LECTURE NOTES ON MATHER'S THEORY FOR LAGRANGIAN SYSTEMS

ALFONSO SORRENTINO

Abstract. These notes are based on a series of lectures that the author gave at the CIMPA Research School *Hamiltonian and Lagrangian Dynamics*, which was held in Salto (Uruguay) in March 2015.

To the memory of Ricardo Mañé
(1948 – 1995)

Contents

1. Introduction

In these lecture notes we provide a brief introduction to John Mather's variational approach to the study of convex and superlinear Hamiltonian systems, what is generally called *Aubry-Mather theory*. Starting from the observation that invariant Lagrangian graphs can be characterized in terms of their "action-minimizing" properties, we then describe how analogue features can be traced in a more general setting, namely the so-called *Tonelli Hamiltonian systems*. This approach brings to light a plethora of compact invariant subsets for the system, which, under many points of view, can be considered as generalization of invariant Lagrangian graphs, despite not being in general either submanifolds or regular. Besides being very significant from a dinamical systems point of view, these objects also appear and play an important role in many other different contexts: PDEs (e.g., Hamilton-Jacobi equation and weak KAM theory), Symplectic geometry, etc...

Date: June 16, 2015.

Since this notes[1] are meant to be a short introduction and a guide to this theory, we will omit most of the proofs. We refer interested readers to [23] for a more systematic and comprehensive presentation of this and other topics.

ACKNOWLEDGEMENT. The author wishes to thank CIMPA and, more specifically, Ezequiel Maderna and Ludovic Rifford, for the organization of this very interesting meeting and for their kind invitation. The author also wishes to thank Princeton University Press for agreeing on the use herein of some of the material from [23].

2. From KAM theory to Aubry-Mather (AM) theory

The celebrated Kolmogorov-Arnol′d -Moser (or KAM) theorem finally settled the old question concerning the existence of *quasi-periodic* motions for *nearly-integrable* Hamiltonian systems, *i.e.,* Hamiltonian systems that are slight perturbation of an integrable one. In the integrable case, in fact, the whole phase space is foliated by invariant Lagrangian submanifolds that are diffeomorphic to tori, and on which the dynamics is conjugate to a rigid rotation. More specifically, let $H : T^*\mathbb{T}^n \longrightarrow \mathbb{R}$ be an integrable Tonelli Hamiltonian in action-angle coordinates, *i.e.*, $H(x,p) = \mathfrak{h}(p)$ with the Hamiltonian depending only on the action variables (see [2])[2]. Let us denote by $\phi_t^{\mathfrak{h}}$ the associated Hamiltonian flow and identify $T^*\mathbb{T}^n$ with $\mathbb{T}^n \times \mathbb{R}^n$, where $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$.

The Hamiltonian flow in this case is very easy to study. Hamilton's equations are:

$$\begin{cases} \dot{x} = \frac{\partial \mathfrak{h}}{\partial p}(p) =: \rho(p) \\ \dot{p} = -\frac{\partial \mathfrak{h}}{\partial x}(p) = 0, \end{cases}$$

therefore $\Phi_t^{\mathfrak{h}}(x_0, p_0) = (x_0 + t\rho(p_0) \mod \mathbb{Z}^n,\, p_0)$. In particular, p is an integral of motion, that is, it remains constant along the orbits. The phase space $T^*\mathbb{T}^n$ is hence foliated by invariant tori $\Lambda^*_{p_0} = \mathbb{T}^n \times \{p_0\}$ on which the motion is a rigid rotation with rotation vector $\rho(p_0)$ (see figure below).

On the other hand, it is natural to ask what happens to such a foliation and to these stable motions once the system is perturbed. In 1954 Kolmogorov [11] - and later Arnol′d [1] and Moser [21] in different contexts - proved that, in spite of the generic disappearance of the invariant tori filled by periodic orbits (already pointed out by Henri Poincaré), for small perturbations of an integrable system it is still possible to find invariant Lagrangian tori corresponding to certain rotation vectors (the so-called *diophantine* rotation vectors). This result is commonly referred to as *KAM theorem*, from the initials of the three main pioneers. In addition to open the way to a new understanding of the nature of Hamiltonian systems and their stable

[1]Portions of this material used with permission from Princeton University Press from "*Action-minimizing Methods in Hamiltonian Dynamics: An Introduction to Aubry-Mather Theory*" by Alfonso Sorrentino, 2015 (see [23]).

[2]In general these coordinates can be defined only locally. For the sake of simplicity, in this example we assume – without affecting its main purpose – that they are defined globally.

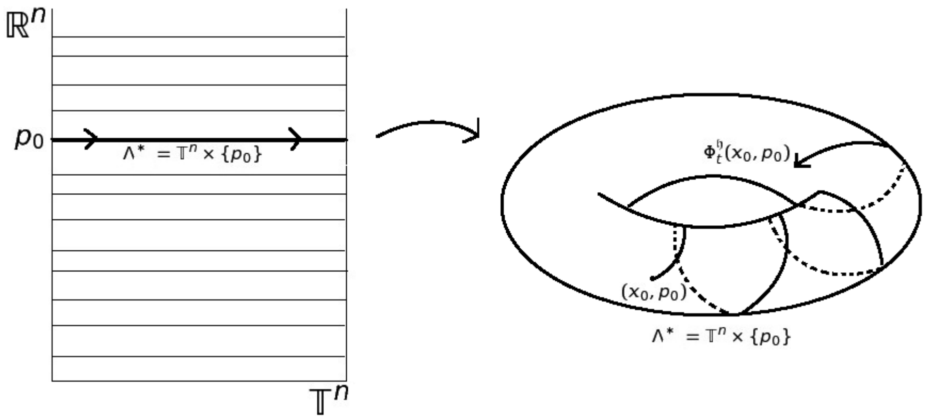

FIGURE 1. The phase space of an integrable system.

motions, this result contributed to raise new interesting questions, such as: what does it happen to the stable motions that are destroyed by effect of the perturbation? Is it possible to identify something reminiscent of their past presence? What can be said for systems that not close to an integrable one?

Aubry-Mather theory provides answers to these questions. Developed independently by Serge Aubry [3] and John Mather [14] in 1980s, this novel approach to the study of the dynamics of *twist diffeomorphisms of the annulus* (which correspond to Poincaré maps of 1-dimensional non-autonomous Hamiltonian systems) pointed out the existence of many invariant sets, which are obtained by means of variational methods and that always exist, even after rotational curves are destroyed. Besides providing a detailed structure theory for these new sets, this powerful approach yielded to a better understanding of the destiny of invariant rotational curves and to the construction of interesting chaotic orbits as a result of their destruction [15, 17].

Motivated by these achievements, John Mather [18, 19] - and later Ricardo Mañé [13, 12] and Albert Fathi [9] in different ways - developed a generalization of this theory to higher dimensional systems. Positive definite superlinear Lagrangians on compact manifolds, also called *Tonelli Lagrangians* (see Definition 3.1), were the appropriate setting to work in. Under these conditions, in fact, it is possible to prove the existence of interesting invariant sets, known as *Mather*, *Aubry* and *Mañé* sets, which generalize KAM tori and invariant Lagrangian graphs, and which continue to exist beyond the nearly-integrable case.

In the following we will provide a brief overview of Mather's theory. We will first discuss an illustrative example (what happens in the integrable case) and then show how similar ideas can be extended to a more general setting.

3. Tonelli Lagrangians and Hamiltonians on compact manifolds

Before starting, let us introduce the basic setting that we will consider in the following. Let M be a compact and connected smooth manifold without boundary. Denote by TM its tangent bundle and T^*M the cotangent one. A point of TM will be denoted by (x,v), where $x \in M$ and $v \in T_xM$, and a point of T^*M by (x,p), where $p \in T_x^*M$ is a linear form on the vector space T_xM. Let us fix a Riemannian metric g on it and denote by d the induced metric on M; let $\|\cdot\|_x$ be the norm induced by g on T_xM; we will use the same notation for the norm induced on T_x^*M.

We will consider functions $L : TM \longrightarrow \mathbb{R}$ of class C^2, which are called *Lagrangians*. Associated to each Lagrangian, there is a flow on TM called the *Euler-Lagrange flow*, defined as follows. Let us consider the action functional A_L from the space of absolutely continuous curves $\gamma : [a,b] \to M$, with $a \leq b$, defined by:

$$A_L(\gamma) := \int_a^b L(\gamma(t), \dot\gamma(t))\, dt.$$

Curves that extremize[3] this functional among all curves with the same end-points (and the same time-length) are solutions of the *Euler-Lagrange equation*:

$$\frac{d}{dt}\frac{\partial L}{\partial v}(\gamma(t),\dot\gamma(t)) = \frac{\partial L}{\partial x}(\gamma(t),\dot\gamma(t)) \qquad \forall\, t \in [a,b]\,.$$

Observe that this equation is equivalent to

$$\frac{\partial^2 L}{\partial v^2}(\gamma(t),\dot\gamma(t))\ddot\gamma(t) = \frac{\partial L}{\partial x}(\gamma(t),\dot\gamma(t)) - \frac{\partial^2 L}{\partial v \partial x}(\gamma(t),\dot\gamma(t))\dot\gamma(t)\,,$$

therefore, if the second partial vertical derivative $\partial^2 L/\partial v^2(x,v)$ is non-degenerate at all points of TM, we can solve for $\ddot\gamma(t)$. This condition

$$\det \frac{\partial^2 L}{\partial v^2} \neq 0$$

is called *Legendre condition* and allows one to define a vector field X_L on TM, such that the solutions of $\ddot\gamma(t) = X_L(\gamma(t),\dot\gamma(t))$ are precisely the curves satisfying the Euler-Lagrange equation. This vector field X_L is called the *Euler-Lagrange vector field* and its flow Φ_t^L is the *Euler-Lagrange flow* associated to L. It turns out that Φ_t^L is C^1 even if L is only C^2 (see Remark 3.3).

