions $u, u' \in C^0(\Sigma)$ satisfying

$$A + u \circ \pi_1 - u \circ \pi_1 \circ \hat{\sigma}^{-1} \le \beta_A \quad \text{and} \quad \beta_A \le A - u' \circ \pi_1 + u' \circ \pi_1 \circ \hat{\sigma}^{-1}.$$

Therefore, we have $(u + u') \circ \pi_1 \le (u + u') \circ \pi_1 \circ \hat{\sigma}^{-1}$. In this case, however, the transitivity hypothesis implies that the function $u + u'$ is identically equal to a constant b. Since $u = b - u'$, from the above two inequalities, it follows that the potential A is cohomologous to β_A via the function u. □

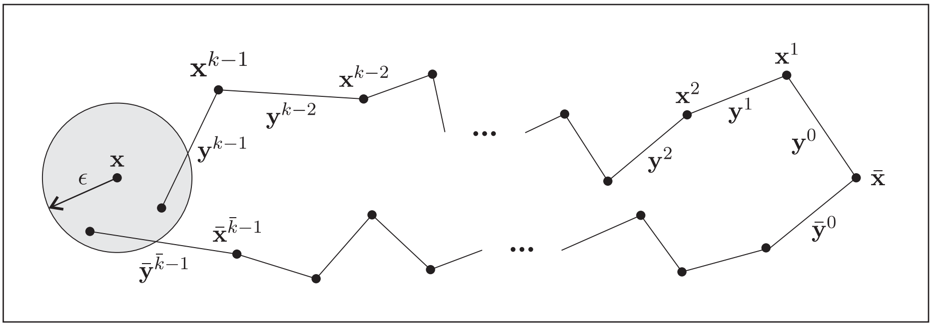

FIGURE 1. Graphical representation of paths belonging to $\mathcal{P}(\mathbf{x}, \bar{\mathbf{x}}, \epsilon)$.

4. *Calibrated sub-actions and Mañé potential*

Using the Mañé potential and the set of non-wandering points, we will be able to introduce a family of Hölder calibrated sub-actions. In the final section, this family will play a crucial role in the classification theorem of calibrated sub-actions.

Definition 6. Given $\epsilon > 0$ and $\mathbf{x}, \bar{\mathbf{x}} \in \Sigma$, we will call a path beginning within ϵ of $\mathbf{x}$ and ending at $\bar{\mathbf{x}}$ an ordered sequence of points

$$(\mathbf{y}^0, \mathbf{x}^0), \ldots, (\mathbf{y}^{k-1}, \mathbf{x}^{k-1}) \in \hat{\Sigma}$$

satisfying $\mathbf{x}^0 = \bar{\mathbf{x}}$, $\mathbf{x}^{j+1} = \tau_{\mathbf{y}^j}(\mathbf{x}^j)$ and $d(\tau_{\mathbf{y}^{k-1}}(\mathbf{x}^{k-1}), \mathbf{x}) < \epsilon$. We will denote by $\mathcal{P}(\mathbf{x}, \bar{\mathbf{x}}, \epsilon)$ the set of such paths (see Figure 1).

Definition 7. Following [**9**], a point $\mathbf{x} \in \Sigma$ will be called non-wandering with respect to the potential $A \in C^0(\hat{\Sigma})$ when, for all $\epsilon > 0$, we can determine a path $\{(\mathbf{y}^0, \mathbf{x}^0), \ldots, (\mathbf{y}^{k-1}, \mathbf{x}^{k-1})\} \in \mathcal{P}(\mathbf{x}, \mathbf{x}, \epsilon)$ such that

$$\left| \sum_{j=0}^{k-1} (A - \beta_A)(\mathbf{y}^j, \mathbf{x}^j) \right| < \epsilon.$$

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

When the potential is Hölder, it is not difficult to see that $\Omega(A)$ is a compact invariant set. We will show that such a set is indeed not empty.

LEMMA 6. *If $\sigma : \Sigma \to \Sigma$ is a transitive subshift of finite type, for any potential $A \in C^\theta(\hat{\Sigma})$, we have $\Omega(A) \neq \emptyset$.*

Proof. Let $u \in C^0(\Sigma)$ be a calibrated sub-action obtained from Theorem 4. Fix any point $\mathbf{x}^0 \in \Sigma$. Take then $\mathbf{y}^0 \in \Sigma^*_{\mathbf{x}^0}$ satisfying the identity $u(\mathbf{x}^0) = u(\tau_{\mathbf{y}^0}(\mathbf{x}^0)) - A(\mathbf{y}^0, \mathbf{x}^0) + \beta_A$. Denote $\mathbf{x}^{j+1} = \tau_{\mathbf{y}^j}(\mathbf{x}^j)$ and proceed in an inductive way determining a point $\mathbf{y}^{j+1} \in \Sigma^*_{\mathbf{x}^{j+1}}$ such that $u(\mathbf{x}^{j+1}) = u(\tau_{\mathbf{y}^{j+1}}(\mathbf{x}^{j+1})) - A(\mathbf{y}^{j+1}, \mathbf{x}^{j+1}) + \beta_A$. Let $\mathbf{x} \in \Sigma$ be a limit of some subsequence $\{\mathbf{x}^{j_m}\}$.

We claim that $\mathbf{x} \in \Omega(A)$. First note that, if $m_2 > m_1$, from the definition of the sequence $\{\mathbf{x}^j\}$, we obtain

$$-\sum_{j=j_{m_1}}^{j_{m_2}-1}(A-\beta_A)(\mathbf{y}^j,\mathbf{x}^j) = u(\mathbf{x}^{j_{m_1}}) - u(\mathbf{x}^{j_{m_2}}).$$

For a fixed $\epsilon > 0$, consider an integer $l > 0$ such that, if $\mathbf{x}', \mathbf{x}'' \in \Sigma$ and $d(\mathbf{x}', \mathbf{x}'') < \lambda^l$, then $|u(\mathbf{x}') - u(\mathbf{x}'')| < \epsilon/2$. We can suppose that l is sufficiently large in such a way that

$$\max\left\{\lambda^l, \frac{\text{Höld}_\theta(A)}{1-\lambda^\theta}\lambda^{\theta l}\right\} < \frac{\epsilon}{2}.$$

Now take an integer m_0 sufficiently large such that $d(\mathbf{x}^{j_m}, \mathbf{x}) < \lambda^l/2$ for all $m > m_0$. Considering integers $m_2 > m_1 > m_0$, put $k = j_{m_2} - j_{m_1}$. Since $\Sigma^*_{\mathbf{x}} = \Sigma^*_{\mathbf{x}^{j_{m_1}}}$, we choose $\bar{\mathbf{y}}^j = \mathbf{y}^{j_{m_1}+j}$ for $0 \le j \le k-1$. Finally, denote $\bar{\mathbf{x}}^0 = \mathbf{x}$ and $\bar{\mathbf{x}}^{j+1} = \tau_{\bar{\mathbf{y}}^j}(\bar{\mathbf{x}}^j)$. Once

$$d(\tau_{\bar{\mathbf{y}}^{k-1}}(\bar{\mathbf{x}}^{k-1}), \mathbf{x}) \le d(\tau_{\bar{\mathbf{y}}^{k-1}}(\bar{\mathbf{x}}^{k-1}), \mathbf{x}^{j_{m_2}}) + d(\mathbf{x}^{j_{m_2}}, \mathbf{x}) < \lambda^{k+l} + \lambda^l < \epsilon,$$

it follows that $\{(\bar{\mathbf{y}}^0, \bar{\mathbf{x}}^0), \ldots, (\bar{\mathbf{y}}^{k-1}, \bar{\mathbf{x}}^{k-1})\} \in \mathcal{P}(\mathbf{x}, \mathbf{x}, \epsilon)$. Moreover, since $d(\mathbf{x}^{j_{m_1}}, \mathbf{x}^{j_{m_2}}) < \lambda^l$, we get

$$\begin{aligned}\left|\sum_{j=0}^{k-1}(A-\beta_A)(\bar{\mathbf{y}}^j,\bar{\mathbf{x}}^j)\right| &\le \left|\sum_{j=0}^{k-1}A(\bar{\mathbf{y}}^j,\bar{\mathbf{x}}^j) - \sum_{j=j_{m_1}}^{j_{m_2}-1}A(\mathbf{y}^j,\mathbf{x}^j)\right| + |u(\mathbf{x}^{j_{m_1}}) - u(\mathbf{x}^{j_{m_2}})| \\ &< \frac{\text{Höld}_\theta(A)}{1-\lambda^\theta}\lambda^{\theta l} + \frac{\epsilon}{2} < \epsilon.\end{aligned}$$

Therefore, $\mathbf{x} \in \Omega(A)$. □

The following definition is also inspired by [**9**].

