Nonlinearity

PAPER • OPEN ACCESS

Aubry–Mather theory on graphs

To cite this article: Antonio Siconolfi and Alfonso Sorrentino 2023 *Nonlinearity* **36** 5819

OPEN ACCESS
IOP Publishing | London Mathematical Society Nonlinearity
Nonlinearity **36** (2023) 5819–5859 https://doi.org/10.1088/1361-6544/acf6ef

Aubry–Mather theory on graphs

Antonio Siconolfi[1] and **Alfonso Sorrentino**[2,*]

[1] Dipartimento di Matematica, Sapienza Università di Roma, Rome, Italy
[2] Dipartimento di Matematica, Università degli Studi di Roma 'Tor Vergata', Rome, Italy

E-mail: sorrentino@mat.uniroma2.it

Received 6 May 2022; revised 2 August 2023
Accepted for publication 5 September 2023
Published 26 September 2023

Recommended by Dr Tere M Seara

Abstract
We formulate Aubry–Mather theory for Hamiltonians/Lagrangians defined on graphs, study the structure of minimizing measures, and discuss the relationship with weak KAM theory developed in Siconolfi and Sorrentino (2018 *Anal. PDE* **1** 171–211). Moreover, we describe how to transport and interpret these results on networks.

Keywords: Aubry–Mather theory, Hamiltonians on networks, weak KAM
Mathematics Subject Classification numbers: 35F21, 35R02, 37J51

1. Introduction

Over the last years there has been an increasing interest in the study of the Hamilton–Jacobi equation on graphs and networks, as well as on related questions. These problems, in fact, besides having a great impact in the applications in various fields (for example to data transmission, traffic management problems, etc…), involve a number of subtle theoretical issues related to the intertwining between the local analysis of the problem and the global structure of the network/graph.

The paper presents the first, as far as we know, systematic detailed account of Aubry–Mather theory for Hamiltonians/Lagrangians defined on graphs recovering the whole theory in this new context and relating it to weak KAM analysis carried out in [29] (see also [25]).

We consider a connected graph $\Gamma = (\mathbf{V}, \mathbf{E})$ with a finite set of vertices $\mathbf{V}$ and a finite set of oriented edges $\mathbf{E}$ among vertices. A Hamiltonian on Γ is defined as a collection of Hamiltonians

* Author to whom any correspondence should be addressed.

$\mathcal{H}(e,\cdot):\mathbb{R}\to\mathbb{R}$, indexed by edges $e\in\mathbf{E}$, which are required, among other things, to be convex and superlinear at infinity, so that a Lagrangian $\mathcal{L}(e,\cdot)$ can be defined through Fenchel transform (see section 3 for more details). We stress that $\mathcal{H}(e,\cdot)$ and $\mathcal{H}(f,\cdot)$ are unrelated whenever $f\neq\pm e$, with $\pm e$ denoting an edge and its opposite (in other words, Hamiltonians and Lagrangians on geometrically distinct edges are independent one from the other).

As it is well-known, Aubry–Mather theory is a variational theory, inspired by the principle of least action [20, 22, 27, 30], whose aim is to find, in suitable spaces, minimizers of the Lagrangian action functional (possible with constraints); we refer the reader to [12,13] for more details on its connection to the study of the Hamilton-Jacobi equation. The passage from manifolds to graphs requires a specific adaptation of the main tools and techniques involved, which is by no means straightforward and, we believe, would be of potential interest for other problems and applications.

Understandably, the first step in our analysis is to define an action functional on each path of Γ, namely each finite sequence of concatenated edges. This requires the notion of a discrete Lagrangian (see (7)) and the one of parametrized paths, that is to say a velocity and a corresponding time to go through any edge of the path (see definition 4.1). The notion of admissible parametrization we introduce has a number of subtleties, especially in the case where the velocity vanishes on some edge. We remark that the notion of parametrization is somehow in duality with that of intrinsic length of a path, employed in the framework of weak KAM theory and related to a choice of a sub-level of the Hamiltonian.

The second issue to be settled is to suitably define the tangent cone of Γ and probability measures on it (see section 4.2).