Definition 3.1 (Tonelli Lagrangian). *A function $L : TM \longrightarrow \mathbb{R}$ is called a Tonelli Lagrangian if:*

i) $L \in C^2(TM)$;
ii) *L is strictly convex in the fibers, in the C^2 sense, i.e., the second partial vertical derivative $\partial^2 L/\partial v^2(x,v)$ is positive definite, as a quadratic form, for all (x,v);*
iii) *L is superlinear in each fiber, i.e.,*

$$\lim_{\|v\|_x \to +\infty} \frac{L(x,v)}{\|v\|_x} = +\infty.$$

[3]These extremals are not in general minima. The existence of global minima and the study of the corresponding motions is the core of Aubry-Mather theory; see section 5.

This condition is equivalent to ask that for each $A \in \mathbb{R}$ *there exists* $B(A) \in \mathbb{R}$ *such that*

$$L(x,v) \geq A\|v\| - B(A) \qquad \forall\,(x,v) \in TM\,.$$

Observe that since the manifold is compact, then condition *iii)* is independent of the choice of the Riemannian metric g.

Examples of Tonelli Lagrangians.

- **Riemannian Lagrangians.** Given a Riemannian metric g on TM, the *Riemannian Lagrangian* on (M,g) is given by the *kinetic energy*:
$$L(x,v) = \frac{1}{2}\|v\|_x^2\,.$$
Its Euler-Lagrange equation is the equation of the geodesics of g:
$$\frac{D}{dt}\dot{x} \equiv 0\,,$$
and its Euler-Lagrange flow coincides with the geodesic flow.
- **Mechanical Lagrangians.** These Lagrangians play a key-role in the study of classical mechanics. They are given by the sum of the kinetic energy and a *potential* $U : M \longrightarrow \mathbb{R}$:
$$L(x,v) = \frac{1}{2}\|v\|_x^2 + U(x)\,.$$
The associated Euler-Lagrange equation is given by:
$$\frac{D}{dt}\dot{x} = \nabla U(x)\,.$$
- **Mañé's Lagrangians.** This is a particular class of Tonelli Lagrangians, introduced by Ricardo Mañé in [12]. If X is a C^k vector field on M, with $k \geq 2$, one can embed its flow φ_t^X into the Euler-Lagrange flow associated to a certain Lagrangian, namely
$$L_X(x,v) = \frac{1}{2}\,\|v - X(x)\|_x^2\,.$$
It is quite easy to check that the integral curves of the vector field X are solutions of the Euler-Lagrange equation. In particular, the Euler-Lagrange flow $\Phi_t^{L_X}$ restricted to $\mathrm{Graph}(X) = \{(x, X(x)),\ x \in M\}$ (which is clearly invariant) is conjugate to the flow of X on M and the conjugacy is given by $\pi|\mathrm{Graph}(X)$, where $\pi : TM \to M$ is the canonical projection. In other words, the following diagram commutes:
$$\begin{tikzcd} \mathrm{Graph}(X) \arrow[r, "\Phi_t^{L_X}"] \arrow[d, "\pi"'] & \mathrm{Graph}(X) \arrow[d, "\pi"] \\ M \arrow[r, "\varphi_t^X"'] & M \end{tikzcd}$$
that is, for every $x \in M$ and every $t \in \mathbb{R}$, $\Phi_t^{L_X}(x, X(x)) = (\gamma_x^X(t), \dot{\gamma}_x^X(t))$, where $\gamma_x^X(t) = \varphi_t^X(x)$.

In the study of classical dynamics it turns often very useful to consider the associated *Hamiltonian system*, which is defined on the cotangent bundle T^*M. Given a Lagrangian L we can define the associated *Hamiltonian* as its *Fenchel transform* (or *Legendre-Fenchel transform*), see [22]:

$$\begin{aligned} H: T^*M &\longrightarrow \mathbb{R} \\ (x,p) &\longmapsto \sup_{v\in T_xM} \{\langle p,\, v\rangle_x - L(x,v)\} \end{aligned}$$

where $\langle\,\cdot,\,\cdot\,\rangle_x$ denotes the canonical pairing between the tangent and cotangent bundles.

If L is a Tonelli Lagrangian, one can easily prove that H is finite everywhere (as a consequence of the superlinearity of L), superlinear and strictly convex in each fiber (in the C^2 sense). Observe that H is also C^2. In fact the Euler-Lagrange vector field corresponds, under the Legendre transformation, to a vector field on T^*M given by Hamilton's equation; it is easily seen that this vector field is C^1 (see [6, p. 207]). Such a Hamiltonian is called a *Tonelli* (or *optical*) *Hamiltonian.*

Definition 3.2 (Tonelli Hamiltonian). *A function $H : T^*M \longrightarrow \mathbb{R}$ is called a Tonelli (or optical) Hamiltonian if:*

i) *H is of class C^2;*
ii) *H is strictly convex in each fiber in the C^2 sense, i.e., the second partial vertical derivative $\partial^2 H/\partial p^2(x,p)$ is positive definite, as a quadratic form, for any $(x,p)\in T^*M$;*
iii) *H is superlinear in each fiber, i.e.,*

$$\lim_{\|p\|_x\to+\infty} \frac{H(x,p)}{\|p\|_x} = +\infty\,.$$

Examples of Tonelli Hamiltonians.
Let us see what are the Hamiltonians associated to the Tonelli Lagrangians that we have introduced in the previous examples.

- **Riemannian Hamiltonians.** If $L(x,v)=\frac{1}{2}\|v\|_x^2$ is the Riemannian Lagrangian associated to a Riemannian metric g on M, the corresponding Hamiltonian will be

  $$H(x,p) = \frac{1}{2}\|p\|_x^2,$$

  where $\|\cdot\|$ represents - in this last expression - the induced norm on the cotangent bundle T^*M.
- **Mechanical Hamiltonians.** If $L(x,v)=\frac{1}{2}\|v\|_x^2+U(x)$ is a mechanical Lagrangian, the associated Hamiltonian is:

  $$H(x,p) = \frac{1}{2}\|p\|_x^2 - U(x).$$

  It is sometimes referred to as *mechanical energy*.
- **Mañé's Hamiltonians.** If X is a C^k vector field on M, with $k\geq 2$, and $L_X(x,v)=\|v-X(x)\|_x^2$ is the associated Mañé Lagrangian, one

can check that the corresponding Hamiltonian is given by:

$$H(x,p) = \frac{1}{2}\|p\|_x^2 + \langle p, X(x)\rangle \,.$$

Given a Hamiltonian one can consider the associated *Hamiltonian flow* Φ_t^H on T^*M. In local coordinates, this flow can be expressed in terms of the so-called *Hamilton's equations*:

$$\begin{cases} \dot{x}(t) = \frac{\partial H}{\partial p}(x(t), p(t)) \\ \dot{p}(t) = -\frac{\partial H}{\partial x}(x(t), p(t))\,. \end{cases}$$

We will denote by $X_H(x,p) := \left(\frac{\partial H}{\partial p}(x,p), -\frac{\partial H}{\partial x}(x,p)\right)$ the *Hamiltonian vector field* associated to H. This has a more intrinsic (geometric) definition in terms of the canonical symplectic structure ω on T^*M, which in local coordinates can be written as $dx \wedge dp$ (see for example [5]). In fact, X_H is the unique vector field that satisfies

$$\omega\left(X_H(x,p), \cdot\right) = d_x H(\cdot) \qquad \forall (x,p) \in T^*M.$$

For this reason, it is sometime called *symplectic gradient of* H. It is easy to check from both definitions that - only in the autonomous case - the Hamiltonian is a *prime integral of the motion*, *i.e.,* it is constant along the solutions of these equations.

Now, we would like to explain what is the relation between the Euler-Lagrange flow and the Hamiltonian one. It follows easily from the definition of Hamiltonian (and Legendre-Fenchel transform) that for each $(x,v) \in TM$ and $(x,p) \in T^*M$ the following inequality holds:

$$\langle p,\, v\rangle_x \leq L(x,v) + H(x,p)\,. \tag{1}$$

This is called *Fenchel inequality* (or *Legendre-Fenchel inequality*, see [22]) and it plays a crucial role in the study of Lagrangian and Hamiltonian dynamics and in the variational methods that we are going to describe. In particular, equality holds if and only if $p = \partial L/\partial v(x,v)$. One can therefore introduce the following diffeomorphism between TM and T^*M, known as *Legendre transform*:

$$\begin{array}{rccc} \mathcal{L}: & TM & \longrightarrow & T^*M \\ & (x,v) & \longmapsto & \left(x,\, \dfrac{\partial L}{\partial v}(x,v)\right)\,. \end{array} \tag{2}$$

Moreover, the following relation with the Hamiltonian holds:

$$H \circ \mathcal{L}(x,v) = \left\langle \frac{\partial L}{\partial v}(x,v),\, v \right\rangle_x - L(x,v)\,.$$

This diffeomorphism $\mathcal{L}$ represents a conjugacy between the two flows, namely the Euler-Lagrange flow on TM and the Hamiltonian flow on T^*M; in other words, the following diagram commutes:

$$\begin{tikzcd} TM \arrow[r, "\Phi_t^L"] \arrow[d, "\mathcal{L}"'] & TM \arrow[d, "\mathcal{L}"] \\ T^*M \arrow[r, "\Phi_t^H"'] & T^*M \end{tikzcd}$$

Remark 3.3. Since $\mathcal{L}$ and the Hamiltonian flow Φ^H are both C^1, then it follows from the commutative diagram above that the Euler-Lagrange flow is also C^1.

4. Action-minimizing properties of integrable systems

Before entering into the details of Mather's work, we would like to discuss a very easy case: properties of invariant measures of an integrable system (see section 2). This will provide us with a better understanding of the ideas behind Mather's theory and will describe clearer in which sense these *action-minimizing sets* – namely, what we will call *Mather sets* (see section 5) – represent a generalization of KAM tori.