Definition 8. We call 'Mañé potential' the function $S_A : \Sigma \times \Sigma \to \mathbb{R} \cup \{\pm\infty\}$ defined by

$$S_A(\mathbf{x}, \bar{\mathbf{x}}) = \lim_{\epsilon \to 0} S_A^\epsilon(\mathbf{x}, \bar{\mathbf{x}}),$$

where

$$S_A^\epsilon(\mathbf{x}, \bar{\mathbf{x}}) = \inf_{\{(\mathbf{y}^0,\mathbf{x}^0),\ldots,(\mathbf{y}^{k-1},\mathbf{x}^{k-1})\}\in\mathcal{P}(\mathbf{x},\bar{\mathbf{x}},\epsilon)}\left[-\sum_{j=0}^{k-1}(A-\beta_A)(\mathbf{y}^j,\mathbf{x}^j)\right].$$

Note that $\Omega(A) = \{\mathbf{x} \in \Sigma : S_A(\mathbf{x}, \mathbf{x}) = 0\}$.

As we will see soon the Mañé potential will provide, for a Hölder potential, a one-parameter family of equally Hölder sub-actions. Before that we need some properties.

Let $u \in C^0(\Sigma)$ be a sub-action for the potential $A \in C^0(\hat{\Sigma})$. We say that the point $\mathbf{x} \in \Sigma$ is u-connected to the point $\bar{\mathbf{x}} \in \Sigma$, and we indicate this by $\mathbf{x} \overset{u}{\to} \bar{\mathbf{x}}$, when, for every $\epsilon > 0$, we can determine a path $\{(\mathbf{y}^0, \mathbf{x}^0), \ldots, (\mathbf{y}^{k-1}, \mathbf{x}^{k-1})\} \in \mathcal{P}(\mathbf{x}, \bar{\mathbf{x}}, \epsilon)$ such that

$$\left|\sum_{j=0}^{k-1}(A-\beta_A)(\mathbf{y}^j,\mathbf{x}^j) - (u(\mathbf{x}) - u(\bar{\mathbf{x}}))\right| < \epsilon.$$

Note that $\mathbf{x} \in \Omega(A)$ implies $\mathbf{x} \overset{u}{\to} \mathbf{x}$ for any sub-action u.

LEMMA 7. *Let $u \in C^0(\Sigma)$ be a sub-action for a potential $A \in C^0(\hat{\Sigma})$. Then, for any $\mathbf{x}, \bar{\mathbf{x}} \in \Sigma$, we have $S_A(\mathbf{x}, \bar{\mathbf{x}}) \geq u(\bar{\mathbf{x}}) - u(\mathbf{x})$. Moreover, the equality is true if, and only if, $\mathbf{x} \stackrel{u}{\to} \bar{\mathbf{x}}$.*

Before the proof of this lemma, we would like just to point out another important property of the Mañé potential: if A is a θ-Hölder potential, then $S_A(\mathbf{x}, \bar{\bar{\mathbf{x}}}) \leq S_A(\mathbf{x}, \bar{\mathbf{x}}) + S_A(\bar{\mathbf{x}}, \bar{\bar{\mathbf{x}}})$ for any points $\mathbf{x}, \bar{\mathbf{x}}, \bar{\bar{\mathbf{x}}} \in \Sigma$. We leave for the reader the demonstration of this simple fact.

Proof. Fix $\rho > 0$. Take $\epsilon \in (0, \rho)$ such that $|u(\mathbf{x}') - u(\mathbf{x}'')| < \rho$, when $\mathbf{x}', \mathbf{x}'' \in \Sigma$ satisfy $d(\mathbf{x}', \mathbf{x}'') < \epsilon$. Consider now any path

$$\{(\mathbf{y}^0, \mathbf{x}^0), \ldots, (\mathbf{y}^{k-1}, \mathbf{x}^{k-1})\} \in \mathcal{P}(\mathbf{x}, \bar{\mathbf{x}}, \epsilon).$$

Once

$$u(\bar{\mathbf{x}}) - u(\mathbf{x}) - \rho < u(\mathbf{x}^0) - u(\tau_{\mathbf{y}^{k-1}}(\mathbf{x}^{k-1})) \leq -\sum_{j=0}^{k-1}(A - \beta_A)(\mathbf{y}^j, \mathbf{x}^j),$$

it follows that $u(\bar{\mathbf{x}}) - u(\mathbf{x}) - \rho \leq S_A(\mathbf{x}, \bar{\mathbf{x}})$. Taking ρ arbitrarily small, we obtain the inequality of the lemma.

If $S_A(\mathbf{x}, \bar{\mathbf{x}}) = u(\bar{\mathbf{x}}) - u(\mathbf{x})$, from the definition of the Mañé potential, immediately we get $\mathbf{x} \stackrel{u}{\to} \bar{\mathbf{x}}$. Reciprocally, suppose that $\mathbf{x}$ is u-connected to $\bar{\mathbf{x}}$. Take then $\rho > 0$. Given $\epsilon \in (0, \rho)$, we can choose a path

$$\{(\mathbf{y}^0, \mathbf{x}^0), \ldots, (\mathbf{y}^{k-1}, \mathbf{x}^{k-1})\} \in \mathcal{P}(\mathbf{x}, \bar{\mathbf{x}}, \epsilon)$$

satisfying

$$\left|\sum_{j=0}^{k-1}(A - \beta_A)(\mathbf{y}^j, \mathbf{x}^j) - (u(\mathbf{x}) - u(\bar{\mathbf{x}}))\right| < \epsilon.$$

Observe that

$$-\sum_{j=0}^{k-1}(A - \beta_A)(\mathbf{y}^j, \mathbf{x}^j) < u(\bar{\mathbf{x}}) - u(\mathbf{x}) + \epsilon < u(\bar{\mathbf{x}}) - u(\mathbf{x}) + \rho.$$

Thus, we verify $S_A(\mathbf{x}, \bar{\mathbf{x}}) \leq u(\bar{\mathbf{x}}) - u(\mathbf{x}) + \rho$. As ρ can be taken arbitrarily small, we finally get the equality claimed by the lemma. □

We now present the main result of this section.

PROPOSITION 8. *Suppose that $\sigma : \Sigma \to \Sigma$ is a transitive subshift of finite type. Let A be a θ-Hölder potential. Then, for each $\mathbf{x} \in \Omega(A)$, the function $S_A(\mathbf{x}, \cdot)$ is a θ-Hölder calibrated sub-action.*

Proof. Fix a point $\mathbf{x} \in \Omega(A)$. We must show first that $S_A(\mathbf{x}, \cdot)$ is a well-defined real function. Thanks to Lemma 7, we only need to assure that $S_A(\mathbf{x}, \bar{\mathbf{x}}) < +\infty$ for any $\bar{\mathbf{x}} \in \Sigma$.

Take $\epsilon > 0$ arbitrary. For a fixed value $\epsilon' \in (0, \lambda]$, consider a path $\{(\mathbf{y}^0, \mathbf{x}^0), \ldots, (\mathbf{y}^{k-1}, \mathbf{x}^{k-1})\} \in \mathcal{P}(\mathbf{x}, \bar{\mathbf{x}}, \epsilon')$ satisfying

$$-\sum_{j=0}^{k-1}(A - \beta_A)(\mathbf{y}^j, \mathbf{x}^j) < S_A^{\epsilon'}(\mathbf{x}, \bar{\mathbf{x}}) + \epsilon.$$

As $\mathbf{x} \in \Omega(A)$, we can take $\{(\bar{\mathbf{y}}^0, \bar{\mathbf{x}}^0), \dots, (\bar{\mathbf{y}}^{\bar{k}-1}, \bar{\mathbf{x}}^{\bar{k}-1})\} \in \mathcal{P}(\mathbf{x}, \mathbf{x}, \epsilon/2)$, with $\lambda^{\bar{k}}\epsilon' < \epsilon/2$, such that

$$\left|\sum_{j=0}^{\bar{k}-1}(A - \beta_A)(\bar{\mathbf{y}}^j, \bar{\mathbf{x}}^j)\right| < \frac{\epsilon}{2}.$$

Thus, we define $\mathbf{y}^j = \bar{\mathbf{y}}^{j-k}$ for $k \le j < k + \bar{k}$. Observe that we have $\mathbf{y}^k = \bar{\mathbf{y}}^0 \in \Sigma^*_{\bar{\mathbf{x}}^0} = \Sigma^*_{\tau_{\mathbf{y}^{k-1}}(\mathbf{x}^{k-1})}$. Therefore, we can put $\mathbf{x}^{j+1} = \tau_{\mathbf{y}^j}(\mathbf{x}^j)$ for $k - 1 \le j < k + \bar{k} - 1$.

