METHODS AND APPLICATIONS OF ANALYSIS.
Vol. 26, No. 1, pp. 037–064, March 2019
003

ON JOHN MATHER'S SEMINAL CONTRIBUTIONS IN HAMILTONIAN DYNAMICS*

ALFONSO SORRENTINO†

To the memory of John N. Mather,
a remarkable man and mathematician. (1942 – 2017)

Key words. Action-minimizing methods, Aubry-Mather theory, Mather measures, Mather sets, Aubry sets, Mañé sets, Minimal average action.

Mathematics Subject Classification. 37J50, 37J45, 37J40, 70H08.

John N. Mather was undoubtedly one of the most influential mathematicians of the last decades, whose exceptional depth and originality had a profound impact on a vast area of mathematics, enhancing cross-field interactions and laying the foundations of novel promising lines of investigation.

I was one of the lucky ones that had the privilege of being John's Ph.D. student at Princeton University: besides learning and benefitting much from his well-rounded knowledge and unmatchable intuition, our weekly interactions profoundly shaped my way of doing and perceiving mathematical research itself. Extremely modest and reserved, he would rather let his striking mathematics speaks on his behalf. His works were a clear reflection of this distinct personality: not a single word was accidental or irrelevant, meant to bring into sharp focus the mathematical vision that he had vivid in his mind and was eager to share.

Trying to provide a comprehensive description of his mathematical legacy is an arduous task: from his earliest works on foliation theory, to his remarkable papers on the theory of singularities – providing the rigorous foundations of this theory –, up to the most recent revolutionary contributions in dynamical systems and Lagrangian/Hamiltonian dynamics, which led to the birth and outgrowth of novel areas of research that nowadays bear his name.

In this note[1] we would like to focus on these latter aspects of his work, trying to provide an overview of what is nowadays called *Mather's theory*: a set of ideas and results for Lagrangian systems, that John started to develop in the early 1990's, following the striking intuition that this was the right path to follow in order to tackle the problem of *Arnol'd diffusion* and to understand the onset of chaos in classical mechanics. Years passed, and we must acknowledge that he was – as usual – very right about that.

More specifically, starting from the observation that invariant Lagrangian graphs can be characterized in terms of their *action-minimizing* properties, John pointed out 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

*Received July 25, 2018; accepted for publication January 4, 2019.
†Dipartimento di Matematica, Università degli Studi di Roma "Tor Vergata", Viale della Ricerca Scientifica 1, 00133 Rome, Italy (sorrentino@mat.uniroma2.it).
[1]Portions of this material is 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 [51]). Since this is meant to be an overview of the theory, we will omit all proofs; we refer interested readers to [51] for a more systematic and comprehensive presentation of this and other topics.

as generalization of invariant Lagrangian graphs, despite not being in general either submanifolds or regular. Besides being very significant from a dynamical 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...

Dear John, your memory will live on through your beautiful mathematics. *Sit tibi terra levis...*

Acknowledgement. The author wishes to thank Rong Du, Yun Gao, Sen Hu, Stanislaw Janeczko, Mina Teicher, Stephen Yau, and Huaiqing Zuo for organizing the meeting "*International Conference on Singularity Theory and Dynamical Systems, in Memory of John Mather*" (11th–15th December 2017), and Tsinghua Sanya International Mathematics Forum (Hainan, China) for their kind hospitality. The author is grateful to Princeton University Press for agreeing on the use herein of some of the material from [51].

Finally, the author acknowledges the MIUR Excellence Department Project awarded to the Department of Mathematics, University of Rome Tor Vergata, CUP E83C18000100006, and the MIUR PRIN project "Regular and stochastic behaviour in dynamical systems" nr. 2017S35EHN.

1. In the beginning there were KAM tori.... 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 [3])[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 $\Phi_t^{\mathfrak{h}}(x_0, p_0) = (x_0 + t\rho(p_0) \bmod \mathbb{Z}^n,\, p_0)$. 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 1).

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 [26] – and later Arnol'd [1] and Moser [42] 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

[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.

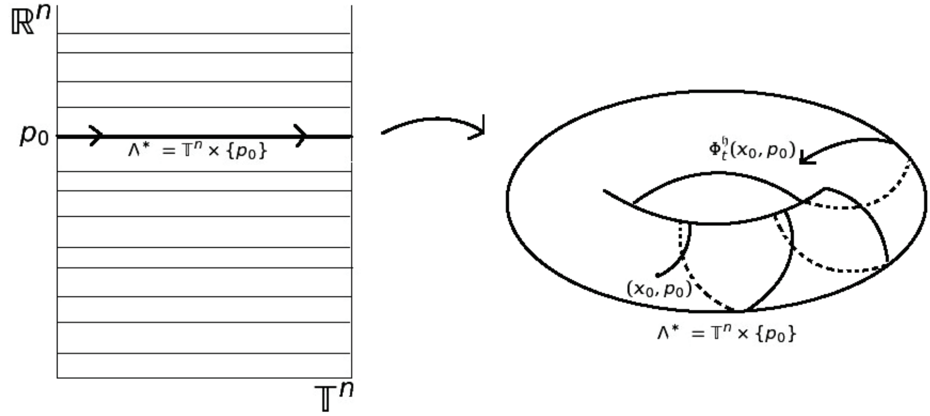

FIG. 1. *The phase space of an integrable system.*

(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 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?*

While all these questions are concerned with the investigation of stable motions of the perturbed system, another interesting issue soon took the stage: *does the breakdown of this stable picture open the way to orbits with unstable or chaotic behaviours?*

An answer to this latter question did not take long to arrive. In 1964 V. I. Arnol'd [2] constructed an example of a perturbed integrable system in which *unstable orbits* – resulting from the breaking of unperturbed KAM tori – coexist with the stable picture drawn by KAM theorem. This striking, and somehow unexpected, phenomenon, yet not completely understood, is nowadays called *Arnol'd diffusion*.

This new insight led to a change of perspective and in order to make sense of the complex balance between stable and unstable motions that was looming out, new approaches needed to be exploited. Amongst these, variational methods turned out to be particularly suitable and successful. Mostly inspired by the so-called *least action principle*,[3] a sort of widely accepted "thriftiness" of Nature in all its actions, they seemed to provide the natural setting to get over the local view given by the analytic methods and make towards a global understanding of the dynamics.