As we have already discussed in section 2, let $H : T^*\mathbb{T}^n \longrightarrow \mathbb{R}$ be an integrable Tonelli Hamiltonian in action-angle coordinates, *i.e.*, $H(x,p) = \mathfrak{h}(p)$ and let $L : T\mathbb{T}^n \longrightarrow \mathbb{R}$, $L(x,v) = \ell(v)$, be the associated Tonelli Lagrangian. We denote by $\Phi^{\mathfrak{h}}$ and Φ^{ℓ} the respective flows, by $\mathcal{L}$ the associated Legendre transform, and identify both $T^*\mathbb{T}^n$ and $T\mathbb{T}^n$ with $\mathbb{T}^n \times \mathbb{R}^n$.

We have recalled in section 3 that the Euler-Lagrange flow can be equivalently defined in terms of a variational principle associated to the *Lagrangian action functional* A_ℓ. We would like to study action-minimizing properties of these invariant manifolds; for, it is much better to work in the Lagrangian setting. Moreover, instead of considering properties of single orbits, it would be more convenient to study "collection" of orbits, in the form of *invariant probability measures*[4] and consider their *average action*. If μ is an invariant probability measure for Φ^ℓ – *i.e.*, $(\Phi_t^\ell)^*\mu = \mu$ for all $t \in \mathbb{R}$, where $(\Phi_t^\ell)^*\mu$ denotes the pull-back of the measure – then we define:

$$A_\ell(\mu) := \int_{T\mathbb{T}^n} \ell(v)\, d\mu.$$

Let us consider any invariant probability measure μ_0 supported on $\tilde{\Lambda}_{p_0} := \mathcal{L}^{-1}(\Lambda_{p_0})$ and compute its action. Observe that on the support of this measure $\ell(v) \equiv \ell(\rho(p_0))$. Then:

$$\begin{aligned} A_\ell(\mu_0) &= \int_{T\mathbb{T}^n} \ell(v)\, d\mu_0 = \int_{T\mathbb{T}^n} \ell(\rho(p_0))\, d\mu_0 = \\ &= \ell(p_0) = p_0 \cdot \rho(p_0) - \mathfrak{h}(p_0), \end{aligned} \tag{3}$$

[4] Actually, it is also possible study directly orbits. See Remark 5.8

where in the last step we have used the Legendre-Fenchel duality between h and ℓ.

Let us now consider a general invariant probability measure μ. In this case it is not true anymore that $\ell(v)$ is constant on the support of μ. However, using Legendre-Fenchel inequality (see (1)), we can conclude that $\ell(v) \geq p_0 \cdot v - \mathfrak{h}(p_0)$ for each $v \in \mathbb{R}^n$. Hence:

$$\begin{aligned} A_\ell(\mu) &= \int_{T\mathbb{T}^n} \ell(v)\, d\mu \geq \int_{T\mathbb{T}^n} \left(p_0 \cdot v - \mathfrak{h}(p_0)\right)\, d\mu \\ &= \int_{T\mathbb{T}^n} p_0 \cdot v\, d\mu - \mathfrak{h}(p_0) = p_0 \cdot \left(\int_{T\mathbb{T}^n} v\, d\mu\right) - \mathfrak{h}(p_0). \end{aligned} \tag{4}$$

We would like to compare expressions (3) and (4). However, in the case of a general measure, we do not know how to evaluate the term $\int_{T\mathbb{T}^n} v\, d\mu$. One possible trick to overcome this problem is the following: instead of considering the action of $\ell(v)$, let us consider the action of $\ell(v) - p_0 \cdot v$. It is easy to see that this new Lagrangian is also Tonelli (we have subtracted a linear term in v) and that it has the same Euler-Lagrange flow as ℓ. In this way we obtain from (3) and (4) that:

$$A_{\ell - p_0 \cdot v}(\mu_0) \;=\; -\mathfrak{h}(p_0) \qquad \text{and} \qquad A_{\ell - p_0 \cdot v}(\mu) \geq -\mathfrak{h}(p_0),$$

which are now comparable. Hence, we have just showed the following fact:

Fact 1: *Every invariant probability measure supported on $\tilde{\Lambda}_{p_0}$ minimizes the action $A_{\ell - p_0 \cdot v}$ amongst all invariant probability measures of Φ^ℓ.*

In particular, we can characterize our invariant tori in a different way:

$$\tilde{\Lambda}_{p_0} = \bigcup \{\operatorname{supp} \mu : \ \mu \text{ minimizes } A_{\ell - p_0 \cdot v}\}.$$

Moreover, there is a relation between the energy (Hamiltonian) of the invariant torus and the minimal action of its invariant probability measures:

$$\mathfrak{h}(p_0) = -\min\{A_{\ell - p_0 \cdot v}(\mu) : \ \mu \text{ is an inv. prob. measure}\}.$$

Observe that it is somehow expectable that we need to modify the Lagrangian in order to obtain information on a specific invariant torus. In fact, in the case of an integrable system we have a foliation of the space made by these invariant tori and it would be unrealistic to expect that they could all be obtained as extremals of the same action functional. In other words, what we did was to add a *weighting term* to our Lagrangian, in order to magnify some motions rather than others.

Is it possible to distinguish these motions in a different way? Let us go back to (3) and (4). The main problem in comparing these two expression was represented by the term $\int_{T\mathbb{T}^n} v\, d\mu$. This can be interpreted as a sort of average rotation vector of orbits in the support of μ. Hence, let us define the *average rotation vector of* μ as:

$$\rho(\mu) := \int_{T\mathbb{T}^n} v \, d\mu \in \mathbb{R}^n.$$

We will give a more precise definition of it (which is also meaningful on manifolds different from the torus) in section 5.

Let now μ be an invariant probability measure of Φ^ℓ with rotation vector $\rho(\mu) = \rho(p_0)$. It follows from (4) that:

$$\begin{aligned} A_\ell(\mu) &\geq p_0 \cdot \left(\int_{T\mathbb{T}^n} v \, d\mu \right) - \mathfrak{h}(p_0) = p_0 \cdot \rho(\mu) - \mathfrak{h}(p_0) = \\ &= p_0 \cdot \rho(p_0) - \mathfrak{h}(p_0) = \ell(\rho(p_0)). \end{aligned}$$

Therefore, comparing with (3) we obtain another characterization of μ_0:

Fact 2: *Every invariant probability measure supported on $\tilde{\Lambda}_{p_0}$ minimizes the action A_ℓ amongst all invariant probability measures of Φ^ℓ with rotation vector $\rho(p_0)$.*

In particular:

$$\tilde{\Lambda}_{p_0} = \bigcup \{\operatorname{supp} \mu : \ \mu \text{ minimizes } A_\ell \text{ amongst measures with rot. vect. } \rho(p_0)\}.$$

Moreover, there is a relation between the value of the Lagrangian at $\rho(p_0)$ and the minimal action of all invariant probability measures with rotation vector $\rho(p_0)$:

$$\ell(\rho(p_0)) = \min\{A_\ell(\mu) : \ \mu \text{ is an inv. prob. meas. with rot. vect. } \rho(p_0)\}.$$

Remark 4.1. One could also study directly orbits on these tori and try to show that their action minimizes a modified Lagrangian action, in the same spirit as we have just discussed for measures. See [23] and Remark 5.8 for more details.

5. Mather's theory for Tonelli Lagrangian systems

In this section we describe Mather's theory for general Tonelli Lagrangians on compact manifolds. As we have already said before, we refer the reader to [23] for all the proofs and for a more detailed presentation of this theory.

Let $\mathfrak{M}(L)$ be the space of probability measures μ on TM that are invariant under the Euler-Lagrange flow of L and such that $\int_{TM} \|v\| \, d\mu < \infty$. We will hereafter assume that $\mathfrak{M}(L)$ is endowed with the *vague topology*, *i.e.,* the weak*–topology induced by the space C^0_ℓ of continuous functions $f : TM \longrightarrow \mathbb{R}$ having at most linear growth:

$$\sup_{(x,v)\in TM} \frac{|f(x,v)|}{1 + \|v\|} < +\infty \,.$$

One can check that $\mathfrak{M}(L) \subset \left(C^0_\ell\right)^*$.

In the case of an autonomous Tonelli Lagrangian, it is easy to see that $\mathfrak{M}(L)$ is non-empty (actually it contains infinitely many measures with distinct supports). In fact, recall that because of the conservation of the energy $E(x,v) := H \circ \mathcal{L}(x,v) = \left\langle \frac{\partial L}{\partial v}(x,v),\, v\right\rangle_x - L(x,v)$ along the orbits, each energy level of E is compact (it follows from the superlinearity condition) and invariant under Φ^L_t. It is a classical result in ergodic theory (sometimes called Kryloff–Bogoliouboff theorem) that a flow on a compact metric space has at least an invariant probability measure, which belongs indeed to $\mathfrak{M}(L)$.

To each $\mu \in \mathfrak{M}(L)$, we may associate its *average action*:

$$A_L(\mu) = \int_{TM} L\, d\mu\,.$$

The action functional $A_L : \mathfrak{M}(L) \longrightarrow \mathbb{R}$ is lower semicontinuous with the vague topology on $\mathfrak{M}(L)$ (this functional might not be necessarily continuous, see [8, Remark 2-3.4]). In particular, this implies that there exists $\mu \in \mathfrak{M}(L)$, which minimizes A_L over $\mathfrak{M}(L)$.

Definition 5.1. *A measure $\mu \in \mathfrak{M}(L)$, such that $A_L(\mu) = \min_{\mathfrak{M}(L)} A_L$, is called an action-minimizing measure of L.*

As we have already seen in section 4, by modifying the Lagrangian (without changing the Euler-Lagrange flow) one can find many other interesting measures besides those found by minimizing A_L. A similar idea can be implemented for a general Tonelli Lagrangian. Observe, in fact, that if η is a 1-form on M, we can interpret it as a function on the tangent bundle (linear on each fiber)

$$\begin{array}{rcl} \hat{\eta} : TM & \longrightarrow & \mathbb{R} \\ (x,v) & \longmapsto & \langle \eta(x),\, v\rangle_x \end{array}$$

and consider a new Tonelli Lagrangian $L_\eta := L - \hat{\eta}$. The associated Hamiltonian will be given by $H_\eta(x,p) = H(x, \eta(x) + p)$.