We claim that $\{(\mathbf{y}^0, \mathbf{x}^0), \dots, (\mathbf{y}^{k+\bar{k}-1}, \mathbf{x}^{k+\bar{k}-1})\} \in \mathcal{P}(\mathbf{x}, \bar{\mathbf{x}}, \epsilon)$. Indeed, we have

$$\begin{aligned} d(\tau_{\mathbf{y}^{k+\bar{k}-1}}(\mathbf{x}^{k+\bar{k}-1}), \mathbf{x}) &\le d(\tau_{\mathbf{y}^{k+\bar{k}-1}}(\mathbf{x}^{k+\bar{k}-1}), \tau_{\bar{\mathbf{y}}^{\bar{k}-1}}(\bar{\mathbf{x}}^{\bar{k}-1})) + d(\tau_{\bar{\mathbf{y}}^{\bar{k}-1}}(\bar{\mathbf{x}}^{\bar{k}-1}), \mathbf{x}) \\ &< \lambda^{\bar{k}}\epsilon' + \frac{\epsilon}{2} < \epsilon. \end{aligned}$$

Besides, without difficulty we verify

$$\left|\sum_{j=k}^{k+\bar{k}-1} A(\mathbf{y}^j, \mathbf{x}^j) - \sum_{j=0}^{\bar{k}-1} A(\bar{\mathbf{y}}^j, \bar{\mathbf{x}}^j)\right| \le \frac{\mathrm{Höld}_\theta(A)}{1 - \lambda^\theta}(\epsilon')^\theta.$$

Hence, we immediately have

$$S^\epsilon_A(\mathbf{x}, \bar{\mathbf{x}}) \le -\sum_{j=0}^{k+\bar{k}-1}(A - \beta_A)(\mathbf{y}^j, \mathbf{x}^j) < \frac{\mathrm{Höld}_\theta(A)}{1 - \lambda^\theta}(\epsilon')^\theta + S^{\epsilon'}_A(\mathbf{x}, \bar{\mathbf{x}}) + \frac{3}{2}\epsilon,$$

which yields

$$S_A(\mathbf{x}, \bar{\mathbf{x}}) \le \frac{\mathrm{Höld}_\theta(A)}{1 - \lambda^\theta}(\epsilon')^\theta + S^{\epsilon'}_A(\mathbf{x}, \bar{\mathbf{x}}).$$

As the right-hand side is finite, the application $S_A(\mathbf{x}, \cdot)$ is well defined.

We claim that it is indeed a θ-Hölder function. Take points $\bar{\mathbf{x}}, \bar{\bar{\mathbf{x}}} \in \Sigma$ such that $d(\bar{\mathbf{x}}, \bar{\bar{\mathbf{x}}}) \le \lambda$. Consider a fixed $\rho > 0$. Given $\epsilon > 0$, we can find a path $\{(\mathbf{y}^0, \mathbf{x}^0), \dots, (\mathbf{y}^{k-1}, \mathbf{x}^{k-1})\} \in \mathcal{P}(\mathbf{x}, \bar{\mathbf{x}}, \epsilon)$, with $\lambda^{k+1} < \epsilon$, such that

$$-\sum_{j=0}^{k-1}(A - \beta_A)(\mathbf{y}^j, \mathbf{x}^j) < S^\epsilon_A(\mathbf{x}, \bar{\mathbf{x}}) + \rho.$$

Taking $\bar{\mathbf{y}}^j = \mathbf{y}^j$ for $0 \le j < k$, we write $\bar{\mathbf{x}}^0 = \bar{\bar{\mathbf{x}}}$ and, finally, we define $\bar{\mathbf{x}}^{j+1} = \tau_{\bar{\mathbf{y}}^j}(\bar{\mathbf{x}}^j)$ when $0 \le j < k - 1$. It is easy to confirm that $\{(\bar{\mathbf{y}}^0, \bar{\mathbf{x}}^0), \dots, (\bar{\mathbf{y}}^{k-1}, \bar{\mathbf{x}}^{k-1})\} \in \mathcal{P}(\mathbf{x}, \bar{\bar{\mathbf{x}}}, 2\epsilon)$, as well as

$$-\sum_{j=0}^{k-1} A(\mathbf{y}^j, \mathbf{x}^j) \ge -\sum_{j=0}^{k-1} A(\bar{\mathbf{y}}^j, \bar{\mathbf{x}}^j) - \frac{\mathrm{Höld}_\theta(A)}{1 - \lambda^\theta} d(\bar{\mathbf{x}}, \bar{\bar{\mathbf{x}}})^\theta.$$

Therefore, we verify the following inequalities:

$$\begin{aligned} S_A(\mathbf{x}, \bar{\mathbf{x}}) &\ge S^\epsilon_A(\mathbf{x}, \bar{\mathbf{x}}) \\ &> -\sum_{j=0}^{k-1}(A - \beta_A)(\mathbf{y}^j, \mathbf{x}^j) - \rho \\ &\ge -\sum_{j=0}^{k-1}(A - \beta_A)(\bar{\mathbf{y}}^j, \bar{\mathbf{x}}^j) - \frac{\mathrm{Höld}_\theta(A)}{1 - \lambda^\theta} d(\bar{\mathbf{x}}, \bar{\bar{\mathbf{x}}})^\theta - \rho \\ &\ge S^{2\epsilon}_A(\mathbf{x}, \bar{\bar{\mathbf{x}}}) - \frac{\mathrm{Höld}_\theta(A)}{1 - \lambda^\theta} d(\bar{\mathbf{x}}, \bar{\bar{\mathbf{x}}})^\theta - \rho. \end{aligned}$$

Since ϵ and ρ can be considered (in such order) arbitrarily small, we get

$$S_A(\mathbf{x}, \bar{\mathbf{x}}) - S_A(\mathbf{x}, \bar{\bar{\mathbf{x}}}) \geq -\frac{\text{Höld}_\theta(A)}{1-\lambda^\theta} d(\bar{\mathbf{x}}, \bar{\bar{\mathbf{x}}})^\theta.$$

It follows at once that $S_A(\mathbf{x}, \cdot) \in C^\theta(\Sigma)$.

It remains to show that the application $S_A(\mathbf{x}, \cdot)$ is a calibrated sub-action.

Fix a point $(\bar{\mathbf{y}}, \bar{\mathbf{x}}) \in \hat{\Sigma}$. When $\{(\mathbf{y}^1, \mathbf{x}^1), \ldots, (\mathbf{y}^k, \mathbf{x}^k)\} \in \mathcal{P}(\mathbf{x}, \tau_{\bar{\mathbf{y}}}(\bar{\mathbf{x}}), \epsilon)$, put $\mathbf{y}^0 = \bar{\mathbf{y}}$, $\mathbf{x}^0 = \bar{\mathbf{x}}$. We point out that

$$\begin{aligned} A(\bar{\mathbf{y}}, \bar{\mathbf{x}}) - \beta_A &= \sum_{j=0}^{k} (A - \beta_A)(\mathbf{y}^j, \mathbf{x}^j) - \sum_{j=0}^{k-1} (A - \beta_A)(\mathbf{y}^{j+1}, \mathbf{x}^{j+1}) \\ &\leq -\sum_{j=0}^{k-1} (A - \beta_A)(\mathbf{y}^{j+1}, \mathbf{x}^{j+1}) - S_A^\epsilon(\mathbf{x}, \bar{\mathbf{x}}). \end{aligned}$$

As the path is arbitrary, we have $A(\bar{\mathbf{y}}, \bar{\mathbf{x}}) - \beta_A \leq S_A^\epsilon(\mathbf{x}, \tau_{\bar{\mathbf{y}}}(\bar{\mathbf{x}})) - S_A^\epsilon(\mathbf{x}, \bar{\mathbf{x}})$. Hence, taking the limit, we show that $S_A(\mathbf{x}, \cdot)$ is indeed a sub-action for the potential A.