Aubry-Mather theory represented undoubtedly a great leap forward in this direction. Developed independently by Serge Aubry [4] and John Mather [33] 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

[3] "*Nature is thrifty in all its actions*", Pierre-Louis Moreau de Maupertuis (1744). A better-known special case of this principle is what is generally called *Maupertius' principle.* Actually, König published a note claiming priority for Leibniz in the Berlin Academy correspondences overseen by Maupertuis. Priority dispute brought in Euler, Voltaire and ultimately a committee convened by the King of Prussia. In 1913, the Berlin Academy reversed its previous decision and found Leibniz had priority.

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 (see also [5, 35] for nice expositions of these results). 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 [34, 36].

Motivated by these achievements, John Mather [37, 38] – and later Ricardo Mañé [28, 27] and Albert Fathi [20] 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 2.1), 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. These invariant sets are obtained as minimizing solutions to variational problems; as a result, these objects present a much richer structure and rigidity than one might generally expect and, quite surprisingly, play a leading role in determining the global dynamics of the system.

Let us remark that these tools – while dealing with stable motions – revealed also quite promising in the construction of *chaotic orbits*, such as for instance *connecting orbits* among the above-mentioned invariant sets [38, 6, 19]. Therefore they set high hopes on the possibility of proving the generic existence of Arnol'd diffusion in nearly integrable Hamiltonian systems [39, 9, 14, 15, 24, 25, 29]. However, differently from the case of twist diffeomorphisms, the situation turns out to be more complicated, due to a general lack of information on the topological structure of these action-minimizing sets. These sets, in fact, play a twofold role. Whereas on the one hand they may provide an obstruction to the existence of "diffusing orbits", on the other hand their topological structure plays a fundamental role in the variational methods that have been developed for the construction of orbits with "prescribed" behaviours. We will not enter further into the discussion of this problematic, but we refer the interested readers to [39, 40, 6, 8, 9, 14, 15, 19, 24, 25, 29].

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 describe how similar ideas can be extended to a more general setting.

We would like to conclude this section, by pointing out that in addition to its fundamental impact on the modern study of classical dynamics, Mather's theory has also contributed to point out interesting and sometimes unexpected links to other fields of research (both pure and applied), fostering a multidisciplinary interest in these ideas and in the techniques involved. The literature is vast, but interested readers could read, for example, [16, 20, 23, 30, 32, 43, 45, 47, 49, 50, 51, 52, 53, 54] and references therein, for more information on these recent developments.

2. Setting: Tonelli Lagrangians and Hamiltonians. 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 *La-*

grangians. 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[4] 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 2.3).

Definition 2.1 (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.,*

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

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.

[4]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 4.

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 [27]. 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
$$L_X(x,v) = \frac{1}{2}\left\|v - X(x)\right\|_x^2\,.$$
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 often turns out to be 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 [44]:
$$\begin{aligned} H :\ & T^*M \longrightarrow \mathbb{R} \\ & (x,p) \longmapsto \sup_{v \in T_xM} \{\langle p,\, v\rangle_x - L(x,v)\} \end{aligned}$$

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 [12, p. 207]). Such a Hamiltonian is called a *Tonelli* (or *optical*) *Hamiltonian*.

Definition 2.2 (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 [11]). 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 [44]) 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*:

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

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 2.3. 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.

3. Cartoon example: Action-minimizing properties of integrable systems. Before entering into the details of Mather's work, we would like to discuss a very easy case: *action-minimizing* properties of invariant measures and orbits of an integrable system (see section 1). This will provide us with a better understanding of the ideas behind Mather's theory and will describe clearer in which sense the *action-minimizing sets* that we are going to construct – namely, what we will call *Mather sets* (see section 4), *Aubry sets* and *Mañé sets* (see section 6) – represent a generalization of KAM tori.

As we have already discussed in section 1, 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 2 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* 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 $\tilde{\Lambda}_{p_0} := \mathcal{L}^{-1}(\Lambda^*_{p_0})$, where $\Lambda^*_{p_0} = \mathbb{T}^n \times \{p_0\}$, and compute its action. Observe that on the support of this measure $\ell(v) \equiv \ell(\rho(p_0))$. Then:

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

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:

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

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 (this will be generalized in section 4):

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

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 4.

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 (see section 4 for the general case):

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 $\rho(p_0)$:

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

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. Let (x_0, p_0) be a point on the KAM torus $\Lambda^*_{p_0}$ and we consider the projection on $\mathbb{T}^n$ of its orbit under the Hamiltonian flow, *i.e.*, $\gamma(t) = x_0 + \frac{\partial\mathfrak{h}}{\partial p_0}(p_0)t$. Let us fix any times $a < b$ and consider the corresponding *Lagrangian action* of this curve. Proceeding as we did before and using Fenchel-Legendre (in)equality we get:

$$\int_a^b \left(\ell(\dot\gamma(t)) - p_0\cdot\dot\gamma(t)\right)\, dt = -\mathfrak{h}(p_0)(b-a)\,.$$

Let us now take any other absolutely continuous curve $\xi : [a,b] \longrightarrow \mathbb{T}^n$ with the same endpoints as γ, *i.e.*, $\xi(a) = \gamma(a)$ and $\xi(b) = \gamma(b)$. Proceeding as before and using Fenchel-Legendre inequality, we obtain:

$$\int_a^b \Big(\ell(\dot\xi(t)) - p_0 \cdot \dot\xi(t)\Big)\, dt \geq -\mathfrak{h}(p_0)(b-a) \;=\; \int_a^b \left(\ell(\dot\gamma(t)) - p_0 \cdot \dot\gamma(t)\right)\, dt.$$

Therefore for any times $a < b$, γ is the curve that minimizes the action of $\ell(v) - p_0 \cdot v$ over all absolutely continuous curves $\xi : [a,b] \longrightarrow \mathbb{T}^n$ with $\xi(a) = \gamma(a)$ and $\xi(b) = \gamma(b)$.