Observe that:

i) If η is closed, then L and L_η have the same Euler-Lagrange flow on TM. See [18].
ii) If $\mu \in \mathfrak{M}(L)$ and $\eta = df$ is an exact 1-form, then $\int \widehat{df}\, d\mu = 0$. Thus, for a fixed L, the minimizing measures will depend only on the de Rham cohomology class $c = [\eta] \in H^1(M;\mathbb{R})$.

Therefore, instead of studying the action minimizing properties of a single Lagrangian, one can consider a family of such "modified" Lagrangians, parameterized over $H^1(M;\mathbb{R})$. Hereafter, for any given $c \in H^1(M;\mathbb{R})$, we will denote by η_c a closed 1-form with that cohomology class.

Definition 5.2. *Let η_c be a closed 1-form of cohomology class c. Then, if $\mu \in \mathfrak{M}(L)$ minimizes $A_{L_{\eta_c}}$ over $\mathfrak{M}(L)$, we will say that μ is a c-action minimizing measure (or c-minimal measure, or Mather measure with cohomology c).*

Compare with Fact 1 in section 4.

Remark 5.3. Observe that the cohomology class of an action-minimizing invariant probability measure is not intrinsic in the measure itself nor in the dynamics, but it depends on the specific choice of the Lagrangian L. Changing the Lagrangian by a closed 1-form η, *i.e.,* $L \longmapsto L - \eta$, we will change all the cohomology classes of its action minimizing measures by $-[\eta] \in H^1(M;\mathbb{R})$. Compare also with Remark 5.5 (*ii*).

One can consider the following function on $H^1(M;\mathbb{R})$ (the minus sign is introduced for a convention that will probably become clearer later on):

$$\begin{array}{rcl} \alpha : H^1(M;\mathbb{R}) & \longrightarrow & \mathbb{R} \\ c & \longmapsto & -\min_{\mu \in \mathfrak{M}(L)} A_{L_{\eta_c}}(\mu)\,. \end{array}$$

This function α is well-defined (it does not depend on the choice of the representatives of the cohomology classes) and it is easy to see that it is convex. This is generally known as *Mather's α-function.* We have seen in section 4 that for an integrable Hamiltonian $H(x,p) = \mathfrak{h}(p)$, $\alpha(c) = \mathfrak{h}(c)$. For this and several other reasons that we will see later on, this function is sometimes called *effective Hamiltonian.* In particular, it can be proven that $\alpha(c)$ is related to the energy level containing such c-action minimizing measures [7].

We will denote by $\mathfrak{M}_c(L)$ the subset of c-action minimizing measures:

$$\mathfrak{M}_c := \mathfrak{M}_c(L) = \{\mu \in \mathfrak{M}(L) :\ A_{L_{\eta_c}}(\mu) = -\alpha(c)\}.$$

We can now define a first important family of invariant sets: the *Mather sets.*

Definition 5.4. *For a cohomology class $c \in H^1(M;\mathbb{R})$, we define the Mather set of cohomology class c as:*

$$\widetilde{\mathcal{M}}_c := \bigcup_{\mu \in \mathfrak{M}_c} \operatorname{supp}\mu \subset TM\,. \tag{5}$$

The projection on the base manifold $\mathcal{M}_c = \pi\left(\widetilde{\mathcal{M}}_c\right) \subseteq M$ is called projected Mather set (with cohomology class c).

Properties of this set:

i) It is non-empty, compact and invariant [18].
ii) It is contained in the energy level corresponding to $\alpha(c)$ [7].
iii) In [18] Mather proved the celebrated *graph theorem*:
Let $\pi : TM \longrightarrow M$ denote the canonical projection. Then, $\pi|\widetilde{\mathcal{M}}_c$ is an injective mapping of $\widetilde{\mathcal{M}}_c$ into M, and its inverse $\pi^{-1} : \mathcal{M}_c \longrightarrow \widetilde{\mathcal{M}}_c$ is Lipschitz.

Now, we would like to shift our attention to a related problem. As we have seen in section 4, instead of considering different minimizing problems over $\mathfrak{M}(L)$, obtained by modifying the Lagrangian L, one can alternatively try to minimize the Lagrangian L by putting some constraint, such as, for instance, fixing the *rotation vector* of the measures. In order to generalize this to Tonelli Lagrangians on compact manifolds, we first need to define what we mean by rotation vector of an invariant measure.

Let $\mu \in \mathfrak{M}(L)$. Thanks to the superlinearity of L, the integral $\int_{TM} \hat{\eta} d\mu$ is well defined and finite for any closed 1-form η on M. Moreover, if η is exact, then this integral is zero, *i.e.,* $\int_{TM} \hat{\eta} d\mu = 0$. Therefore, one can define a linear functional:

$$\begin{array}{rcl} H^1(M;\mathbb{R}) & \longrightarrow & \mathbb{R} \\ c & \longmapsto & \displaystyle\int_{TM} \hat{\eta} d\mu \,, \end{array}$$

where η is any closed 1-form on M with cohomology class c. By duality, there exists $\rho(\mu) \in H_1(M;\mathbb{R})$ such that

$$\int_{TM} \hat{\eta}\, d\mu = \langle c, \rho(\mu) \rangle \qquad \forall\, c \in H^1(M;\mathbb{R})$$

(the bracket on the right–hand side denotes the canonical pairing between cohomology and homology). We call $\rho(\mu)$ the *rotation vector* of μ. This rotation vector is the same as the Schwartzman's asymptotic cycle of μ (see [24] and [23] for more details).

Remark 5.5. (*i*) It is possible to provide a more geometric interpretation of this. Suppose for the moment that μ is ergodic. Then, it is known that a generic orbit $\gamma(t) := \pi \Phi_t^L(x,v)$, where $\pi : TM \longrightarrow M$ denotes the canonical projection, will return infinitely often close (as close as we like) to its initial point $\gamma(0) = x$. We can therefore consider a sequence of times $T_n \to +\infty$ such that $d(\gamma(T_n), x) \to 0$ as $n \to +\infty$, and consider the closed loops σ_n obtained by closing $\gamma|[0,T_n]$ with the shortest geodesic connecting $\gamma(T_n)$ to x. Denoting by $[\sigma_n]$ the homology class of this loop, one can verify (see [24]) that $\lim_{n\to\infty} \frac{[\sigma_n]}{T_n} = \rho(\mu)$, independently of the chosen sequence $\{T_n\}_n$. In other words, in the case of ergodic measures, the rotation vector tells us how on average a generic orbit winds around TM. If μ is not ergodic, $\rho(\mu)$ loses this neat geometric meaning, yet it may be interpreted as the average of the rotation vectors of its different ergodic components.

(*ii*) It is clear from the discussion above that the rotation vector of an invariant measure depends only on the dynamics of the system (*i.e.,* on the Euler-Lagrange flow) and not on the chosen Lagrangian. Therefore, it does not change when we modify our Lagrangian by adding a closed one form.

Using that the action functional $A_L : \mathfrak{M}(L) \longrightarrow \mathbb{R}$ is lower semicontinuous, one can prove that the map $\rho : \mathfrak{M}(L) \longrightarrow H_1(M;\mathbb{R})$ is continuous and surjective, *i.e.*, for every $h \in H_1(M;\mathbb{R})$ there exists $\mu \in \mathfrak{M}(L)$ with $A_L(\mu) < \infty$ and $\rho(\mu) = h$ (see [18]).

Following Mather [18], let us consider the minimal value of the average action A_L over the probability measures with rotation vector h. Observe that this minimum is actually achieved because of the lower semicontinuity of A_L and the compactness of $\rho^{-1}(h)$ (ρ is continuous and L superlinear).

Let us define

$$\begin{array}{rcl} \beta : H_1(M;\mathbb{R}) & \longrightarrow & \mathbb{R} \\ h & \longmapsto & \min_{\mu\in\mathfrak{M}(L):\,\rho(\mu)=h} A_L(\mu)\,. \end{array} \tag{6}$$

This function β is what is generally known as *Mather's β-function* and it is immediate to check that it is convex. We have seen in section 4 that if we have an integrable Tonelli Hamiltonian $H(x,p) = \mathfrak{h}(p)$ and the associated Lagrangian $L(x,v) = \ell(v)$, then $\beta(h) = \ell(h)$. For this and several other reasons, this function is sometime called *effective Lagrangian*.

We can now define what we mean by action minimizing measure with a given rotation vector.

Definition 5.6. *A measure $\mu \in \mathfrak{M}(L)$ realizing the minimum in (6), i.e., such that $A_L(\mu) = \beta(\rho(\mu))$, is called an action minimizing (or minimal, or Mather) measure with rotation vector $\rho(\mu)$.*

Compare with Fact 2 in section 4.

We will denote by $\mathfrak{M}^h(L)$ the subset of action minimizing measures with rotation vector h:

$$\mathfrak{M}^h := \mathfrak{M}^h(L) = \{\mu \in \mathfrak{M}(L) :\ \rho(\mu) = h \text{ and } A_L(\mu) = \beta(h)\}.$$

This allows us to define another important familty of invariant sets.

Definition 5.7. *For a homology class (or rotation vector) $h \in H_1(M;\mathbb{R})$, we define the Mather set corresponding to a rotation vector h as*

$$\widetilde{\mathcal{M}}^h := \bigcup_{\mu\in\mathfrak{M}^h} \operatorname{supp}\mu \subset TM\,, \tag{7}$$

and the projected one as $\mathcal{M}^h = \pi\left(\widetilde{\mathcal{M}}^h\right) \subseteq M$.

Similarly to what we have already seen above, this set satisfies the following properties:

i) It is non-empty, compact and invariant.
ii) It is contained in a given energy level.
iii) It also satisfies the *graph theorem*:
let $\pi : TM \longrightarrow M$ denote the canonical projection. Then, $\pi|\widetilde{\mathcal{M}}^h$ is an injective mapping of $\widetilde{\mathcal{M}}^h$ into M, and its inverse $\pi^{-1} : \mathcal{M}^h \longrightarrow \widetilde{\mathcal{M}}^h$ is Lipschitz.