In order to verify that it is a calibrated sub-action, we should be able to determine, for each $\bar{\mathbf{x}} \in \Sigma$, a point $\bar{\mathbf{y}} \in \Sigma_{\bar{\mathbf{x}}}^*$ accomplishing the equality $S_A(\mathbf{x}, \bar{\mathbf{x}}) = S_A(\mathbf{x}, \tau_{\bar{\mathbf{y}}}(\bar{\mathbf{x}})) - A(\bar{\mathbf{y}}, \bar{\mathbf{x}}) + \beta_A$. Given $\epsilon > 0$, consider a path $\{(\mathbf{y}^0, \mathbf{x}^0), \ldots, (\mathbf{y}^{k-1}, \mathbf{x}^{k-1})\} \in \mathcal{P}(\mathbf{x}, \bar{\mathbf{x}}, \epsilon)$ such that

$$-\sum_{j=0}^{k-1} (A - \beta_A)(\mathbf{y}^j, \mathbf{x}^j) < S_A^\epsilon(\mathbf{x}, \bar{\mathbf{x}}) + \epsilon.$$

This defines a family $\{\mathbf{y}^0\}_{\epsilon>0} \subset \Sigma_{\bar{\mathbf{x}}}^*$. Take $\bar{\mathbf{y}} \in \Sigma_{\bar{\mathbf{x}}}^*$ an accumulation point of this family when ϵ tends to zero. Observe that

$$S_A^\epsilon(\mathbf{x}, \tau_{\mathbf{y}^0}(\bar{\mathbf{x}})) - (A - \beta_A)(\mathbf{y}^0, \bar{\mathbf{x}}) \leq -\sum_{j=0}^{k-1} (A - \beta_A)(\mathbf{y}^j, \mathbf{x}^j).$$

As $\tau_{\mathbf{y}^0}(\bar{\mathbf{x}}) = \tau_{\bar{\mathbf{y}}}(\bar{\mathbf{x}})$ for ϵ sufficiently small, we can focus on

$$S_A^\epsilon(\mathbf{x}, \tau_{\bar{\mathbf{y}}}(\bar{\mathbf{x}})) - (A - \beta_A)(\mathbf{y}^0, \bar{\mathbf{x}}) < S_A^\epsilon(\mathbf{x}, \bar{\mathbf{x}}) + \epsilon.$$

So taking ϵ arbitrarily small, we finish the proof. □

5. *Sub-actions and supports*

This section is dedicated to the analysis of relationships between sub-actions and supports of holonomic probabilities. A unifying element of these concepts continues to be the notion of contact locus.

Definition 9. Given a sub-action $u \in C^0(\Sigma)$ for a potential $A \in C^0(\hat{\Sigma})$, consider the function $A^u = A + u \circ \pi_1 - u \circ \pi_1 \circ \hat{\sigma}^{-1}$. We call the set $\mathbb{M}_A(u) = (A^u)^{-1}(\beta_A)$ the contact locus of the sub-action u.

The contact locus is just the set where the usual inequality defining a sub-action becomes an equality. It plays an important role in the localization of the support of maximizing holonomic probabilities.

PROPOSITION 9. *If $u \in C^0(\Sigma)$ is a sub-action for a potential $A \in C^0(\hat{\Sigma})$, then*

$$m_A = \{\hat{\mu} \in \mathcal{M}_0 : \operatorname{supp}(\hat{\mu}) \subset \mathbb{M}_A(u)\}.$$

The proof of this statement is reduced to the well-known fact that, if its integral is zero, a measurable non-negative function is zero almost everywhere.

We now require a classification theorem for calibrated sub-actions. We start by presenting a result which supplies a representation formula for these sub-actions.

THEOREM 10. *If $u \in C^0(\Sigma)$ is a calibrated sub-action for a θ-Hölder potential A, then*

$$u(\bar{\mathbf{x}}) = \inf_{\mathbf{x} \in \Omega(A)} [u(\mathbf{x}) + S_A(\mathbf{x}, \bar{\mathbf{x}})].$$

Proof. Thanks to Lemma 7, it immediately follows that

$$u(\bar{\mathbf{x}}) \leq \inf_{\mathbf{x} \in \Omega(A)} [u(\mathbf{x}) + S_A(\mathbf{x}, \bar{\mathbf{x}})].$$

Besides, the identity will be true if there exists a point $\mathbf{x} \in \Omega(A)$ satisfying $\mathbf{x} \xrightarrow{u} \bar{\mathbf{x}}$.

Consider $\{(\mathbf{y}^j, \mathbf{x}^j)\} \subset \hat{\Sigma}$ an optimal trajectory associated to the potential A such that $\mathbf{x}^0 = \bar{\mathbf{x}}$. Denote by $\mathbf{x} \in \Sigma$ the limit of a subsequence $\{\mathbf{x}^{j_m}\}$.

Lemma 6 shows that $\mathbf{x} \in \Omega(A)$. So we only have to prove that $\mathbf{x} \xrightarrow{u} \bar{\mathbf{x}}$. Fix $\epsilon > 0$ and choose an integer $l > 0$ in such a way that $|u(\mathbf{x}') - u(\mathbf{x}'')| < \epsilon$ when $\mathbf{x}', \mathbf{x}'' \in \Sigma$ satisfy $d(\mathbf{x}', \mathbf{x}'') < \lambda^l$. Assume that l also accomplishes $\lambda^l < \epsilon$. Take m sufficiently large such that $d(\mathbf{x}^{j_m}, \mathbf{x}) < \lambda^l$. Put $k = j_m$.

Observe that $d(\tau_{\mathbf{y}^{k-1}}(\mathbf{x}^{k-1}), \mathbf{x}) = d(\mathbf{x}^{j_m}, \mathbf{x}) < \epsilon$. Therefore, we assure $\{(\mathbf{y}^0, \mathbf{x}^0), \ldots, (\mathbf{y}^{k-1}, \mathbf{x}^{k-1})\} \in \mathcal{P}(\mathbf{x}, \bar{\mathbf{x}}, \epsilon)$. As

$$\sum_{j=0}^{k-1} (A - \beta_A)(\mathbf{y}^j, \mathbf{x}^j) - (u(\mathbf{x}^k) - u(\bar{\mathbf{x}})) = 0,$$

we obtain

$$\left| \sum_{j=0}^{k-1} (A - \beta_A)(\mathbf{y}^j, \mathbf{x}^j) - (u(\mathbf{x}) - u(\bar{\mathbf{x}})) \right| = |u(\mathbf{x}^{j_m}) - u(\mathbf{x})| < \epsilon,$$

which finishes the proof. □

The following immediate corollary indicates the importance of the set $\Omega(A)$ in the analysis of calibrated sub-actions.

COROLLARY 11. *Let $u, u' \in C^0(\Sigma)$ be calibrated sub-actions for a potential $A \in C^\theta(\hat{\Sigma})$. If $u \leq u'$ on $\Omega(A)$, then $u \leq u'$ everywhere on Σ. In particular, if we have $u|_{\Omega(A)} = u'|_{\Omega(A)}$, then both sub-actions are equal.*

Theorem 10 admits a reciprocal.

THEOREM 12. *Let $\sigma : \Sigma \to \Sigma$ be a transitive subshift of finite type. Consider a potential $A \in C^\theta(\hat{\Sigma})$. Assume that the function $f : \Omega(A) \to \mathbb{R}$ has a finite lower bound. Then*

$$u(\bar{\mathbf{x}}) = \inf_{\mathbf{x} \in \Omega(A)} [f(\mathbf{x}) + S_A(\mathbf{x}, \bar{\mathbf{x}})]$$

defines a θ-Hölder calibrated sub-action. Moreover, if $f(\bar{\mathbf{x}}) - f(\mathbf{x}) \leq S_A(\mathbf{x}, \bar{\mathbf{x}})$ for any $\mathbf{x}, \bar{\mathbf{x}} \in \Omega(A)$, then $u = f$ on $\Omega(A)$.

Proof. The good definition of $u : \Sigma \to \mathbb{R}$ is clear. We will show that it is a Hölder function. Fix $\epsilon > 0$. Given $\bar{\mathbf{x}}, \bar{\bar{\mathbf{x}}} \in \Sigma$ with $d(\bar{\mathbf{x}}, \bar{\bar{\mathbf{x}}}) \le \lambda$, take a point $\mathbf{x} \in \Omega(A)$ such that $f(\mathbf{x}) + S_A(\mathbf{x}, \bar{\bar{\mathbf{x}}}) < u(\bar{\bar{\mathbf{x}}}) + \epsilon$. It follows from the proof of Proposition 8 that

$$u(\bar{\mathbf{x}}) - u(\bar{\bar{\mathbf{x}}}) - \epsilon < S_A(\mathbf{x}, \bar{\mathbf{x}}) - S_A(\mathbf{x}, \bar{\bar{\mathbf{x}}}) \le \frac{\text{Höld}_\theta(A)}{1 - \lambda^\theta} d(\bar{\mathbf{x}}, \bar{\bar{\mathbf{x}}})^\theta.$$

As ϵ is arbitrary, we get $u \in C^\theta(\Sigma)$.