Once this frame has been established, the main contributions in this article can be summarized as follows:

- We introduce the notion of occupation measures (see definition 4.11), which can be thought as measure representations of parametrized paths, and the one of closed probability measures and their rotation vectors, by adapting the corresponding definitions given on manifolds (see section 4.3). These are the measures involved in the minimization procedures in which we are interested.
  Significantly, we extend to our setting the density result, with respect to the first Wasserstein topology, of closed occupation measures in the set of closed measures (see theorem 4.15).
- We prove the existence of minimizing measures—named after Mather—obtained as minimizers of suitably constrained/modified variational problems (see theorem 5.3 and section 5.2). We define the corresponding Mather sets as the union of the supports of these minimizing measures (see (17) and (19)); these families of sets are parametrized, respectively, over the first homology and cohomology groups of the graph.
- We define the minimal average actions (the so-called Mather's α and β functions) (see (16) and (18)), prove that they are in convex duality and use them to relate Mather sets/measures corresponding to different homology and cohomology classes (proposition 5.8 and corollary 5.9).
- We thoroughly investigate structural properties of Mather measures; in particular, we prove that Mather measures are convex combinations of Dirac deltas (theorem 6.1) and deduce that they are convex combinations of occupation measures supported on parametrized circuits (theorem 6.6).
- We prove the analogue of Mather's graph theorem (proposition 6.2 and corollary 6.4).
- We introduce the notion of irreducible Mather's measures (section 6.2), prove that they correspond to occupation measures on parametrized circuits (proposition 6.8) and discuss for which homology classes they do exist (proposition 6.11). In theorem 6.10, we describe the set of Mather measures as a convex polytope generated by irreducible Mather measures.

- We relate Mather's α function to the critical value of certain Hamilton–Jacobi equation and show that (irreducible) Mather measures are supported on circuits of vanishing intrinsic length (theorem 8.1).
- We prove that the (projected) Mather sets are included in the corresponding Aubry sets (see (30) and corollary 8.3) and use viscosity solutions and subsolutions to provide a more explicit description of Mather's graph theorem (theorem 8.5).
- In section 8.2 we discuss some properties of Mather measures corresponding to the minimum of Mather's α-function.
- In appendix A we describe how to develop an Aubry–Mather theory on networks, and look from the point of view of networks to the notions that we have introduced on graphs and the corresponding results.

1.1. Final remarks and future directions

The investigation carried out in this paper and [29] is part of a more general project to prove a homogenization result for Hamilton–Jacobi equations on networks, following the homological approach introduced in [11]. The development of an Aubry–Mather theory on graphs allows determining the limit problem and proving the convergence result. Note that, even if the approximated equations in the homogenization problem are posed on a network, the natural setting where the approximation procedure should take place is the corresponding abstract graph. See [14, 15] for different interesting models of partial homogenization on junctures, mainly devoted to applications to traffic theory.

More generally, there is a broad interest in the recent literature on probability measures supported on graphs/networks, see for instance [8, 23]. One of the goal being, for instance, to extend mean field games models to graphs (see [1, 7, 16, 17]).

Passing to a related field, connections between Aubry–Mather theory and optimal transport have been pointed out in different contexts by various authors, see [3–6] (see also [9] for an application of Aubry-Mather theory to statistical mechanics). The outputs of the present paper can be seen as a further step to explore these directions of research in the graph/network setting.

1.2. Organization of the article

The article is organized as follows.

In section 2 we provide a brief introduction to graph theory, in order to set the terminology and introduce the main concepts that will be needed. In particular, we define the algebraic topological notions of chains, cochains, homology and cohomology of the graph, that are crucial importance for the full implementation of the variational analysis.

In section 3 we give the notion of Hamiltonian on a graph and introduce the associated Lagrangian which allows us to define the action functional to be minimized under appropriate constraints.

In section 4 we provide the relaxed setting on which the variational analysis will occur.

Sections 5 and 6 are the core of the development of Aubry–Mather theory in the context of graphs. We set, in analogy to the classical setting, a family of variational problems, show that they admit global minimizers and discuss their significance and their structural properties.

After having recalled in section 7 the basic results of weak KAM theory from [29], in section 8 we discuss the relation between Aubry–Mather theory and weak KAM theory on graphs. As in the classical case, these two approaches turn out to be tightly intertwined, each providing a different and interesting perspective on the other.

In appendix A we describe how to develop an Aubry–Mather theory on networks.

In appendix B we provide the proof of the density result of closed occupation measures.