Remark 5.8. (i) In the above discussion we have only discussed properties of invariant probability measures associated to the system. Actually, one could study directly orbits of the systems and look for orbits that globally minimize the action of a modified Lagrangian (in the same spirit as before). This would lead to the definition of two other families of invariant compact sets, the *Aubry sets* $\widetilde{\mathcal{A}}_c$ and the *Mañé sets* $\widetilde{\mathcal{N}}_c$, which are also parameterized by $H^1(M;\mathbb{R})$ (the parameter which describes the modification of the

Lagrangian, exactly in the same way as before). For a given $c \in H^1(M;\mathbb{R})$, these sets contain the Mather set $\widetilde{\mathcal{M}}_c$, and this inclusion may be strict. In fact, while the motion on the Mather sets is *recurrent* (it is the union of the supports of invariant probability measures), the Aubry and the Mañé sets may contain non-recurrent orbits as well.
(*ii*) Differently from what happens with invariant probability measures, it will not be always possible to find *action-minimizing orbits* for any given rotation vector (not even possible to define a rotation vector for every action minimizing orbit). For instance, an example due to Hedlund [10] provides the existence of a Riemannian metric on a three-dimensional torus, for which minimal geodesics exist only in three directions. The same construction can be extended to any dimension larger than three.

6. Mather's α and β-functions

The discussion in section 5 led to two equivalent formulations of the minimality of an invariant probability measure μ:

- there exists a homology class $h \in \mathrm{H}_1(M;\mathbb{R})$, namely its rotation vector $\rho(\mu)$, such that μ minimizes A_L amongst all measures in $\mathfrak{M}(L)$ with rotation vector h, *i.e.,* $A_L(\mu) = \beta(h)$.
- There exists a cohomology class $c \in \mathrm{H}^1(M;\mathbb{R})$, such that μ minimizes $A_{L_{\eta_c}}$ amongst all probability measures in $\mathfrak{M}(L)$, *i.e.,* $A_{L_{\eta_c}}(\mu) = -\alpha(c)$.

What is the relation between two these different approaches? Are they equivalent, *i.e.,* $\bigcup_{h\in \mathrm{H}_1(M;\mathbb{R})} \mathfrak{M}^h = \bigcup_{c\in \mathrm{H}^1(M;\mathbb{R})} \mathfrak{M}_c$?

In order to comprehend the relation between these two families of action-minimizing measures, we need to understand better the properties of the these two functions that we have introduced above:

$$\alpha : H^1(M;\mathbb{R}) \longrightarrow \mathbb{R} \quad \text{and} \quad \beta : H_1(M;\mathbb{R}) \longrightarrow \mathbb{R}.$$

Let us start with the following trivial remark.

Remark 6.1. As we have previously pointed out, if we have an integrable Tonelli Hamiltonian $H(x,p) = \mathfrak{h}(p)$ and the associated Lagrangian $L(x,v) = \ell(v)$, then $\alpha(c) = \mathfrak{h}(c)$ and $\beta(h) = \ell(h)$. In this case, the cotangent bundle $T^*\mathbb{T}^n$ is foliated by invariant tori $\mathcal{T}_c^* := \mathbb{T}^n \times \{c\}$ and the tangent bundle $T\mathbb{T}^n$ by invariant tori $\widetilde{\mathcal{T}}^h := \mathbb{T}^n \times \{h\}$. In particular, we proved that

$$\widetilde{\mathcal{M}}_c = \mathcal{L}^{-1}(\mathcal{T}_c) = \widetilde{\mathcal{T}}^h = \widetilde{\mathcal{M}}^h,$$

where h and c are such that $h = \nabla\mathfrak{h}(c) = \nabla\alpha(c)$ and $c = \nabla\ell(h) = \nabla\beta(h)$.

We would like to investigate whether a similar relation linking Mather sets of a certain cohomology class to Mather sets with a certain rotation vector, continues to exist beyond the specificity of this situation. Of course, one main difficulty is that in general the *effective Hamiltonian* α and the *effective Lagrangian* β, although being convex and superlinear (see Proposition 6.2), are not necessarily differentiable.

Before stating the main relation between these two functions, let us recall some definitions and results from classical convex analysis (see [22]). Given a convex function $\varphi : V \longrightarrow \mathbb{R} \cup \{+\infty\}$ on a finite dimensional vector space V, one can consider a *dual* (or *conjugate*) function defined on the dual space V^*, via the so-called *Fenchel transform*: $\varphi^*(p) := \sup_{v \in V} \big(p \cdot v - \varphi(v)\big)$. In our case, the following holds.

Proposition 6.2. *α and β are convex conjugate, i.e., $\alpha^* = \beta$ and $\beta^* = \alpha$. In particular, it follows that α and β have superlinear growth.*

Next proposition will allow us to clarify the relation (and duality) between the two minimizing procedures described above. To state it, recall that, like any convex function on a finite-dimensional space, β admits a subderivative at each point $h \in H_1(M;\mathbb{R})$, *i.e.,* we can find $c \in H^1(M;\mathbb{R})$ such that

$$\forall h' \in H_1(M;\mathbb{R}), \quad \beta(h') - \beta(h) \geq \langle c, h' - h \rangle.$$

As it is usually done, we will denote by $\partial\beta(h)$ the set of $c \in H^1(M;\mathbb{R})$ that are subderivatives of β at h, *i.e.,* the set of c's which satisfy the above inequality. Similarly, we will denote by $\partial\alpha(c)$ the set of subderivatives of α at c. Actually, Fenchel's duality implies an easier characterization of subdifferentials: $c \in \partial\beta(h)$ *if and only if* $\langle c, h\rangle = \alpha(c) + \beta(h)$ (similarly for $h \in \partial\alpha(c)$).

We can now state precisely in which sense what observed in Remark 6.1 continues to hold in the general case

Proposition 6.3. *Let $\mu \in \mathfrak{M}(L)$ be an invariant probability measure. Then:*
(i) *$A_L(\mu) = \beta(\rho(\mu))$ if and only if there exists $c \in \mathrm{H}^1(M;\mathbb{R})$ such that μ minimizes $A_{L_{\eta_c}}$ (i.e., $A_{L_{\eta_c}}(\mu) = -\alpha(c)$).*
(ii) *If μ satisfies $A_L(\mu) = \beta(\rho(\mu))$ and $c \in H^1(M;\mathbb{R})$, then μ minimizes $A_{L_{\eta_c}}$ if and only if $c \in \partial\beta(\rho(\mu))$ (or equivalently $\langle c, h\rangle = \alpha(c) + \beta(\rho(\mu))$).*

Remark 6.4. (*i*) It follows from the above proposition that both minimizing procedures lead to the same sets of invariant probability measures:

$$\bigcup_{h \in \mathrm{H}_1(M;\mathbb{R})} \mathfrak{M}^h = \bigcup_{c \in \mathrm{H}^1(M;\mathbb{R})} \mathfrak{M}_c\,.$$

In other words, minimizing over the set of invariant measures with a fixed rotation vector or globally minimizing the modified Lagrangian (corresponding to a certain cohomology class) are dual problems, as the ones that often appears in linear programming and optimization. In some sense, modifying the Lagrangian by a closed 1-form is analog to the method of Lagrange multipliers for searching constrained critical points of a function.

(*ii*) In particular we have the following inclusions between Mather sets:

$$c \in \partial\beta(h) \quad \Longleftrightarrow \quad h \in \partial\alpha(c) \quad \Longleftrightarrow \quad \widetilde{\mathcal{M}}^h \subseteq \widetilde{\mathcal{M}}_c\,.$$

Moreover, for any $c \in H^1(M;\mathbb{R})$:

$$\widetilde{\mathcal{M}}_c = \bigcup_{h \in \partial\alpha(c)} \widetilde{\mathcal{M}}^h\,.$$

Observe that the non-differentiability of α at some c produces the presence in $\widetilde{\mathcal{M}}_c$ of (ergodic) invariant probability measures with different rotation vectors. On the other hand, the non-differentiability of β at some h implies that there exist $c \neq c'$ such that $\widetilde{\mathcal{M}}_c \cap \widetilde{\mathcal{M}}_{c'} \neq \emptyset$ (compare with the integrable case discussed in section 4, where these phenomena do not appear).

(*iii*) The minimum of the α-function is sometime called *Mañé's strict critical value*. Observe that if $\alpha(c_0) = \min \alpha(c)$, then $0 \in \partial\alpha(c_0)$ and $\beta(0) = -\alpha(c_0)$. Therefore, the measures with zero homology are contained in the least possible energy level containing Mather sets and $\widetilde{\mathcal{M}}^0 \subseteq \widetilde{\mathcal{M}}_{c_0}$. This inclusion might be strict, unless α is differentiable at c_0; in fact, there may be other action minimizing measures with non-zero rotation vectors corresponding to the other subderivatives of α at c_0.

(iv) Note that measures of trivial homology are not necessarily supported on orbits with trivial homology or fixed points. For instance, one can consider the following example. Let $M = \mathbb{T}^2$ equipped with the flat metric and consider a vector field X with norm 1 and such that X has two closed orbits γ_1 and γ_2 in opposite (non-trivial) homology classes and any other orbit asymptotically approaches γ_1 in forward time and γ_2 in backward time; for example one can consider $X(x_1, x_2) = (\cos(2\pi x_1), \sin(2\pi x_1))$, where $(x_1, x_2) \in \mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$.