Remark 3.1. Actually something more is true. Let us consider a curve with the same endpoints, but a different time-length, *i.e.*, $\xi : [a', b'] \longrightarrow \mathbb{T}^d$ with $a' < b'$ and such that $\xi(a') = \gamma(a)$ and $\xi(b') = \gamma(b)$. Proceeding as above, one obtains

$$\int_a^b \left(\ell(\dot\gamma(t)) - p_0 \cdot \dot\gamma(t) + \mathfrak{h}(p_0)\right) \leq \int_{a'}^{b'} \Big(\ell(\dot\xi(t)) - p_0 \cdot \dot\xi(t) + \mathfrak{h}(p_0)\Big)\,.$$

Hence, for any times $a < b$, γ minimizes the action of $\ell(v) - p_0 \cdot v + \mathfrak{h}(p_0)$ amongst all absolutely continuous curves ξ that connect $\gamma(a)$ to $\xi(b) = \gamma(b)$ in any given time (adding a constant does not change the Euler-Lagrange flow).

We have just proven the following fact:

For any given $a < b$, the projection γ of any orbit on $\Lambda^*_{p_0}$ minimizes the action of $\ell(v) - p_0 \cdot v$ amongst all absolutely continuous curves that connect $\gamma(a)$ to $\gamma(b)$ in time $b - a$.

Such a curve is called a p_0-*global minimizer* of ℓ.

In particular, we obtain that:

$$\mathcal{L}^{-1}(\Lambda^*_{p_0}) = \bigcup \{(\gamma(t), \dot\gamma(t)) :\ \gamma \text{ is a } p_0\text{-global minimizer of } \ell \text{ and } t \in \mathbb{R}\}.$$

The set on the right-hand side is often denoted by $\widetilde{\mathcal{N}}_{p_0}$ and called the *Mañé set of cohomology class* p_0 (it will be defined, in the general case, in section 6).

Actually these curves satisfy a stronger property that just being p_0-global minimizers, namely they are *regular minimizers*. This latter property is not very clear in this simple context: it happens to coincide with being a global minimizer; we will describe it more clearly in section 6. For the time being, let us just provide a sketchy idea of what we would like to look at.

Let $x_1, x_2 \in \mathbb{T}^n$ and let us denote by $h^T_{p_0}(x_1, x_2)$ the minimimal action of $\ell(v) - p_0 \cdot v$ along curves that connect x_1 to x_2 in time T. It follows from what we discussed above, that if there exist an orbit on $\Lambda^*_{p_0}$ connecting (x_1, p_0) to (x_2, p_0) in time T, then $h^T_{p_0}(x_1, x_2) = -\mathfrak{h}(p_0)T$, or equivalently $h^T_{p_0}(x_1, x_2) + \mathfrak{h}(p_0)T = 0$.

Inspired by this, we define the so-called *Pieierls' barrier*:

$$h_{p_0}(x_1, x_2) := \liminf_{T \to +\infty} \left(h^T_{p_0}(x_1, x_2) + \mathfrak{h}(p_0)T\right).$$

It is not difficult to show that in our case $h_{p_0}(x_1, x_2) \equiv 0$ for every $x_1, x_2 \in \mathbb{T}^n$.

In particular, $h_{p_0}(x, x) = 0$ for every $x \in \mathbb{T}^n$, namely we can find closed loops $\gamma_k : [0, T_k] \longrightarrow \mathbb{T}^n$, with $\gamma_k(0) = \gamma_k(T_k) = x$ and $T_k \to +\infty$, such that

$$\int_0^{T_k} \left(\ell(\dot\gamma_k(t)) - p_0 \cdot \dot\gamma_k(t) + \mathfrak{h}(p_0)\right)\, dt \longrightarrow 0 \qquad \text{as } k \to +\infty.$$

We say that a p_0-global minimizer γ is a *p_0-regular minimizer*, if $h_{p_0}(x_\alpha, x_\omega) + h_{p_0}(x_\omega, x_\alpha) = 0$ for each x_α in the α-limit set of γ and x_ω in the ω-limit set of γ.

It is trivial that in our case all minimizers are regular, since $h_{p_0}(x_1, x_2) \equiv 0$; however, this is not the case in general and justifies the attention to this special class of minimizers and the invariant set formed by their supports. In the special case we are considering,

$$\mathcal{L}^{-1}(\Lambda^*_{p_0}) = \bigcup \{(\gamma(t), \dot\gamma(t)) : \ \gamma \text{ is a } p_0\text{-regular minimizer of } \ell \text{ and } t \in \mathbb{R}\}.$$

The set on the right-hand side is often denoted by $\widetilde{\mathcal{A}}_{p_0}$ and called the *Aubry set of cohomology class c* (it will be defined, in the general case, in section 6).

4. Action-minimizing measures and Mather sets. 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 [51] 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 [18, 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 4.1. *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 3, 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)

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

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 [37].

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 4.2. *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 3.

Remark 4.3. 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 4.5 (*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):

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

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 3 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 [13].

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 4.4. *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 [37].
ii) It is contained in the energy level corresponding to $\alpha(c)$ [13].
iii) In [37] 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 3, 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:

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

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 [46] and [51] for more details).

Remark 4.5. (*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 [46]) 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 [37]).

Following Mather [37], 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

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

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 3 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 4.6. *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 3.

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 4.7. *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.

5. Mather's α and β-functions. The discussion in section 4 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 5.1. 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 5.2), 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 [44]). 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 5.2. *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 5.1 continues to hold in the general case

Proposition 5.3. *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 5.4. (*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 3, 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 $(x_1, x_2) \in \mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$ (see Figure 2).