In fact, u is a sub-action for the potential A. Consider a point $(\bar{\mathbf{y}}, \bar{\mathbf{x}}) \in \hat{\Sigma}$ and $\epsilon > 0$. Choose $\mathbf{x} \in \Omega(A)$ satisfying $f(\mathbf{x}) + S_A(\mathbf{x}, \tau_{\bar{\mathbf{y}}}(\bar{\mathbf{x}})) < u(\tau_{\bar{\mathbf{y}}}(\bar{\mathbf{x}})) + \epsilon$. Since

$$u(\bar{\mathbf{x}}) - u(\tau_{\bar{\mathbf{y}}}(\bar{\mathbf{x}})) - \epsilon < S_A(\mathbf{x}, \bar{\mathbf{x}}) - S_A(\mathbf{x}, \tau_{\bar{\mathbf{y}}}(\bar{\mathbf{x}})) \le \beta_A - A(\bar{\mathbf{y}}, \bar{\mathbf{x}}),$$

the claim follows when ϵ tends to zero.

The calibrated character of u is also a consequence of Proposition 8. Indeed, take $\bar{\mathbf{x}} \in \Sigma$, and choose a point $\mathbf{x}^j \in \Omega(A)$ such that

$$f(\mathbf{x}^j) + S_A(\mathbf{x}^j, \bar{\mathbf{x}}) < u(\bar{\mathbf{x}}) + \frac{1}{j}.$$

Now, for each index j, take a point $\mathbf{y}^j \in \Sigma^*_{\bar{\mathbf{x}}}$ satisfying

$$S_A(\mathbf{x}^j, \bar{\mathbf{x}}) = S_A(\mathbf{x}^j, \tau_{\mathbf{y}^j}(\bar{\mathbf{x}})) - A(\mathbf{y}^j, \bar{\mathbf{x}}) + \beta_A.$$

Finally, let $\bar{\mathbf{y}} \in \Sigma^*_{\bar{\mathbf{x}}}$ be an accumulation point of the sequence $\{\mathbf{y}^j\}$. As $u(\tau_{\mathbf{y}^j}(\bar{\mathbf{x}})) \le f(\mathbf{x}^j) + S_A(\mathbf{x}^j, \tau_{\mathbf{y}^j}(\bar{\mathbf{x}}))$, we verify

$$u(\tau_{\mathbf{y}^j}(\bar{\mathbf{x}})) - A(\mathbf{y}^j, \bar{\mathbf{x}}) + \beta_A < u(\bar{\mathbf{x}}) + \frac{1}{j}.$$

Therefore, $u(\tau_{\bar{\mathbf{y}}}(\bar{\mathbf{x}})) - A(\bar{\mathbf{y}}, \bar{\mathbf{x}}) + \beta_A \le u(\bar{\mathbf{x}})$.

At last, suppose that $f(\bar{\mathbf{x}}) - f(\mathbf{x}) \le S_A(\mathbf{x}, \bar{\mathbf{x}})$ for any $\mathbf{x}, \bar{\mathbf{x}} \in \Omega(A)$. Hence, the inequalities $u(\bar{\mathbf{x}}) \le f(\bar{\mathbf{x}}) \le f(\mathbf{x}) + S_A(\mathbf{x}, \bar{\mathbf{x}})$ are valid for all $\mathbf{x} \in \Omega(A)$, which implies immediately $u = f$ on $\Omega(A)$. □

One of the main consequences of the previous theorem is a kind of *Hölder supremacy* for sub-actions that we will state below. This result corresponds to the well-known fact in Lagrangian Aubry–Mather theory according to which a weak KAM solution is differentiable in the Aubry set (see [**7**]).

COROLLARY 13. *Suppose $\sigma : \Sigma \to \Sigma$ is a transitive subshift of finite type. If $u \in C^0(\Sigma)$ is a sub-action for a potential $A \in C^\theta(\hat{\Sigma})$, then $u|_{\Omega(A)}$ is θ-Hölder.*

Allow us to indicate another immediate consequence of Theorem 12.

COROLLARY 14. *Let $\sigma : \Sigma \to \Sigma$ be a transitive subshift of finite type. Assume that $u \in C^0(\Sigma)$ is a sub-action for a θ-Hölder potential A. Then, for every point $\mathbf{x} \in \Omega(A)$, we verify*

$$u(\mathbf{x}) = \min_{\mathbf{y} \in \Sigma^*_{\mathbf{x}}} [u(\tau_{\mathbf{y}}(\mathbf{x})) - A(\mathbf{y}, \mathbf{x}) + \beta_A].$$

Theorems 10 and 12 assure that every calibrated sub-action for a Hölder potential A is also Hölder. Moreover, we have a complete description of the set of these sub-actions.

THEOREM 15. *Consider $\sigma : \Sigma \to \Sigma$ a transitive subshift of finite type and $A : \Sigma \to \mathbb{R}$ a θ-Hölder potential. Then, there exists a bijective and isometric correspondence between the set of calibrated sub-actions for A and the set of functions $f \in C^0(\Omega(A))$ satisfying $f(\bar{\mathbf{x}}) - f(\mathbf{x}) \le S_A(\mathbf{x}, \bar{\mathbf{x}})$, for all points $\mathbf{x}, \bar{\mathbf{x}} \in \Omega(A)$.*

Proof. Let us analyze the correspondence

$$f \mapsto u_f = \inf_{\mathbf{x} \in \Omega(A)} [f(\mathbf{x}) + S_A(\mathbf{x}, \cdot)].$$

It follows from Theorem 12 that such correspondence is well defined and injective. From Theorem 10 we get that it is surjective. Besides, the correspondence is an isometry. Indeed, fixing $\epsilon > 0$, if $\bar{\mathbf{x}} \in \Sigma$, take a point $\mathbf{x} \in \Omega(A)$ such that $f(\mathbf{x}) + S_A(\mathbf{x}, \bar{\mathbf{x}}) < u_f(\bar{\mathbf{x}}) + \epsilon$. Therefore, we have

$$u_g(\bar{\mathbf{x}}) - u_f(\bar{\mathbf{x}}) - \epsilon < g(\mathbf{x}) - f(\mathbf{x}) \le \|f - g\|_0.$$

When ϵ tends to zero, since $\bar{\mathbf{x}}$ is arbitrary and since we can interchange the roles of f and g, we see that $\|u_f - u_g\|_0 \le \|f - g\|_0$. On the other hand, as $u_f|_{\Omega(A)} = f$ and $u_g|_{\Omega(A)} = g$, we verify $\|u_f - u_g\|_0 \ge \|f - g\|_0$. □

In [**6**], Contreras characterizes the weak KAM solutions of the Hamilton–Jacobi equation in terms of their values at each static class and the values of the action potential of Mañé. The result we presented above describes a similar property for our holonomic setting.

As announced just before the statement of Theorem 4, under the transitive hypothesis, there always exists a calibrated sub-action of maximal character for a Hölder potential. We only need to consider the following one:

$$u_0 = \inf_{\mathbf{x} \in \Omega(A)} S_A(\mathbf{x}, \cdot).$$

Indeed, it is clear that $u_0 \le 0$ on $\Omega(A)$. Moreover, if we take any sub-action $u \in C^0(\Sigma)$ satisfying $u|_{\Omega(A)} \le 0$, since $u(\bar{\mathbf{x}}) \le u(\mathbf{x}) + S_A(\mathbf{x}, \bar{\mathbf{x}}) \le S_A(\mathbf{x}, \bar{\mathbf{x}})$ for $\mathbf{x} \in \Omega(A)$ and $\bar{\mathbf{x}} \in \Sigma$, we verify $u \le u_0$.

Now we will focus also on the support of maximizing holonomic probabilities in order to complete our investigation. We need just two lemmas.

LEMMA 16. *Suppose $\hat{\mu} \in \mathcal{M}_0$. Then, almost every point $(\mathbf{y}, \mathbf{x}) \in \mathrm{supp}(\hat{\mu})$ is of the form $(\mathbf{y}, \tau_{\bar{\mathbf{y}}}(\bar{\mathbf{x}}))$, with $(\bar{\mathbf{y}}, \bar{\mathbf{x}}) \in \mathrm{supp}(\hat{\mu})$.*

Proof. Consider the set

$$\hat{R} = \{(\mathbf{y}, \mathbf{x}) \in \mathrm{supp}(\hat{\mu}) : \mathbf{x} \ne \tau_{\bar{\mathbf{y}}}(\bar{\mathbf{x}}) \ \forall\ (\bar{\mathbf{y}}, \bar{\mathbf{x}}) \in \mathrm{supp}(\hat{\mu})\}.$$

Suppose $\hat{\mu}(\hat{R}) = \epsilon > 0$. Put $R = \pi_1(\hat{R})$. Consider $D \subset \Sigma$ a compact subset and $E \subset \Sigma$ an open subset satisfying $D \subset R \subset E$ with $(\hat{\mu} \circ \pi_1^{-1})(E - D) < \epsilon/2$. Take then a function $f \in C^0(\Sigma, [0, 1])$ such that $f|_D \equiv 1$ and $f|_{\Sigma - E} \equiv 0$. Once $\pi_1^{-1}(R) \cap \mathrm{supp}(\hat{\mu}) = \hat{R}$, we get

$$\int_{\hat{\Sigma}} f(\mathbf{x})\, d\hat{\mu}(\mathbf{y}, \mathbf{x}) \ge \hat{\mu}(\pi_1^{-1}(D)) \ge \hat{\mu}(\pi_1^{-1}(R)) - \hat{\mu}(\pi_1^{-1}(E - D)) > \frac{\epsilon}{2}.$$