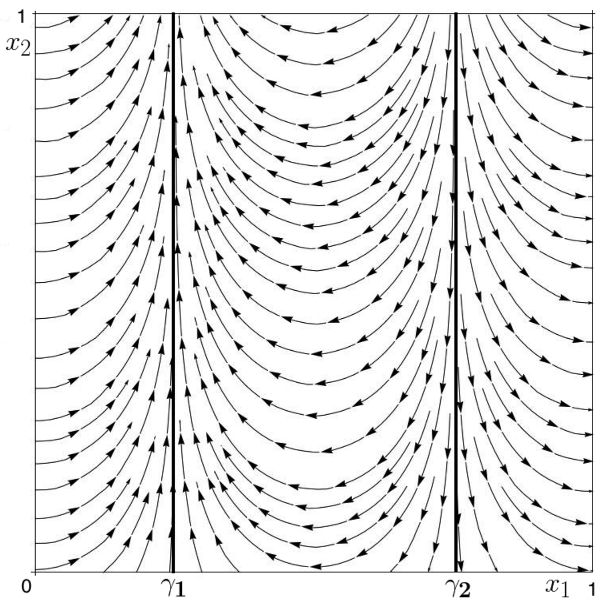

FIGURE 2. Plot of the vector field X.

As we have described in section 3, we can embed this vector field into the Euler-Lagrange vector field given by the Tonelli Lagrangian $L_X(x, v) = \frac{1}{2}\|v - X(x)\|^2$. Let us now consider the probability measure μ_{γ_1} and μ_{γ_2}, uniformly distributed respectively on $(\gamma_1, \dot{\gamma}_1)$ and $(\gamma_2, \dot{\gamma}_2)$. Since these two

curves have opposite homologies, then $\rho(\mu_{\gamma_1}) = -\rho(\mu_{\gamma_2}) =: h_0 \neq 0$. Moreover, it is easy to see that $A_{L_X}(\mu_{\gamma_1}) = A_{L_X}(\mu_{\gamma_2}) = 0$, since the Lagrangian vanishes on Graph(X). Using the fact that $L_X \geq 0$ (in particular it is strictly positive outside of Graph(X)) and that there are no other invariant ergodic probability measures contained in Graph(X), we can conclude that $\mathcal{M}_0 = \gamma_1 \cup \gamma_2$ and $\alpha(0) = 0$. Moreover, $\mu_0 := \frac{1}{2}\mu_{\gamma_1} + \frac{1}{2}\mu_{\gamma_2}$ has zero homology and its support is contained in $\widetilde{\mathcal{M}}_0$. Therefore (see Proposition 6.3 (i)), μ_0 is action minimizing with rotation vector 0 and $\widetilde{\mathcal{M}}^0 \subseteq \widetilde{\mathcal{M}}_0$; in particular, $\widetilde{\mathcal{M}}^0 = \widetilde{\mathcal{M}}_0$. This also implies that $\beta(0) = 0$ and $\alpha(0) = \min \alpha(c) = 0$.
Observe that α is not differentiable at 0. In fact, reasoning as we have done before for the zero homology class, it is easy to see that for all $t \in [-1, 1]$ $\widetilde{\mathcal{M}}^{th_0} = \widetilde{\mathcal{M}}_0$. It is sufficient to consider the convex combination $\mu_\lambda = \lambda\mu_{\gamma_1} + (1-\lambda)\mu_{\gamma_2}$ for any $\lambda \in [0, 1]$. Therefore, $\partial\alpha(0) = \{th_0,\ t \in [-1, 1]\}$ and $\beta(th_0) = 0$ for all $t \in [-1, 1]$.

As we have just seen in item (*iv*) of Remark 6.4, it may happen that the Mather sets corresponding to different homology (resp. cohomology) classes coincide or are included one into the other. This is something that, for instance, cannot happen in the integrable case: in this situation, in fact, these sets form a foliation and are disjoint. The problem in the above mentioned example, seems to be related to a lack of *strict convexity* of β and α. See also the discussion on the simple pendulum in section 7: in this case the Mather sets, corresponding to a non-trivial interval of cohomology classes about 0, coincide.
In the light of this, let us try to understand better what happens when α and β are not strictly convex, *i.e.,* when we are in the presence of *flat* pieces. Let us first fix some notation. If V is a real vector space and $v_0, v_1 \in V$, we will denote by $\sigma(v_0, v_1)$ the segment joining v_0 to v_1, that is $\sigma(v_0, v_1) := \{tv_0 + (1-t)v_1 :\ t \in [0, 1]\}$. We will say that a function $f : V \longrightarrow \mathbb{R}$ is *affine* on $\sigma(v_0, v_1)$, if there exists $v^* \in V^*$ (the dual of V), such that $f(v) = f(v_0) + \langle v^*, v - v_0\rangle$ for each $v \in \sigma(v_0, v_1)$. Moreover, we will denote by $\mathrm{Int}(\sigma(v_0, v_1))$ the *interior* of $\sigma(v_0, v_1)$, *i.e.,* $\mathrm{Int}(\sigma(v_0, v_1)) := \{tv_0 + (1-t)v_1 : t \in (0, 1)\}$.

Proposition 6.5. (i) *Let $h_0, h_1 \in H_1(M; \mathbb{R})$; β is affine on $\sigma(h_0, h_1)$ if and only if for any $h \in \mathrm{Int}(\sigma(h_0, h_1))$ we have $\widetilde{\mathcal{M}}^h \supseteq \widetilde{\mathcal{M}}^{h_0} \cup \widetilde{\mathcal{M}}^{h_1}$.*
(ii) *Let $c_0, c_1 \in H^1(M; \mathbb{R})$; α is constant on $\sigma(c_0, c_1)$ if and only if for any $c \in \mathrm{Int}(\sigma(c_0, c_1))$ we have $\widetilde{\mathcal{M}}_c \subseteq \widetilde{\mathcal{M}}_{c_0} \cap \widetilde{\mathcal{M}}_{c_1}$.*

Remark 6.6. The inclusion in Proposition 6.5 (i) may not be true at the end points of σ. For instance, Remark 6.4 (iv) provides an example in which the inclusion in Proposition 6.5 (i) is not true at the end-points of $\sigma(-h_0, h_0)$.

Remark 6.7. It follows from the previous remarks and Proposition 6.5, that, in general, the action minimizing measures (and consequently the Mather sets $\widetilde{\mathcal{M}}_c$ or $\widetilde{\mathcal{M}}^h$) are not necessarily ergodic. Recall that an invariant probability measure is said to be *ergodic*, if all invariant Borel sets have measure 0 or 1. These measures play a special role in the study of the dynamics of the system, therefore one could ask what are the ergodic action-minimizing measures. It is a well-known result from ergodic theory, that the ergodic measures of a flow correspond to the *extremal points* of the set of invariant probability measures, where by extremal point of a convex set, we mean an element that cannot be obtained as a non-trivial convex combination of other elements of the set. Since β has superlinear growth, its epigraph $\{(h,t) \in H_1(M;\mathbb{R}) \times \mathbb{R} : \ t \geq \beta(h)\}$ has infinitely many extremal points. Let $(h, \beta(h))$ denote one of these extremal points. Then, there exists at least one ergodic action minimizing measure with rotation vector h. It is in fact sufficient to consider any extremal point of the set $\{\mu \in \mathfrak{M}^h(L) : \ A_L(\mu) = \beta(h)\}$: this measure will be an extremal point of $\mathfrak{M}(L)$ and hence ergodic. Moreover, as we have already recalled in Remark 5.5, for such an ergodic measure μ, Birkhoff's ergodic theorem implies that for μ-almost every initial datum, the corresponding trajectory has rotation vector h.

7. An example: the simple pendulum

In this section we would like to describe the Mather sets, the α-function and the β-function, in a specific example: the *simple pendulum.* This system can be described in terms of the Lagrangian:

$$\begin{array}{rcl} L : T\mathbb{T} & \longrightarrow & \mathbb{R} \\ (x,v) & \longmapsto & \dfrac{1}{2}|v|^2 + \big(1-\cos(2\pi x)\big). \end{array}$$

It is easy to check that the Euler-Lagrange equation provides exactly the equation of the pendulum:

$$\dot{v} = 2\pi \sin(2\pi x) \qquad \Longleftrightarrow \qquad \left\{ \begin{array}{l} v = \dot{x} \\ \ddot{x} - 2\pi\sin(2\pi x) = 0. \end{array} \right.$$

The associated Hamiltonian (or energy) $H : T^*\mathbb{T} \longrightarrow \mathbb{R}$ is given by $H(x,p) := \frac{1}{2}|p|^2 - (1-\cos(2\pi x))$. Observe that in this case the Legendre transform is $(x,p) = \mathcal{L}_L(x,v) = (x,v)$, therefore we can easily identify the tangent and cotangent bundles. In the following we will consider $T\mathbb{T} \simeq T^*\mathbb{T} \simeq \mathbb{T}\times\mathbb{R}$ and identify $H^1(M;\mathbb{R}) \simeq H_1(M;\mathbb{R}) \simeq \mathbb{R}$.

First of all, let us study what are the invariant probability measures of this system.

- Observe that $(0,0)$ and $(\frac{1}{2},0)$ are fixed points for the system (respectively *unstable* and *stable*). Therefore, the Dirac measures concentrated on each of them are invariant probability measures. Hence, we have found two first invariant measures: $\delta_{(0,0)}$ and $\delta_{(\frac{1}{2},0)}$, both with zero rotation vector: $\rho(\delta_{(0,0)}) = \rho(\delta_{(\frac{1}{2},0)}) = 0$. As far as their energy is concerned (*i.e.,* the energy levels in which they are contained), it is easy to check that $E(\delta_{(0,0)}) = H(0,0) = 0$ and

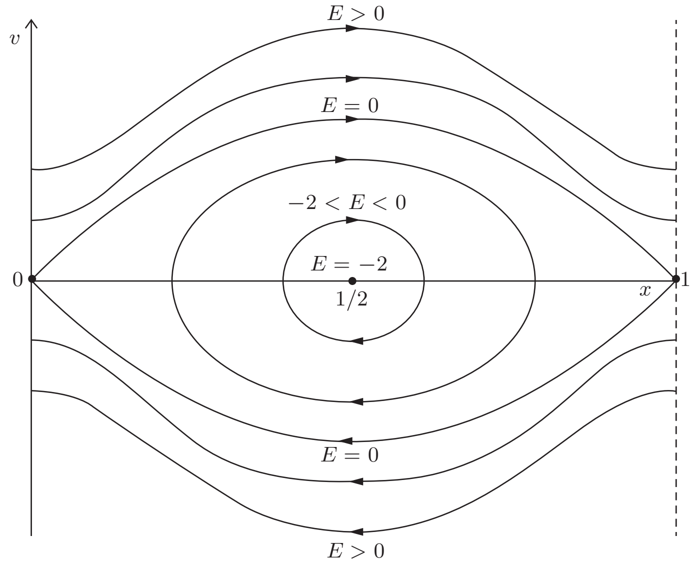

FIGURE 3. The phase space of the simple pendulum.