As we have described in section 2, 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, $\mu_0 := \frac{1}{2}\mu_{\gamma_1} + \frac{1}{2}\mu_{\gamma_2}$

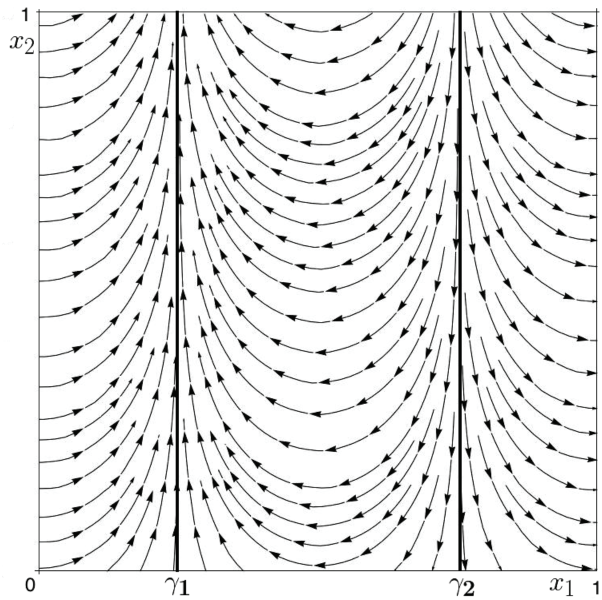

FIG. 2. *Plot of the vector field X.*

has zero homology and its support is contained in $\widetilde{\mathcal{M}}_0$. Therefore (see Proposition 5.3 (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 5.4, 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.
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 5.5. (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 5.6. The inclusion in Proposition 5.5 (i) may not be true at the end points of σ. For instance, Remark 5.4 (iv) provides an example in which the inclusion in Proposition 5.5 (i) is not true at the end-points of $\sigma(-h_0, h_0)$.

REMARK 5.7. It follows from the previous remarks and Proposition 5.5, 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 4.5, 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.

6. Action-minimizing curves and more invariant sets. In section 4 we have described the construction and the main properties of the Mather sets. One of the main limitations of these sets is that, being the support of invariant probability measures, they are *recurrent* under the flow (Poincaré recurrence theorem), *i.e.,* each orbit after a sufficiently long time (and therefore infinitely many often) will return arbitrarily close to its initial point. This property excludes many interesting invariant sets, which are somehow "invisible" to such a construction; for instance, think about the stable and unstable manifolds of some hyperbolic invariant set, or about heteroclinic and homoclinic orbits between invariant sets.

We will describe how to construct other (possibly) "larger" compact invariant sets and discuss their significance for the dynamics: the *Aubry sets* and the *Mañé sets*.

The key idea is the following: instead of considering action-minimizing invariant probability measures, one can look at *action-minimizing curves* for some modified Lagrangian. We showed in section 3 that orbits on KAM tori can be characterized in terms of this property. In this section we will imitate that construction in the general case of a Tonelli Lagrangian.

Similarly to what observed in section 4, let us fix a cohomology class $c \in H^1(M;\mathbb{R})$ and choose a smooth 1-form η on M that represents c. As we have already pointed out in section 2, there is a close relation between solutions of the Euler-Lagrange flow and extremals of the action functional A_{L_η} for the fixed end-point problem (which are the same as the extremals of A_L). In general, these extremals are not minima (they are local minima only if the time length is very short [20, Section 3.6]). One could wonder if such minima exist, namely if for any given end-points $x, y \in M$ and any given positive time T, there exists a minimizing curve connecting x to y in time T. From what already said, this curve will correspond to an orbit for the Euler-Lagrange flow. Under our hypothesis on the Lagrangian, the answer to this question turns out to be affirmative. This is a classical result in calculus of variations, known as *Tonelli Theorem*.

Theorem 6.1 (Tonelli Theorem, [37]). *Let M be a compact manifold and L a Tonelli Lagrangian on TM. For all $a < b \in \mathbb{R}$ and $x, y \in M$, there exists, in the set of absolutely continuous curves $\gamma : [a, b] \longrightarrow M$ such that $\gamma(a) = x$ and $\gamma(b) = y$, a curve that minimizes the action $A_{L_\eta}(\gamma) = \int_a^b L_\eta(\gamma(t), \dot{\gamma}(t))\,dt$.*

We refer the reader to [37, Appendix 1] for details on its proof.

Remark 6.2. (*i*) A curve minimizing $A_{L\eta}(\gamma) = \int_a^b L_\eta(\gamma(t), \dot{\gamma}(t))\,dt$ subject to the fixed end-point condition $\gamma(a) = x$ and $\gamma(b) = y$, is called a *c-Tonelli minimizer*. Recall that such minimizers do only depend on c and not on the chosen representative η. In fact, adding an exact 1-form df to L will contribute with a constant term $f(y) - f(x)$, that does not play any role in selecting the minimizers.

(*ii*) As Mañé pointed out in [28], in order for these minimizers to exist it is not necessary to assume the compactness of M: the superlinear growth condition with respect to some complete Riemannian metric on M is enough.

(*iii*) A Tonelli minimizer which is C^1 is in fact C^r (if the Lagrangian L is C^r) and satisfies the Euler-Lagrange equation; this follows from the usual elementary argument in calculus of variations, together with Caratheodory's remark on differentiability. In the autonomous case, Tonelli minimizers will be always C^1; in the non-autonomous time-periodic case (Tonelli Theorem holds also in this case [37]), one needs to require that the Euler-Lagrange flow is also complete.

In the following we will be interested in particular Tonelli minimizers that are defined for all times and whose action is minimal with respect to any given time length. We will see that these curves present a very rich structure.

Definition 6.3 (c-minimizers). *An absolutely continuous curve $\gamma : \mathbb{R} \longrightarrow M$ is a c-*(global) minimizer *for L, if for any given $a < b \in \mathbb{R}$*

$$A_{L_\eta}(\gamma\big|[a, b]) = \min A_{L_\eta}(\sigma)$$

where the minimum is taken over all $\sigma : [a, b] \to M$ such that $\sigma(a) = \gamma(a)$ and $\sigma(b) = \gamma(b)$.

Remark 6.4. Differently from what happens with invariant probability measures, it will not be always possible to find *action-minimizing orbits* for any given rotation vector (it is not even possible to define, in general, a rotation vector for every action minimizing orbit). For instance, an example due to Hedlund [22] provides the existence of a Riemannian metric on a three-dimensional torus, for which minimal geodesics exist only in three asymptotic directions; this example can be extended to any dimension larger than three.

One can prove that in the setting we are considering, c-minimizers always exist.

Proposition 6.5 (Mather, [37]). *For any $c \in H1(M;\mathbb{R})$, there exist c-global minimizers for L. Moreover, every trajectory of the Euler-Lagrange flow contained in $\widetilde{\mathcal{M}}_c$ is a c-global minimizer.*

See [37, Proposition 3] for a detailed proof.