Thus, consider a sequence of functions $\{f_j\} \subset C^0(\Sigma, [0, 1])$ such that $f_j \uparrow \chi_{E-D}$. By the monotonic convergence theorem, we obtain

$$\begin{aligned}\int_{\hat{\Sigma}} \chi_{E-D}(\tau_{\mathbf{y}}(\mathbf{x}))\, d\hat{\mu}(\mathbf{y}, \mathbf{x}) &= \lim_{j\to\infty} \int_{\hat{\Sigma}} f_j(\tau_{\mathbf{y}}(\mathbf{x}))\, d\hat{\mu}(\mathbf{y}, \mathbf{x}) \\ &= \lim_{j\to\infty} \int_{\hat{\Sigma}} f_j(\mathbf{x})\, d\hat{\mu}(\mathbf{y}, \mathbf{x}) \\ &= \hat{\mu}(\pi_1^{-1}(E-D)) < \frac{\epsilon}{2}.\end{aligned}$$

Note that, from the definition of R, we have $\int_{\mathrm{supp}(\hat{\mu})} \chi_R(\tau_{\mathbf{y}}(\mathbf{x}))\, d\hat{\mu}(\mathbf{y}, \mathbf{x}) = 0$. Hence, as $0 \le f \le \chi_E$, we verify

$$\begin{aligned}\int_{\hat{\Sigma}} f(\tau_{\mathbf{y}}(\mathbf{x}))\, d\hat{\mu}(\mathbf{y}, \mathbf{x}) &\le \int_{\mathrm{supp}(\hat{\mu})} \chi_{E-R}(\tau_{\mathbf{y}}(\mathbf{x}))\, d\hat{\mu}(\mathbf{y}, \mathbf{x}) \\ &\le \int_{\mathrm{supp}(\hat{\mu})} \chi_{E-D}(\tau_{\mathbf{y}}(\mathbf{x}))\, d\hat{\mu}(\mathbf{y}, \mathbf{x}) < \frac{\epsilon}{2}.\end{aligned}$$

However, since $f \in C^0(\Sigma)$ and $\hat{\mu} \in \mathcal{M}_0$, it follows that $\int_{\hat{\Sigma}} f(\mathbf{x})\, d\hat{\mu}(\mathbf{y}, \mathbf{x}) < \epsilon/2$.

We then get a contradiction. Therefore, $\hat{\mu}(\hat{R}) = 0$. □

We need also a result on numerical sequences.

LEMMA 17. *Consider a sequence $\{a_j\} \subset \mathbb{R}$ for which the following is true:*

$$\lim_{k\to\infty} \frac{1}{k} \sum_{j=1}^{k} a_j = b.$$

Let R be a subset of the set of positive integers satisfying

$$\lim_{k\to\infty} \frac{1}{k} \#\{j \in R : j \le k\} > 0.$$

Then, for any $\epsilon > 0$ and any positive integer K, there exist $k_1, k_2 \in R$ such that $k_2 > k_1 \ge K$ and

$$\left| \sum_{j=k_1+1}^{k_2} a_j - (k_2 - k_1)b \right| < \epsilon.$$

The previous lemma was used by Mañé in [**21**]. We can now present the following result.

PROPOSITION 18. *Suppose $\sigma : \Sigma \to \Sigma$ is a transitive subshift of finite type. Let A be a θ-Hölder potential. Assume $\hat{\mu} \in m_A$ with $\hat{\mu} \circ \pi_1^{-1}$ ergodic. Then $\pi_1(\mathrm{supp}(\hat{\mu})) \subset \Omega(A)$.*

Proof. It is enough to show that $(\hat{\mu} \circ \pi_1^{-1})(\Omega(A)) = 1$. Fix $\epsilon > 0$. Denote by $\Omega(A, \epsilon)$ the set of the points $\mathbf{x} \in \Sigma$ for which we can find a path $\{(\mathbf{y}^0, \mathbf{x}^0), \ldots, (\mathbf{y}^{k-1}, \mathbf{x}^{k-1})\} \in \mathcal{P}(\mathbf{x}, \mathbf{x}, \epsilon)$ satisfying

$$\left| \sum_{j=0}^{k-1} (A - \beta_A)(\mathbf{y}^j, \mathbf{x}^j) \right| < \epsilon.$$

As $\Omega(A) = \bigcap \Omega(A, 1/j)$, it is enough to show that $(\hat{\mu} \circ \pi_1^{-1})(\Omega(A, \epsilon)) = 1$.

Suppose, however, that $(\hat{\mu} \circ \pi_1^{-1})(\pi_1(\mathrm{supp}(\hat{\mu})) - \Omega(A, \epsilon)) > 0$. Take an integer $l > 0$ sufficiently large in such a way that $2\lambda^l < \epsilon$. So there exists $\mathbf{x} \in \pi_1(\mathrm{supp}(\hat{\mu}))$ such that $(\hat{\mu} \circ \pi_1^{-1})(D_l - \Omega(A, \epsilon)) > 0$, where D_l is the open ball of radius λ^l centered at the point $\mathbf{x}$.

Thus, consider a point $\bar{\mathbf{x}} \in \pi_1(\mathrm{supp}(\hat{\mu}))$ such that

$$\lim_{k\to\infty} \frac{1}{k} \#\{0 \le j < k : \sigma^j(\bar{\mathbf{x}}) \in D_l - \Omega(A, \epsilon)\} > 0.$$

Thanks to Lemma 16, we can assume that, for every index $j > 0$, there exists a point $\bar{\mathbf{y}}^j \in \Sigma^*$ such that $(\bar{\mathbf{y}}^j, \sigma^j(\bar{\mathbf{x}})) \in \mathrm{supp}(\hat{\mu})$ and $\sigma^{j-1}(\bar{\mathbf{x}}) = \tau_{\bar{\mathbf{y}}^j}(\sigma^j(\bar{\mathbf{x}}))$.

As $u \in C^0(\Sigma)$ is an arbitrary sub-action for A, from Proposition 9 we get that $A(\bar{\mathbf{y}}^j, \sigma^j(\bar{\mathbf{x}})) - \beta_A = u(\sigma^{j-1}(\bar{\mathbf{x}})) - u(\sigma^j(\bar{\mathbf{x}}))$. Define, finally,

$$a_j = u(\sigma^{j-1}(\bar{\mathbf{x}})) - u(\sigma^j(\bar{\mathbf{x}})) \quad \text{and} \quad R = \{j : \sigma^j(\bar{\mathbf{x}}) \in D_l - \Omega(A, \epsilon)\}.$$

Using Lemma 17, we obtain integers $k_1, k_2 \in R$, with $1 \le k_1 < k_2$, accomplishing

$$\left| \sum_{j=k_1+1}^{k_2} (A - \beta_A)(\bar{\mathbf{y}}^j, \sigma^j(\bar{\mathbf{x}})) \right| = \left| \sum_{j=k_1+1}^{k_2} a_j \right| < \epsilon.$$

However, once $\sigma^{k_1}(\bar{\mathbf{x}}), \sigma^{k_2}(\bar{\mathbf{x}}) \in D_l$, it follows that $d(\sigma^{k_1}(\bar{\mathbf{x}}), \sigma^{k_2}(\bar{\mathbf{x}})) \le 2\lambda^l$. Therefore, $\{(\bar{\mathbf{y}}^{k_2}, \sigma^{k_2}(\bar{\mathbf{x}})), \ldots, (\bar{\mathbf{y}}^{k_1+1}, \sigma^{k_1+1}(\bar{\mathbf{x}}))\} \in \mathcal{P}(\sigma^{k_2}(\bar{\mathbf{x}}), \sigma^{k_2}(\bar{\mathbf{x}}), \epsilon)$ yields $\sigma^{k_2}(\bar{\mathbf{x}}) \in \Omega(A, \epsilon)$. This is a contradiction because $k_2 \in R$.

Hence, $(\hat{\mu} \circ \pi_1^{-1})(\Omega(A, \epsilon)) = 1$. □

Remember that the addition of a constant does not change the role played by a sub-action. Thus, the next proposition indicates a kind of rigidity created by the previous ergodic assumption.