$E(\delta_{(\frac{1}{2},0)}) = H(\frac{1}{2},0) = -2$. Observe that these two energy levels cannot contain any other invariant probability measure.

- If $E > 0$, then the energy level $\{H(x,v) = E\}$ consists of two homotopically non-trivial periodic orbits (*rotation motions*):

$$\mathcal{P}_E^{\pm} := \{(x,v) :\ v = \pm\sqrt{2[(1+E) - \cos(2\pi x)]},\ \forall\, x \in \mathbb{T}\}.$$

The probability measures evenly distributed along these orbits – which we will denote $\mu_E^{\pm}$ – are invariant probability measures of the system. If we denote by

(8) $$T(E) := \int_0^1 \frac{1}{\sqrt{2[(1+E) - \cos(2\pi x)]}}\, dx$$

the period of such orbits, then it is easy to check that $\rho(\mu_E^{\pm}) = \frac{\pm 1}{T(E)}$ (see Remark 5.5). Observe that this function $T : (0,+\infty) \longrightarrow (0,+\infty)$, which associates to a positive energy E the period of the corresponding periodic orbits $\mathcal{P}_E^{\pm}$, is continuous and strictly decreasing. Moreover, $T(E) \to \infty$ as $E \to 0$ (it is easy to see this, by noticing that motions on the separatrices take an infinitely long time to connect 0 to $1 \equiv 0 \bmod \mathbb{Z}$). Therefore, $\rho(\mu_E^{\pm}) \to 0$ as $E \to 0$.

- If $-2 < E < 0$, then the energy level $\{H(x,v) = E\}$ consists of one contractible periodic orbit (*libration motion*):

$$\mathcal{P}_E := \{(x,v) :\ v^2 = 2(1+E) - 2\cos(2\pi x), \quad x \in [x_E, 1 - x_E]\},$$

where $x_E := \frac{1}{2\pi}\arccos(1+E)$. The probability measure evenly distributed along this orbit – which we will denote by μ_E – is an invariant probability measure of the system. Moreover, since this orbit is contractible, its rotation vector is zero: $\rho(\mu_E) = 0$.

The measures above are the only ergodic invariant probability measures of the system. Other invariant measures can be easily obtained as a convex combination of them.

Now we want to understand which of these are action-minimizing for some cohomology class.

Remark 7.1. (*i*) Let us start by remarking that for $-2 < E < 0$ the support of the measure μ_E is not a graph over $\mathbb{T}$, therefore it cannot be action-minimizing for any cohomology class, since otherwise it would violate Mather's graph theorems (see section 5). Therefore all action-minimizing measures will be contained in energy levels corresponding to energy bigger than zero. It follows from what said in sections 5 and 6 that $\alpha(c) \geq 0$ for all $c \in \mathbb{R}$.

(*ii*) Another interesting property of the α-function (in this specific case) is that it is an even function: $\alpha(c) = \alpha(-c)$ for all $c \in \mathbb{R}$. This is a consequence of the particular symmetry of the system, *i.e.,* $L(x,v) = L(x,-v)$. In fact, let us denote $\tau : \mathbb{T} \times \mathbb{R} \longrightarrow \mathbb{T} \times \mathbb{R}$, $(x,v) \longmapsto (x,-v)$ and observe that if μ is an invariant probability measure, then also $\tau^*\mu$ is still an invariant probability measure. Moreover, $\tau^*\mathfrak{M}(L) = \mathfrak{M}(L)$, where $\mathfrak{M}(L)$ denotes the set of all invariant probability measures of L. It is now sufficient to notice that for each $\mu \in \mathfrak{M}(L)$, $\int (L - c\cdot v)\, d\mu = \int (L + c\cdot v) d\tau^*\mu$, and hence conclude that

$$\alpha(c) = -\inf_{\mathfrak{M}(L)} \int (L - c\cdot v)\, d\mu = -\inf_{\mathfrak{M}(L)} \int (L + c\cdot v) d\tau^*\mu = \alpha(-c)\,.$$

(*iii*) It follows from the above symmetry and the convexity of α, that $\min_{\mathbb{R}} \alpha(c) = \alpha(0)$.

Let us now start by studying the 0-action minimizing measures, *i.e.,* invariant probability measures that minimize the action of L without any modification. Since $L(x,v) \geq 0$ for each $(x,v) \in \mathbb{T} \times \mathbb{R}$, then $A_L(\mu) \geq 0$ for all $\mu \in \mathfrak{M}(L)$. In particular, $A_L(\delta_{(0,0)}) = 0$, therefore $\delta_{(0,0)}$ is a 0-action minimizing measure and $\alpha(0) = 0$. Since there are not other invariant probability measures supported in the energy level $\{H(x,v) = 0\}$ (*i.e.,* on the separatrices), then we can conclude that:

$$\widetilde{\mathcal{M}}_0 = \{(0,0)\}\,.$$

Moreover, since $\alpha'(0) = 0$ (see Remark 7.1 (*iii*)), then it follows from Remark 6.4 that:

$$\widetilde{\mathcal{M}}^0 = \widetilde{\mathcal{M}}_0 = \{(0,0)\}\,.$$

On the other hand, this could be also deduced from the fact that the only other measures with rotation vector 0, cannot be action minimizing since they do not satisfy the graph theorem (Remark 7.1).

Now let us investigate what happens with other cohomology classes. A naïve observation is that since the α-function is superlinear and continuous, all energy levels for $E \geq 0$ must contain some Mather set; in other words, all energy levels $E \geq 0$ must be obtained as $\alpha(c)$, for some c.

Let $E > 0$ and consider the periodic orbit $\mathcal{P}_E^+$ and the invariant probability measure μ_E^+ evenly distributed on it. The graph of this orbit can be seen as the graph of a closed 1-form $\eta_E^+ := \sqrt{2[(1+E)-\cos(2\pi x)]}\,dx$, whose cohomology class is

$$c^+(E) := [\eta_E^+] = \int_0^1 \sqrt{2[(1+E)-\cos(2\pi x)]}\,dx, \tag{9}$$

which can be interpreted as the (signed) area between the curve and the positive x-semiaxis. This value is clearly continuous and strictly increasing with respect to E (for $E>0$) and as $E \to 0^+$:

$$c^+(E) \longrightarrow \int_0^1 \sqrt{2[1-\cos(2\pi x)]}\,dx = \frac{4}{\pi}\,.$$

Therefore, it defines an invertible function $c^+ : (0,+\infty) \longrightarrow (\frac{4}{\pi}, +\infty)$.
We want to prove that μ_E^+ is $c^+(E)$-action minimizing. The proof will be an imitation of what already seen for KAM tori in section 4.
Let us consider the Lagrangian $L_{\eta_E^+}(x,v) := L(x,v) - \eta_E^+(x)\cdot v$. Then, using Legendre-Fenchel inequality (1) (on the support of μ_E^+, because of our choice of η_E^+, this is indeed an equality):

$$\begin{aligned}\int L_{\eta_E^+}(x,v)d\mu_E^+ &= \int \big(L(x,v)-\eta_E^+(x)\cdot v\big)\,d\mu_E^+ = \\ &= \int -H(x,\eta_E^+(x))d\mu_E^+ = -E\,.\end{aligned}$$

Now, let ν be any other invariant probability measure and apply again the same procedure as above (warning: this time Legendre-Fenchel inequality is not an equality anymore!):

$$\begin{aligned}\int L_{\eta_E^+}(x,v)d\nu &= \int \big(L(x,v)-\eta_E^+(x)\cdot v\big)\,d\nu \geq \\ &\geq \int -H(x,\eta_E^+(x))d\nu = -E\,.\end{aligned}$$

Therefore, we can conclude that μ_E^+ is $c^+(E)$-action minimizing. Since it already projects over the whole $\mathbb{T}$, it follows from the graph theorem that it is the only one:

$$\widetilde{\mathcal{M}}_{c^+(E)} = \mathcal{P}_E^+ = \{(x,v):\ v = \sqrt{2[(1+E)-\cos(2\pi x)]},\ \forall\, x\in\mathbb{T}\}.$$

Furthermore, since $\rho(\mu_E^+) = \frac{1}{T(E)}$, then:

$$\widetilde{\mathcal{M}}^{\frac{1}{T(E)}} = \widetilde{\mathcal{M}}_{c^+(E)} = \mathcal{P}_E^+.$$

Similarly, one can consider the periodic orbit $\mathcal{P}_E^-$ and the invariant probability measure μ_E^- evenly distributed on it. The graph of this orbit can be seen as the graph of a closed 1-form $\eta_E^- := -\sqrt{2[(1+E)-\cos(2\pi x)]}\,dx = -\eta_E^+$, whose cohomolgy class is $c^-(E) = -c^+(E)$. Then (see also Remark 7.1 (*ii*)):

$$\widetilde{\mathcal{M}}_{c^-(E)} = \mathcal{P}_E^- = \{(x,v):\ v = -\sqrt{2[(1+E)-\cos(2\pi x)]},\ \forall\, x\in\mathbb{T}\},$$

and

$$\widetilde{\mathcal{M}}^{-\frac{1}{T(E)}} = \widetilde{\mathcal{M}}_{c^-(E)} = \mathcal{P}_E^-.$$