Definition 6.6 (Mañé set). *The* Mañé set *(with cohomology class c) is defined as:*

$$\widetilde{\mathcal{N}}_c = \bigcup \{(\gamma(t), \dot{\gamma}(t)) : \ \gamma \text{ is a c-global minimizer and } t \in \mathbb{R}\}.$$

REMARK 6.7. (*i*) Clearly, this set is invariant. Moreover, it follows from Proposition 6.5 that $\widetilde{\mathcal{M}}_c \subseteq \widetilde{\mathcal{N}}_c$. This inclusion might be strict (in general this set is much larger).

(*ii*) It was proved by Carneiro [13] that there is a very interesting link between the energy of global minimizers and their average action. Namely, c-minimizers are contained in the energy level $\alpha(c)$ (see sections 4 and 5 for the definition and the main properties of Mather's α function). Hence,

$$\widetilde{\mathcal{N}}_c \subseteq \widetilde{\mathcal{E}}_c := \{(H \circ \mathcal{L})(x,v) = \alpha(c)\},$$

where $\mathcal{L}$ denotes the Legendre transform associated to L, as defined in (2).

(*iii*) It is possible to prove that if the Hamiltonian flow associated to H admits an invariant Lagrangian graphs Λ_c of cohomology class c, then each orbit on Λ_c is a c-minimizer and $\mathcal{L}^{-1}(\Lambda_c) \subseteq \widetilde{\mathcal{N}}_c$; see, for instance, [21].

We want now to introduce a special class of minimizers. The main ingredient is the notion of *Peierls barrier*, introduced in [38].[5]

For $t > 0$ and x, $y \in M$, let us consider :

$$h_{\eta,t}(x,y) = \min \int_0^t L_\eta(\gamma(s), \dot{\gamma}(s))\, ds\,, \tag{8}$$

where the minimum is taken over all piecewise C^1 paths $\gamma : [0,t] \longrightarrow M$, such that $\gamma(0) = x$ and $\gamma(t) = y$. This minimum is achieved because of Tonelli theorem (Theorem 6.1). We define the *Peierls barrier* as:

$$h_\eta(x,y) = \liminf_{t \to +\infty} (h_{\eta,t}(x,y) + \alpha(c)t)\,. \tag{9}$$

REMARK 6.8. (*i*) Observe that h_η does not depend only on the cohomology class c, but also on the choice of the representative η; namely, if $\eta' = \eta + df$, then $h_{\eta'}(x,y) = h_\eta(x,y) + f(y) - f(x)$. Anyhow, this dependence will not be harmful for what we are going to do in the following (it will not affect the set of action-minimizing curves).

(*ii*) This function h_η is a generalization of Peierls barrier introduced by Aubry [4] and Mather [33, 34, 36, 35] in their study of twist maps. In some sense we are comparing, in the limit, the action of Tonelli minimizers of time length T with the corresponding average c-minimimal action $-\alpha(c)T$. Remember, in fact, that $-\alpha(c)$ is the "average action" of a c-minimal measure.

(*iv*) Albert Fathi [20] showed that – in the autonomous case – this lim inf can be replaced by a limit. This is not generally true in the non-autonomous time-periodic case. Tonelli Lagrangians for which this convergence result holds are called *regular*; Patrick Bernard [6] showed that under suitable assumptions on the Mather set, it is possible to prove that the Lagrangian is regular. For instance, if the Mather set $\widetilde{\mathcal{M}}_c$ is union of 1-periodic orbits, then L_η is regular. This problem turned out to be strictly related to the convergence of the so-called *Lax-Oleinik semigroup* (see [20] for its definition).

[5]The function that we are defining here is actually a variant of h_c^∞ defined in [38]. Pay attention that throughout article [38], the sign of the α-function is wrong: wherever there is $\alpha(c)$, it should be substituted by $-\alpha(c)$.

It is interesting to consider the following symmetrization:

$$\begin{aligned} \delta_c : \; M \times M &\longrightarrow \mathbb{R} \\ (x,y) &\longmapsto h_\eta(x,y) + h_\eta(y,x). \end{aligned} \tag{10}$$

Observe that this function does now depend only on the cohomology class c and moreover it is non-negative, symmetric and satisfies the triangle inequality.
An interesting property of δ_c is the following (see [38, Section 8]). If d denotes the distance induced on M by the Riemannian metric g, then there exists $C > 0$ such that for each $x, y \in M$ we have

$$\delta_c(x,y) \leq C d(x,y)^2.$$

Let us see now some relation between this Peierls barrier (or equivalently δ_c) and c-action minimizing curves. Let $\gamma : \mathbb{R} \longrightarrow M$ be a c-minimizer and consider x_α, x'_α in the α-limit set[6] of γ and x_ω, x'_ω in the ω-limit set[7] of γ. Mather in [38, Section 6] proved that $\delta_c(x_\alpha, x'_\alpha) = \delta_c(x_\omega, x'_\omega) = 0$. In general, it is not true that $\delta_c(x_\alpha, x_\omega)=0$; what one can prove is that this value does not depend on the particular x_α and x_ω, *i.e.*, $\delta_c(x_\alpha, x_\omega) = \delta_c(x'_\alpha, x'_\omega)$: it is a property of the limit sets rather than of their elements. Nevertheless, there will exist particular c-minimizers for which this value is equal to 0 and these will be the c-minimizers that we want to single out.

Definition 6.9 (c-regular minimizers). *A c-minimizer $\gamma : \mathbb{R} \longrightarrow M$ is called a c-regular minimizer, if $\delta_c(x_\alpha, x_\omega) = 0$ for each x_α in the α-limit set of γ and x_ω in the ω-limit set of γ.*

It can be shown that orbits contained in $\widetilde{\mathcal{M}}_c$ are c-regular minimizers (see [38, 28, 51]), hence these special kind of curves do exist (see [38, 28, 20]).

Remark 6.10. (*i*) Observe that the adjective *regular* in the alternative appelation (coined by John Mather) has no relation to the smoothness of the curve, since, like all solutions of the Euler-Lagrange flow, this curve will be as smooth as the Lagrangian.

(*ii*) It follows from the fact that orbits in $\widetilde{\mathcal{M}}_c$ are c-regular minimizers, that all orbits on a KAM torus of cohomology class c, are c regular minimizers. Observe, however, that if the Hamiltonian flow associated to H admits an invariant Lagrangian graphs Λ_c of cohomology class c, although orbits on Λ_c are c- minimizer, yet it is not automatically true that they are c-regular minimizers (see [21, 51] for more details).