PROPOSITION 19. *Consider a probability $\hat{\mu} \in \mathcal{m}_A$ such that $\hat{\mu} \circ \pi_1^{-1}$ is ergodic. If $u, u' \in C^0(\Sigma)$ are sub-actions for $A \in C^0(\hat{\Sigma})$, then $u - u'$ is identically constant on $\pi_1(\mathrm{supp}(\hat{\mu}))$.*

Proof. Suppose $\mathbf{x} \in \pi_1(\mathrm{supp}(\hat{\mu}))$. We can use Lemma 16 in order to get a point $(\bar{\mathbf{y}}, \bar{\mathbf{x}}) \in \mathrm{supp}(\hat{\mu})$ such that $\mathbf{x} = \tau_{\bar{\mathbf{y}}}(\bar{\mathbf{x}})$.

From Proposition 9, we verify

$$u(\bar{\mathbf{x}}) - u(\mathbf{x}) = \beta_A - A(\bar{\mathbf{y}}, \bar{\mathbf{x}}) = u'(\bar{\mathbf{x}}) - u'(\mathbf{x}).$$

So $(u - u')(\mathbf{x}) = (u - u')(\bar{\mathbf{x}}) = (u - u') \circ \sigma(\mathbf{x})$. Therefore, we have $u - u' = (u - u') \circ \sigma$ on $\pi_1(\mathrm{supp}(\hat{\mu}))$. As the probability $\hat{\mu} \circ \pi_1^{-1}$ is ergodic, it follows immediately that $u - u'$ is constant on $\pi_1(\mathrm{supp}(\hat{\mu}))$. □

Let us consider again the transitivity hypothesis and assume that A is Hölder. Given u a sub-action for A, let $\mathbb{M}_A(u)$ be its corresponding contact locus. Then, we claim that $\Omega(A) \subset \pi_1(\mathbb{M}_A(u))$. This is completely obvious when u is a calibrated sub-action, because in such a case $\pi_1(\mathbb{M}_A(u)) = \Sigma$. Besides, Corollary 14 tells us that every sub-action $u \in C^0(\Sigma)$ for the potential A behaves as a calibrated sub-action on $\Omega(A)$.

Therefore, the following inclusions are true:

$$\bigcup_{\substack{\hat{\mu}\in m_A \\ \hat{\mu}\circ\pi_1^{-1} \text{ ergodic}}} \pi_1(\mathrm{supp}(\hat{\mu})) \subset \Omega(A) \subset \bigcap_{\substack{u\in C^0(\Sigma) \\ u \text{ sub-action}}} \pi_1(\mathbb{M}_A(u)).$$

In some situations for the standard model $(X, T, \mathcal{M}_T)$, it is known that, given a Hölder potential A, a probability is A-maximizing if, and only if, its support is contained in the set of non-wandering points (with respect to A). See, for instance, the case of expanding maps of the circle in [**9,** Proposition 15(ii)] and also the case of Anosov diffeomorphisms in [**19,** Lemmas 12 and 13].

Hence, it is natural to ask the following: In order to verify that $\hat{\mu} \in m_A$, would it be enough to check that $\hat{\mu} \circ \pi_1^{-1}$ is ergodic and $\pi_1(\mathrm{supp}(\hat{\mu})) \subset \Omega(A)$? The answer is 'no'.

Indeed, here is a counter-example. Take a potential $A : \{0, 1\}^{\mathbb{Z}} \to \mathbb{R}$ depending just on three coordinates in such a way that $A(1, 1 \mid 1) > A(s, s' \mid s'')$ whenever $s + s' + s'' \leq 2$. If we denote by $\underline{ss'}$ either the periodic point $(s, s', \ldots, s, s', \ldots) \in \Sigma$, or the periodic point $(\ldots, s, s', \ldots, s, s') \in \Sigma^*$, then we have $\delta_{(\underline{11},\underline{11})}, \delta_{(\underline{01},\underline{11})} \in \mathcal{M}_0$ with $\delta_{(\underline{11},\underline{11})} \circ \pi_1^{-1} = \delta_{\underline{11}} = \delta_{(\underline{01},\underline{11})} \circ \pi_1^{-1}$. Nevertheless, observe that $\delta_{(\underline{11},\underline{11})}$ is a maximizing probability, but clearly $\delta_{(\underline{01},\underline{11})} \notin m_A$.

The second inclusion above also brings us an interesting question: What can be said about $\pi_1(\mathbb{M}_A(u)) - \Omega(A)$? The next proposition gives a partial answer.

PROPOSITION 20. *Let $\sigma : \Sigma \to \Sigma$ be a transitive subshift of finite type and assume $A \in C^\theta(\hat{\Sigma})$ is not cohomologous to a constant. Take $u \in C^0(\Sigma)$ an arbitrary sub-action for A. Then, for each positive integer k, there exists a sub-action $U_k \in C^0(\Sigma)$ satisfying*

$$\pi_1(\mathbb{M}_A(U_k)) \subset \bigcap_{j=0}^{k-1} \sigma^{-j}(\pi_1(\mathbb{M}_A(u))).$$

Moreover, if u is θ-Hölder, then we can also take U_k as a θ-Hölder function.

Proof. We begin with $A^u = A + u \circ \pi_1 - u \circ \pi_1 \circ \hat{\sigma}^{-1} \leq \beta_A$.

Given $k > 0$ and $\mathbf{x} \in \Sigma$, we call a path of size k ending at the point $\mathbf{x}$ an ordered sequence of points $(\mathbf{y}^0, \mathbf{x}^0), \ldots, (\mathbf{y}^{k-1}, \mathbf{x}^{k-1}) \in \hat{\Sigma}$ which verifies $\mathbf{x}^0 = \mathbf{x}$ and $\mathbf{x}^{j+1} = \tau_{\mathbf{y}^j}(\mathbf{x}^j)$ for $0 \leq j < k - 1$. Denote by $\mathcal{P}_k(\mathbf{x})$ the set of such paths. Note that

$$\sum_{j=0}^{k-1} A^u(\mathbf{y}^j, \mathbf{x}^j) \leq k\beta_A$$

for $\{(\mathbf{y}^0, \mathbf{x}^0), \ldots, (\mathbf{y}^{k-1}, \mathbf{x}^{k-1})\} \in \mathcal{P}_k(\mathbf{x})$.

Taking $\{(\mathbf{y}^0, \sigma^{k-1}(\mathbf{x})), (\mathbf{y}^1, \sigma^{k-2}(\mathbf{x})), \ldots, (\mathbf{y}^{k-1}, \mathbf{x})\} \in \mathcal{P}_k(\sigma^{k-1}(\mathbf{x}))$, we have the identity

$$\sum_{j=0}^{k-1} A(\mathbf{y}^j, \sigma^{k-1-j}(\mathbf{x}))$$
$$= kA(\mathbf{y}^{k-1}, \mathbf{x}) + \sum_{j=0}^{k-1} jA(\mathbf{y}^{j-1}, \sigma^{k-j}(\mathbf{x})) - \sum_{j=0}^{k-1} jA(\mathbf{y}^j, \sigma^{k-1-j}(\mathbf{x})).$$

Now we define $W : \Sigma \to \mathbb{R}$ in the following way:

$$W(\mathbf{x}) = \max_{\{(\mathbf{y}^0, \sigma^{k-1}(\mathbf{x})), \ldots, (\mathbf{y}^{k-1}, \mathbf{x})\} \in \mathcal{P}_k(\sigma^{k-1}(\mathbf{x}))} \left[\frac{1}{k} \sum_{j=1}^{k-1} j A(\mathbf{y}^{j-1}, \sigma^{k-j}(\mathbf{x})) \right].$$

Once the correspondence $\mathbf{x} \mapsto \max_{y_0 = x_0} A(\mathbf{y}, \sigma(\mathbf{x}))$ is θ-Hölder, the same is true for the function W.