Note that this completes the study of the Mather sets for any given rotation vector, since

$$\rho(\mu_E^{\pm}) = \pm\frac{1}{T(E)} \stackrel{E\to+\infty}{\longrightarrow} \pm\infty \qquad \text{and} \qquad \rho(\mu_E^{\pm}) = \pm\frac{1}{T(E)} \stackrel{E\to 0^+}{\longrightarrow} 0\,.$$

What remains to study is what happens for non-zero cohomology classes in $[-\frac{4}{\pi}, \frac{4}{\pi}]$. The situation turns out to be quite easy. Observe that $\alpha(c^{\pm}(E)) = E$. Thefore, from the continuity of α it follows that (take the limit as $E \to 0$): $\alpha(\pm\frac{4}{\pi}) = 0$. Moreover, since α is convex and $\min \alpha(c) = \alpha(0) = 0$, then: $\alpha(c) \equiv 0$ on $[-\frac{4}{\pi}, \frac{4}{\pi}]$. Therefore, the corresponding Mather sets will lie in the zero energy level. From the above discussion, it follows that in this energy level there is a unique invariant probability measure, namely $\delta_{(0,0)}$, and consequently:

$$\widetilde{\mathcal{M}}_c = \{(0,0)\} \qquad \text{for all } -\frac{4}{\pi} \leq c \leq \frac{4}{\pi}.$$

Let us summarize what we have found so far. Recall that in (8) and (9) we have introduced these two functions: $T : (0, +\infty) \longrightarrow (0, +\infty)$ and $c^+ : (0, +\infty) \longrightarrow (\frac{4}{\pi}, +\infty)$ representing respectively the period and the cohomology (area below the curve) of the upper periodic orbit of energy E. These functions (for which we have an explicit formula in terms of E) are continuous and strictly monotone (respectively, decreasing and increasing). Therefore, we can define their inverses which provide the energy of the periodic orbit with period T (for all positive periods) or the energy of the periodic orbit with cohomology class c (for $|c| > \frac{4}{\pi}$). We will denote them $E(T)$ and $E(c)$ (observe that this last quantity is exactly $-\alpha(c)$). Then:

$$\widetilde{\mathcal{M}}_c = \begin{cases} \{(0,0)\} & \text{if } -\frac{4}{\pi} \leq c \leq \frac{4}{\pi} \\ \mathcal{P}^+_{E(c)} & \text{if } c > \frac{4}{\pi} \\ \mathcal{P}^-_{E(-c)} & \text{if } c < -\frac{4}{\pi} \end{cases}$$

and

$$\widetilde{\mathcal{M}}^h = \begin{cases} \{(0,0)\} & \text{if } h = 0 \\ \mathcal{P}^+_{E(\frac{1}{h})} & \text{if } h > 0 \\ \mathcal{P}^-_{E(-\frac{1}{h})} & \text{if } h < 0\,. \end{cases}$$

We can provide an expression for these functions in terms of the quantities introduced above:

$$\alpha(c) = \begin{cases} 0 & \text{if } -\frac{4}{\pi} \leq c \leq \frac{4}{\pi} \\ E(|c|) & \text{if } |c| > \frac{4}{\pi} \end{cases}$$

and

$$\beta(h) = \begin{cases} 0 & \text{if } h = 0 \\ c(E(\frac{1}{|h|}))|h| - E(\frac{1}{|h|}) & \text{if } h \neq 0\,. \end{cases}$$

Observe that the α-function is C^1. In fact, the only problem might be at $c = \pm\frac{4}{\pi}$, but also there it is differentiable, with derivative 0. If it were not differentiable, then there would exist a subderivative $h \neq 0$ and consequently $\widetilde{\mathcal{M}}^h \subseteq \widetilde{\mathcal{M}}_{\pm\frac{4}{\pi}}$, which is absurd since the set on the right-hand side consists of a single point. However, α is not strictly convex, since there is a flat piece on which it is zero.

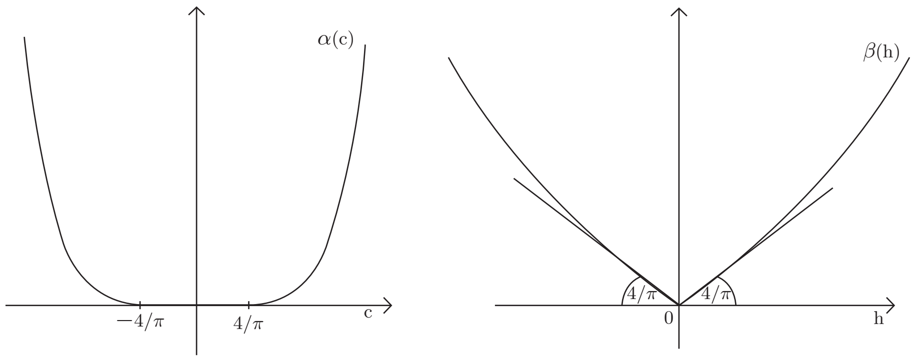

FIGURE 4. Sketch of the graphs of the α and β-functions of the simple pendulum.

As far as β is concerned, it is strictly convex (as a consequence of α being C^1), but it is differentiable everywhere except at the origin. At the origin, in fact, there is a corner and the set of subderivatives (*i.e.,* the slopes of tangent lines) is given by $\partial\beta(0) = [-\frac{4}{\pi}, \frac{4}{\pi}]$ (this is related to the fact that α has a flat on this interval).

REFERENCES

[1] Vladimir I. Arnol′d. Proof of a theorem of A. N. Kolmogorov on the preservation of conditionally periodic motions under a small perturbation of the Hamiltonian. *Uspehi Mat. Nauk*, 18 (5): 13–40, 1963.

[2] Vladimir I. Arnol′d. Mathematical methods of classical mechanics *Graduate Texts in Mathematics*, 60. Springer-Verlag, New York, 1989.

[3] Serge Aubry and P. Y. Le Daeron. The discrete Frenkel-Kontorova model and its extensions. I. Exact results for the ground-states. *Phys. D*, 8 (3): 381–422, 1983.

[4] Victor Bangert. Mather sets for twist maps and geodesics on tori. *Dyn. Rep.* 1, 1–56, 1988.

[5] Ana Cannas da Silva. Lecture on Symplectic Geometry. *Lecture Notes in Mathematics* 1764, Springer-Verlag, 2001.

[6] Constantin Caratheodory. Variationsrechnung und partielle Differentialgleichung erster Ordnung. Leipzig-Berlin: B.G. Teubner, 1935.

[7] Mario J. Dias Carneiro. On minimizing measures of the action of autonomous Lagrangians. *Nonlinearity*, 8 (6): 1077–1085, 1995.

[8] Gonzalo Contreras and Renato Iturriaga. Global minimizers of autonomous Lagrangians. Preprint, 1999.

[9] Albert Fathi. The Weak KAM theorem in Lagrangian dynamics. Cambridge University Press (to appear).

[10] Gustav A. Hedlund. Geodesics on a two-dimensional Riemannian manifold with periodic coefficients. *Ann. of Math. (2)*, 33 (4): 719–739, 1932.

[11] Andrey N. Kolmogorov. On conservation of conditionally periodic motions for a small change in Hamilton's function. *Dokl. Akad. Nauk SSSR (N.S.)*, 98: 572–530, 1954.

[12] Ricardo Mañé. On the minimizing measures of Lagrangian dynamical systems. *Nonlinearity*, 5(3): 623–638, 1992.

[13] Ricardo Mañé. Lagrangian flows: the dynamics of globally minimizing orbits. *Bol. Soc. Brasil. Mat. (N.S.)*, 28 (2): 141–153, 1997.

[14] John N. Mather. Existence of quasiperiodic orbits for twist homeomorphisms of the annulus. *Topology*, 21 (4): 457–467, 1982.

[15] John N. Mather. A criterion for the nonexistence of invariant circles. *Inst. Hautes Études Sci. Publ. Math.*, 63: 153–204, 1986.

[16] John N. Mather and Giovanni Forni. Action minimizing orbits in Hamiltonian systems. *Transition to chaos in classical and quantum mechanics.* Lectures given at the 3rd session of the Centro Internazionale Matematico Estivo (C.I.M.E.) held in Montecatini Terme, Italy, July 6–13, 1991. Berlin: Springer-Verlag. Lect. Notes Math. 1589, 92-186, 1994.

[17] John N. Mather. Variational construction of orbits of twist diffeomorphisms. *J. Amer. Math. Soc.*, 4 (2): 207–263, 1991.

[18] John N. Mather. Action minimizing invariant measures for positive definite Lagrangian systems. *Math. Z.*, 207 (2): 169–207, 1991.

[19] John N. Mather. Variational construction of connecting orbits. *Ann. Inst. Fourier (Grenoble)*, 43 (5): 1349–1386, 1993.

[20] John N. Mather. Order structure on action minimizing orbits. *Symplectic topology and measure preserving dynamical systems.* Papers of the AMS-IMS-SIAM joint summer research conference, Snowbird, UT, USA, July 1–5, 2007. Providence, RI: American Mathematical Society (AMS). Contemporary Mathematics 512, 41–125, 2010.

[21] Jürgen Moser. Convergent series expansions for quasi-periodic motions. *Math. Ann.*, 169: 136–176, 1967.

[22] R. Tyrell Rockafellar. Convex analysis. *Princeton Mathematical Series, No. 28*, Princeton University Press, xviii+451, 1970.

[23] Alfonso Sorrentino. Action-minimizing methods in Hamiltonian dynamics: an introduction to Aubry-Mather theory. *Mathematical Notes* Vol. 50, Princeton and Oxford University Press, pp. xi + 128, May 2015.

[24] Sol Schwartzman. Asymptotic cycles. *Ann. of Math. (2)*, 66: 270–284, 1957.

Dipartimento di Matematica, Università degli Studi di Roma "Tor Vergata", Viale della Ricerca Scientifica 1, 00133 Rome (Italy).

E-mail address: `sorrentino@mat.uniroma2.it`