By means of this special kind of minimizers, Mather defined a new invariant set consisting of c-regular minimizers, namely what is nowadays called the *Aubry set*.

Definition 6.11 (Aubry set). *The* Aubry set *(with cohomology class c) is:*

$$\widetilde{\mathcal{A}}_c = \bigcup \left\{ (\gamma(t), \dot\gamma(t)) : \; \gamma \textit{ is a c-regular minimizer and } t \in \mathbb{R} \right\}.$$

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

Properties of this set:

[6] Recall that a point z is in the α-limit set of γ, if there exists a sequence $t_n \to -\infty$ such that $\gamma(t_n) \to z$.

[7] Recall that a point z is in the ω-limit set of γ, if there exists a sequence $t_n \to +\infty$ such that $\gamma(t_n) \to z$.

i) It is non-empty, compact and invariant [38].
ii) It is contained in the energy level corresponding to $\alpha(c)$ [13].
iii) In [37, 38] Mather proved the celebrated *graph theorem*:
Let $\pi : TM \longrightarrow M$ denote the canonical projection. Then, $\pi|\widetilde{\mathcal{A}}_c$ is an injective mapping of $\widetilde{\mathcal{A}}_c$ into M, and its inverse $\pi^{-1} : \mathcal{A}_c \longrightarrow \widetilde{\mathcal{A}}_c$ is Lipschitz.

We summarize in this diagram[8] the main properties of this set and its relations to the other invariant sets that we have defined so-far.

$$\begin{tikzcd}
\widetilde{\mathcal{M}}_c \arrow[d, shift right, "\pi"] & \subseteq & \widetilde{\mathcal{A}}_c \arrow[d, shift right, "\pi"'] & \subseteq & \widetilde{\mathcal{N}}_c & \subseteq & \widetilde{\mathcal{E}}_c := \{H \circ \mathcal{L}(x,v) = \alpha(c)\} & \subseteq & \mathrm{T}M \arrow[d, "\pi"] \\
\mathcal{M}_c \arrow[u, shift right, "(\pi|\widetilde{\mathcal{M}}_c)^{-1}"] & \subseteq & \mathcal{A}_c \arrow[u, shift right, "(\pi|\widetilde{\mathcal{A}}_c)^{-1}"'] & \subseteq & & & & \subseteq & M
\end{tikzcd}$$

Remark 6.12. (*i*) The above inclusions may not be strict (see [51] for some examples).

(*ii*) The Lipschitz graph property of $\widetilde{\mathcal{M}}_c$ and $\widetilde{\mathcal{A}}_c$ are generally called *Mather's graph theorem(s)*; namely, the Mather set and the Aubry set are contained in a Lipschitz graph over M. This is probably the most important property of these sets and it has many dynamical consequences. The original proof by Mather exploits the so-called "crossing" Lemma (see [37, Lemma p. 186]), inspired by similar properties of Riemannian geodesic flows. The graph property can be also proved in a different way by means of viscosity (sub)solutions of the associated Hamilton-Jacobi equation; this is the content of the so-called *weak KAM Theory* (see [20]).

(*iii*) The graph property does not hold in general for the Mañé set (see [51] for some counterexamples).

Observe that one can also provide an alternative definition of the (projected) Aubry set (compare with what discussed in section 3 in the case of KAM tori):

Proposition 6.13 (See [20, Proposition 5.3.8]). *The following properties are equivalent.*

i) $x \in \mathcal{A}_c$;
ii) $h_\eta(x,x) = 0$;
iii) *there exists a sequence of absolutely continuous curves $\gamma_n : [0, t_n] \to M$ such that:*
   - *for each n, we have $\gamma_n(0) = \gamma_n(t_n) = x$;*
   - *the sequence $t_n \to +\infty$, as $n \to +\infty$;*
   - *as $n \to +\infty$, $\int_0^{t_n} L_\eta(\gamma_n(s), \dot\gamma_n(s))\, ds + \alpha(c) t_n \to 0$.*

Remark 6.14. (*i*) Therefore, the Aubry set consists of points that are contained in loops with periods as long as we wish and actions as close as we want to the minimal average one.

(*ii*) Moreover, it follows from *ii*) in Proposition 6.13 that δ_c is a pseudometric on the projected Aubry set

$$\mathcal{A}_c = \{x \in M : \ \delta_c(x,x) = 0\}.$$

[8] This (unintentional?) typographical "coincidence" honoring Ricardo Mañé was first pointed out by Albert Fathi.

(*iii*) One can easily construct a metric space out of $(\mathcal{A}_c, \delta_c)$. We call *quotient Aubry set* (or *Mather quotient*) the metric space $(\bar{\mathcal{A}}_c, \bar{\delta}_c)$ obtained by identifying two points in $\mathcal{A}_c$, if their δ_c-pseudodistance is zero. This set plays quite an interesting role in the study of the dynamics; see for example [48] for more details.

7. Some topological and symplectic properties of the Aubry and Mañé sets. To conclude this survey, let us describe some other properties of these invariant sets.

We have remarked that the Mather sets, being the support of invariant probability measures, are recurrent under the flow. This is not true anymore for the Aubry and Mañé sets, but something can still be said. Let us first recall the definition of *ε-pseudo orbit*. Given a (compact) metric space X and a flow φ on it, we say that there exists an ε-pseudo orbit between two points $x, y \in X$, if we can find $\{x_n\}_{n=0}^{k_\varepsilon} \subset X$ and positive times $t_1, \ldots, t_{k_\varepsilon} > 0$ such that $x_0 = x$, $x_{k_\varepsilon} = y$ and $dist\big(\varphi_{t_{i+1}}(x_i), x_{i+1}\big) \leq \varepsilon$ for all $i = 0, \ldots, k_\varepsilon$.