Fix a point $(\mathbf{y}, \mathbf{x}) \in \hat{\Sigma}$. Then consider a path

$$\{(\mathbf{y}^0, \sigma^{k-1}(\mathbf{x})), \ldots, (\mathbf{y}^{k-2}, \sigma(\mathbf{x})), (\mathbf{y}, \mathbf{x})\} \in \mathcal{P}_k(\sigma^{k-1}(\mathbf{x}))$$

accomplishing

$$\frac{1}{k} \sum_{j=1}^{k-1} j A(\mathbf{y}^{j-1}, \sigma^{k-j}(\mathbf{x})) = W(\mathbf{x}).$$

Put $\mathbf{y}^{k-1} = \mathbf{y}$. As $\{(\mathbf{y}^1, \sigma^{k-2}(\mathbf{x})), \ldots, (\mathbf{y}^{k-1}, \mathbf{x})\} \in \mathcal{P}_{k-1}(\sigma^{k-1}(\tau_{\mathbf{y}}(\mathbf{x})))$, without difficulty we get

$$\begin{aligned} A(\mathbf{y}, \mathbf{x}) + W(\mathbf{x}) - W(\tau_{\mathbf{y}}(\mathbf{x})) &\leq A(\mathbf{y}^{k-1}, \mathbf{x}) + \frac{1}{k} \sum_{j=0}^{k-1} j A(\mathbf{y}^{j-1}, \sigma^{k-j}(\mathbf{x})) \\ &\quad - \frac{1}{k} \sum_{j=0}^{k-1} j A(\mathbf{y}^{j}, \sigma^{k-1-j}(\mathbf{x})) \\ &= \frac{1}{k} \sum_{j=0}^{k-1} A(\mathbf{y}^{j}, \sigma^{k-1-j}(\mathbf{x})). \end{aligned}$$

Therefore, if we denote $U_k = W + k^{-1} S_k u$, we obtain

$$\begin{aligned} A(\mathbf{y}, \mathbf{x}) + U_k(\mathbf{x}) - U_k(\tau_{\mathbf{y}}(\mathbf{x})) &\leq \frac{1}{k} \sum_{j=0}^{k-1} A(\mathbf{y}^{j}, \sigma^{k-1-j}(\mathbf{x})) + \frac{1}{k} S_k u(\mathbf{x}) - \frac{1}{k} S_k u(\tau_{\mathbf{y}}(\mathbf{x})) \\ &= \frac{1}{k} \sum_{j=0}^{k-1} A^u(\mathbf{y}^{j}, \sigma^{k-1-j}(\mathbf{x})) \leq \beta_A. \end{aligned}$$

Hence, U_k is a sub-action for the potential A.

Let us check that such a sub-action U_k accomplishes the claim of the proposition. We just follow the itinerary of the construction of U_k in the opposite direction. If $\mathbf{x} \in \pi_1(\mathbb{M}_A(U_k))$, then there exists a path

$$\{(\mathbf{y}^0, \sigma^{k-1}(\mathbf{x})), \ldots, (\mathbf{y}^{k-1}, \mathbf{x})\} \in \mathcal{P}_k(\sigma^{k-1}(\mathbf{x}))$$

such that

$$\frac{1}{k} \sum_{j=0}^{k-1} A^u(\mathbf{y}^{j}, \sigma^{k-1-j}(\mathbf{x})) = \beta_A,$$

which yields $A^u(\mathbf{y}^j, \sigma^{k-1-j}(\mathbf{x})) = \beta_A$. Thus, clearly $\sigma^{k-1-j}(\mathbf{x}) \in \pi_1(\mathbb{M}_A(u))$ for all $j \in \{0, \ldots, k-1\}$. □

The proof described above found inspiration in the strategy used by Bousch in [**5**].

The previous proposition brings our attention to the following question: Does a non-calibrated sub-action exist? The answer is ‘yes’.

Under the same hypotheses as in Proposition 20, assume that $u \in C^\theta(\Sigma)$ is a calibrated sub-action. Suppose the existence of a point $(\mathbf{y}^0, \mathbf{x}^0) \in \hat{\Sigma}$ satisfying both $A(\mathbf{y}^0, \mathbf{x}^0) = \max_{y_0 = y_0^0} A(\mathbf{y}, \mathbf{x}^0)$ and

$$A(\mathbf{y}^0, \mathbf{x}^0) + u(\mathbf{x}^0) - u(\tau_{\mathbf{y}^0}(\mathbf{x}^0)) < \beta_A.$$

(These assumptions are obviously verified by any potential $A \in C^\theta(\Sigma)$ not cohomologous to a constant.) We claim that the function $U \in C^\theta(\Sigma)$ defined by

$$U(\mathbf{x}) = \tfrac{1}{2}[u(\sigma(\mathbf{x})) + u(\mathbf{x})] + \tfrac{1}{2} \max_{y_0 = x_0} A(\mathbf{y}, \sigma(\mathbf{x}))$$

is a sub-action for A which is not calibrated. Indeed, the function U is nothing other than the sub-action U_2 described in the proof of the previous proposition. Moreover, note that, for all $\mathbf{y} \in \Sigma^*_{\tau_{\mathbf{y}^0}(\mathbf{x}^0)}$,

$$\begin{aligned} A(\mathbf{y}, \tau_{\mathbf{y}^0}(\mathbf{x}^0)) &+ U(\tau_{\mathbf{y}^0}(\mathbf{x}^0)) - U(\tau_{\mathbf{y}}(\tau_{\mathbf{y}^0}(\mathbf{x}^0))) \\ &\leq \tfrac{1}{2}[A(\mathbf{y}, \tau_{\mathbf{y}^0}(\mathbf{x}^0)) + u(\tau_{\mathbf{y}^0}(\mathbf{x}^0)) - u(\tau_{\mathbf{y}}(\tau_{\mathbf{y}^0}(\mathbf{x}^0)))] \\ &\quad + \tfrac{1}{2}[A(\mathbf{y}^0, \mathbf{x}^0) + u(\mathbf{x}^0) - u(\tau_{\mathbf{y}^0}(\mathbf{x}^0))] < \beta_A, \end{aligned}$$

and therefore $\tau_{\mathbf{y}^0}(\mathbf{x}^0) \notin \pi_1(\mathbb{M}_A(U))$.

A deeper study of non-calibrated sub-actions is the aim of a subsequent paper [**14**]. Finally, we would like to mention that the possibility of adapting our holonomic setting to the case of iterated function systems has been recently announced [**22**].

Acknowledgements. E. Garibaldi was supported by a CAPES PhD scholarship. A. O. Lopes was partially supported by CNPq, PRONEX – Sistemas Dinâmicos, Instituto do Milênio, and is the beneficiary of CAPES financial support.

REFERENCES

[1] V. Bangert. Mather sets for twist maps and geodesics on tori. *Dynamics Reported* **1** (1988), 1–56.

[2] A. Baraviera, A. O. Lopes and P. Thieullen. A large deviation principle for equilibrium states of Holder potentials: the zero temperature case. *Stoch. Dyn.* **6** (2006), 77–96.

[3] T. Bousch. Le poisson n’a pas d’arêtes. *Ann. Inst. H. Poincaré Probab. Statist.* **36** (2000), 489–508.

[4] T. Bousch. La condition de Walters. *Ann. Sci. École Norm. Sup.* (4) **34** (2001), 287–311.

[5] T. Bousch. Un lemme de Mañé bilatéral. *C. R. Math. Acad. Sci. Paris* **335** (2002), 533–536.

[6] G. Contreras. Action potential and weak KAM solutions. *Calc. Var. Partial Differential Equations* **13** (2001), 427–458.

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

[8] G. Contreras, R. Iturriaga, G. P. Paternain and M. Paternain. Lagrangian graphs, minimizing measures and Mañé’s critical values. *Geom. Funct. Anal.* **8** (1998), 788–809.

[9] G. Contreras, A. O. Lopes and P. Thieullen. Lyapunov minimizing measures for expanding maps of the circle. *Ergod. Th. & Dynam. Sys.* **21** (2001), 1379–1409.

[10] J. P. Conze and Y. Guivarc'h. Croissance des sommes ergodiques et principe variationnel. *Technical Report*, Université de Rennes 1, 1993.

[11] A. Fathi. Théorème KAM faible et théorie de Mather sur les systèmes lagrangiens. *C. R. Acad. Sci. Paris, Sér. I, Math.* **324** (1997), 1043–1046.

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

[13] E. Garibaldi and A. O. Lopes. Functions for relative maximization. *Preprint*, 2006.

[14] E. Garibaldi, A. O. Lopes and P. Thieullen. On separating sub-actions. *Preprint*, 2006.

[15] D. A. Gomes. Viscosity solution method and the discrete Aubry–Mather problem. *Discrete Contin. Dyn. Syst. Ser. A* **13** (2005), 103–116.

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

[17] O. Jenkinson. Ergodic optimization. *Discrete Contin. Dyn. Syst. Ser. A* **15** (2006), 197–224.

[18] O. Jenkinson. Every ergodic measure is uniquely maximizing. *Discrete Contin. Dyn. Syst. Ser. A* **16** (2006), 383–392.

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

[20] A. O. Lopes and P. Thieullen. Mather measures and the Bowen–Series transformation. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **23** (2006), 663–682.

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

[22] E. R. Oliveira. Propriedades genéricas de lagrangianos e problemas variacionais holonômicos em sistemas de funções iteradas. *PhD Thesis* (Preliminary Version), Universidade Federal do Rio Grande do Sul, 2007.

[23] W. Parry and M. Pollicott. Zeta functions and the periodic orbit structure of hyperbolic dynamics. *Astérisque* **187–188** (1990), 1–268.

[24] L. Radu. Duality in thermodynamic formalism. *Preprint*, 2004.