PROPOSITION 7.1. (i) $\Phi^L\big|\widetilde{\mathcal{N}}_c$ *is chain transitive, i.e., for each $\varepsilon > 0$ and for all $(x,v),(y,w) \in \widetilde{\mathcal{N}}_c$, there exists an ε-pseudo-orbit for the flow Φ^L connecting them.*
(ii) $\Phi^L\big|\widetilde{\mathcal{A}}_c$ *is chain recurrent, i.e., for each $\varepsilon > 0$ and for all $(x,v) \in \widetilde{\mathcal{A}}_c$, there exists an ε-pseudo-orbit for the flow Φ^L connecting (x,y) to itself.*

The proof of this result can be found for instance in [17, Theorem V].

As a consequence of the chain-transitivity it follows that the Mañé set must be connected (the Aubry set in general not).

COROLLARY 7.2. *The Mañé set is connected.*

Moreover, one can also prove some topological and symplectic properties of the Aubry and Mañé sets, similar to what we have alread seen and proven for the Mather sets (see section 5).

In Proposition 5.5 we have related the intersection of Mather sets corresponding to different cohomology classes, to the "flatness" of the α-function. The same result holds for the Aubry set and has been proven by Daniel Massart in [31, Proposition 6]. However, the proof in this case is less straightforward and more involved.

Massart proved that it is possible to relate the dimension of a "face" of the epigraph of the α-function to the topological complexity of the Aubry sets corresponding to cohomologies in that face (see [31, Theorem 1]). More precisely, for any sufficiently small $\varepsilon > 0$, let us define $C_c(\varepsilon)$ be the set of integer homology classes which are represented by a piecewise C^1 closed curve made with arcs contained in $\mathcal{A}_c$ except for a remainder of total length less than ε. Let $C_c := \bigcap_{\varepsilon>0} C_c(\varepsilon)$. Let V_c be the space spanned in $H_1(M;\mathbb{R})$ by C_c. Note that V_c is an integer subspace of $H_1(M;\mathbb{R})$, that is it has a basis of integer elements (images in $H_1(M;\mathbb{R})$ of elements in $H_1(M;\mathbb{Z})$).

We denote by:

- F_c the maximal face (flat piece) of the epigraph of α, containing c in its interior;

- Vect F_c the underlying vector space of the affine subspace generated by F_c in $H^1(M;R)$;

- $V_c^\perp$ the vector space of cohomology classes of C^1 1- forms that vanish on V_c;

- G_c the vector space of cohomology classes of C^1 1-forms that vanish in T_xM for each $x \in \mathcal{A}_c$;

- E_c the space of cohomology classes of 1-forms of class C^1, the supports of which are disjoint from $\mathcal{A}_c$.

Massart in [31] proved that the following inclusions hold:

$$E_c \subseteq \mathrm{Vect}\, F_c \subseteq G_c \subseteq V_c^{\perp}.$$

Moreover:

THEOREM 7.3 (Massart, [31]). *Let $c \in H^1(M;\mathbb{R})$ and denote by F_c maximal face of the epigraph of α containing c in its interior.*
(i) *If a cohomology class c_1 belongs F_c, then $\mathcal{A}_c \subseteq \mathcal{A}_{c_1}$. In particular, if c_1 belongs to the interior of F_c, then they coincide, i.e., $\mathcal{A}_c = \mathcal{A}_{c_1}$.*
(ii) *Conversely, if two cohomology classes c and c_1 are such that $\widetilde{\mathcal{A}}_c \cap \widetilde{\mathcal{A}}_{c_1} \neq \emptyset$, then for each $\lambda \in [0,1]$ we have $\alpha(c) = \alpha(\lambda c + (1-\lambda)c_1)$, i.e., the epigraph of α has a face containing c and c_1.*

See also [32], where the relation between the differentiability of Mather's beta function and the integrability of the system has been thoroughly investigated.

To conclude this section, let us point out that these sets are symplectic invariant. Let us denote by $\mathcal{M}_c^*(H)$, $\mathcal{A}_c^*(H)$ and $\mathcal{N}_c^*(H)$ the Mather, Aubry and Mañé sets associated to a Tonelli Hamiltonian H (in the sense of the Legendre transform of the corresponding ones for the associated Lagrangian). Then:

THEOREM 7.4 (Bernard, [7]). *Let $L : TM \longrightarrow \mathbb{R}$ be a Tonelli Lagrangian and $H : T^*M \longrightarrow \mathbb{R}$ the associated Hamiltonian. If $\Phi : T^*M \longrightarrow T^*M$ is an exact symplectomorphism, then*

$$\begin{aligned}
\mathcal{M}_0^*(H \circ \Phi) &= \Phi^{-1}\left(\mathcal{M}_0^*(H)\right), \\
\mathcal{A}_0^*(H \circ \Phi) &= \Phi^{-1}\left(\mathcal{A}_0^*(H)\right) \\
\mathcal{N}_0^*(H \circ \Phi) &= \Phi^{-1}\left(\mathcal{N}_0^*(H)\right).
\end{aligned}$$

REMARK 7.5. The above result can be extended to non-exact symplectomorphisms and to other cohomology classes, under the assumption that Φ is isotopic to the identity. In general, there will be a linear reparametrization of the cohomology classes, depending on how the symplectomorphism acts on the first cohomology group; see [41, 51] for a more precise discussion.

For more symplectic geometric aspects of Mather theory, see for example [7, 10, 30, 41, 49, 50, 54].

REFERENCES

[1] V. 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 (1963), pp. 13–40.
[2] V. I. ARNOL'D, *Instability of dynamical systems with many degrees of freedom*, Dokl. Akad. Nauk SSSR, 156 (1964), pp. 9–12.
[3] V. I. ARNOL'D, *Mathematical methods of classical mechanics*, Graduate Texts in Mathematics, 60. Springer-Verlag, New York, 1989.
[4] S. 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 (1983), pp. 381–422.
[5] V. BANGERT, *Mather sets for twist maps and geodesics on tori*, Dyn. Rep. 1 (1988), pp. 1– 56.
[6] P. BERNARD, *Connecting orbits of time dependent Lagrangian systems*, Ann. Inst. Fourier (Grenoble), 52:5 (2002), pp. 1533–1568.
[7] P. BERNARD, *Symplectic aspects of Mather theory*, Duke Math. J., 136:3 (2007), pp. 401–420.

[8] P. Bernard, *The dynamics of pseudographs in convex Hamiltonian systems*, J. Amer. Math. Soc., 21:3 (2008), pp. 615–669.

[9] P. Bernard, V. Kaloshin, and K. Zhang, *Arnold diffusion in arbitrary degrees of freedom and 3-dimensional normally hyperbolic invariant cylinders*, Acta Math., 217:1 (2016), pp. 1–79.

[10] L. T. Butler and A. Sorrentino, *Weak Liouville–Arnol′d Theorems and their implications*, Comm. Math. Phys., 315:1 (2012), pp. 109–133.

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

[12] C. Caratheodory, *Variationsrechnung und partielle Differentialgleichung erster Ordnung*, Leipzig-Berlin: B.G. Teubner, 1935.

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

[14] C.-Q. Cheng and J. Y. Jun, *Existence of diffusion orbits in a priori unstable Hamiltonian systems*, J. Differential Geom., 67:3 (2004), pp. 457–517.

[15] C.-Q. Cheng, *Dynamics around the double resonance*, Camb. J. Math., 5:2 (2017), pp. 153–228. (See also Addendum to this article in Camb. J. Math., 5:4 (2017), pp. 571–571.

[16] K. Cieliebak, U. Frauenfelder, and G. P. Paternain, *Symplectic topology of Mañé's critical values*, Geom. Topol., 14:3 (2010), pp. 765–1870.

[17] G. Contreras, J. Delgado, and R. Iturriaga, *Lagrangian flows: the dynamics of globally minimizing orbits. II*, Bol. Soc. Brasil. Mat. (N.S.), 28:2 (1997), pp. 155–196.

[18] G. Contreras and R. Iturriaga, *Global minimizers of autonomous Lagrangians*, preprint, 1999.

[19] G. Contreras and G. P. Paternain, *Connecting orbits between static classes for generic Lagrangian systems*, Topology, 41:4 (2002), pp. 645–666.

[20] A. Fathi, *The Weak KAM theorem in Lagrangian dynamics*, Cambridge University Press (to appear).

[21] A. Fathi, A. Giuliani, and A. Sorrentino, *Uniqueness of invariant Lagrangian graphs in a homology or a cohomology class*, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5), 8:4 (2009), pp. 659–680.

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

[23] G. Huang, V. Kaloshin, and A. Sorrentino, *On the marked length spectrum of generic strictly convex billiard tables*, Duke Math. J., 167:1 (2018), pp. 175–209.

[24] V. Kaloshin and K. Zhang, *A strong form of Arnold diffusion for two and a half degrees of freedom*, preprint, 2012.

[25] V. Kaloshin and K. Zhang, *A strong form of Arnold diffusion for three and a half degrees of freedom*, preprint, 2014.

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

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

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

[29] J.-P. Marco, *Arnold diffusion for cusp-generic nearly integrable convex systems on $\mathbb{A}^3$*, preprint, 2016.

[30] S. Maro and A. Sorrentino, *Aubry-Mather theory for conformally symplectic systems*, Comm. Math. Phys., 354:2 (2017), pp. 775–808.

[31] D. Massart, *On Aubry sets and Mather's action functional*, Israel J. Math., 134 (2003), pp. 157–171.

[32] D. Massart and A. Sorrentino, *Differentiability of Mather's average action and integrability on closed surfaces*, Nonlinearity, 24 (2011), pp. 1777–1793.

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

[34] J. N. Mather, *A criterion for the nonexistence of invariant circles*, Inst. Hautes Études Sci. Publ. Math., 63 (1986), pp. 153–204.

[35] J. N. Mather and G. 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, pp. 92–186, 1994.

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

[37] J. N. Mather, *Action minimizing invariant measures for positive definite Lagrangian systems*, Math. Z., 207:2 (1991), pp. 169–207.
[38] J. N. Mather, *Variational construction of connecting orbits*, Ann. Inst. Fourier (Grenoble), 43:5 (1993), pp. 1349–1386.
[39] J. N. Mather, *Arnol′d diffusion. I. Announcement of results*, Sovrem. Mat. Fundam. Napravl., 2 (2003), pp. 116–130.
[40] J. N. Mather, *Arnol′d diffusion. II.*, Unpublished Manuscript, 2007.
[41] M. Mazzucchelli and A. Sorrentino, *Remarks on the symplectic invariance of Aubry-Mather sets*, C. R. Math. Acad. Sci. Paris, 354:4 (2016), pp. 419–423.
[42] J. Moser, *Convergent series expansions for quasi-periodic motions*, Math. Ann., 169 (1967), pp. 136–176.
[43] G. P. Paternain and A. Sorrentino, *Symplectic and contact properties of the Mañé critical value of the universal cover*, NoDEA Nonlinear Differential Equations Appl., 21:5 (2014), pp. 679–708.
[44] R. T. Rockafellar, *Convex analysis*, Princeton Mathematical Series, No. 28, Princeton University Press, xviii+451, 1970.
[45] K. F. Siburg, *The principle of least action in geometry and dynamics*, Lecture Notes in Mathematics, 1844. Springer-Verlag, Berlin, xii+128 pp., 2004.
[46] S. Schwartzman, *Asymptotic cycles*, Ann. of Math. (2), 66 (1957), pp. 270–284.
[47] A. Siconolfi and A. Sorrentino, *Global results for eikonal Hamilton-Jacobi equations on networks*, Anal. PDE, 11:1 (2018), pp. 171–211.
[48] A. Sorrentino, *On the total disconnectedness of the quotient Aubry set*, Ergodic Theory Dynam. Systems, 28:1 (2008), pp. 267–290.
[49] A. Sorrentino, *On the integrability of Tonelli Hamiltonians*, Trans. Amer. Math. Soc., 363:10 (2011), pp. 5071–5089.
[50] A. Sorrentino and C. Viterbo, *Action minimizing properties and distances on the group of Hamiltonian diffeomorphisms*, Geom. Topol., 14:4 (2010), pp. 2383–2403.
[51] A. 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.
[52] A. Sorrentino, *Lecture notes on Mather's theory for Lagrangian systems*, Publ. Mat. Urug., 16 (2016), pp. 169–192.
[53] A. Sorrentino and A. P. Veselov, *Markov numbers, Mather's beta function and stable norm*, Nonlinearity, 32:6 (2019), pp. 2147–2156.
[54] C. Viterbo, *Symplectic Homogenization*, Unpublished Manuscript, 2007.