2. Prerequisites on graphs

2.1. Definition and terminology

A graph $\Gamma = (\mathbf{V}, \mathbf{E})$ is an ordered pair of disjoint non-empty sets $\mathbf{V}$ and $\mathbf{E}$, which are called, respectively, *vertices* and (directed) *edges*, plus two functions:

$$\mathrm{o} : \mathbf{E} \longrightarrow \mathbf{V}$$

which associates to each edge its *origin* (initial vertex), and

$$\begin{aligned} - \quad : \mathbf{E} &\longrightarrow \mathbf{E} \\ e &\longmapsto -e, \end{aligned}$$

which changes direction and is a fixed point free involution, namely

$$-e \neq e \qquad \text{and} \qquad -(-e) = e \qquad \text{for any } e \in \mathbf{E}.$$

We define the *terminal vertex* of e as

$$\mathrm{t}(e) := \mathrm{o}(-e).$$

We further denote by $|\mathbf{V}|$, $|\mathbf{E}|$, the number of vertices and edges, respectively. For any vertex $x \in \mathbf{V}$, we denote by

$$\mathbf{E}_x := \{e \in \mathbf{E} : \ \mathrm{o}(e) = x\}$$

the set of edges originating from x; this is sometimes called the *star centred at* x.

An *orientation* of Γ is a subset $\mathbf{E}^+$ of the edges satisfying

$$-\mathbf{E}^+ \cap \mathbf{E}^+ = \emptyset \qquad \text{and} \qquad -\mathbf{E}^+ \cup \mathbf{E}^+ = \mathbf{E}.$$

In other words, an orientation of Γ consists of a choice of exactly one edge in each pair $\{e, -e\}$.

We define a *path* $\xi := (e_1, \ldots, e_M) = (e_i)_{i=1}^M$ as a finite sequence of concatenated edges in $\mathbf{E}$, namely $\mathrm{t}(e_j) = \mathrm{o}(e_{j+1})$ for any $j = 1, \ldots, M-1$.

We define the *length of a path* as the number of its edges. We set $\mathrm{o}(\xi) := \mathrm{o}(e_1)$, $\mathrm{t}(\xi) := \mathrm{t}(e_M)$. We call a path *closed*, or a *cycle*, if $\mathrm{o}(\xi) = \mathrm{t}(\xi)$.

Throughout the paper, we assume Γ to be

(G1) *finite*, namely with $|\mathbf{E}|$, $|\mathbf{V}|$ finite,
(G2) *connected*, in the sense that any two vertices are linked by some path.

It follows from the connectedness assumption, that the functions o and t are surjective.

In order to ease the presentation, in the following we also assume that

(G3) Γ *does not contain loops*, namely $\mathrm{o}(e) \neq \mathrm{t}(e)$ for any $e \in \mathbf{E}$.

This assumption is not essential, but it allows us to avoid some technical details. We will point out, throughout the article, the relevant parts that need to be modified in order include the presence of loops (see remarks 6.7, 7.1, and 8.6).

We call *simple* a path without repetition of vertices, except possibly the initial and terminal vertex, in other terms $\xi = (e_i)_{i=1}^M$ is simple if

$$\mathrm{t}(e_i) = \mathrm{t}(e_j) \Rightarrow i = j.$$

Clearly, there are finitely many simple paths in a finite graph. We call *circuit* a simple closed path. Given any edge e, we call *equilibrium circuit* (based on e) the path $(e,-e)$.

2.2. Homology of a graph

Throughout the paper we will take homology and cohomology with coefficients in $\mathbb{R}$. We refer to [31, chapter 4] for a more detailed and general presentation.

We define the 0–*chain* group as the free Abelian group on the vertices with coefficients in $\mathbb{R}$. We denote it by $\mathfrak{C}_0(\Gamma,\mathbb{R})$. We have

$$\mathfrak{C}_0(\Gamma,\mathbb{R}) \sim \mathbb{R}^{|\mathbf{V}|}.$$

We do the same operation with edges, making the reversed edge $-e$ coincide with the opposite of e with respect to the group operation, and we obtain the 1–*chain group*, denoted by $\mathfrak{C}_1(\Gamma,\mathbb{R})$. A basis is given by any orientation $\mathbf{E}^+$, in the sense that any element of the 1–chain group can be uniquely expressed as a linear combination of elements in $\mathbf{E}^+$ with real coefficients. We consequently have

$$\mathfrak{C}_1(\Gamma,\mathbb{R}) \sim \mathbb{R}^{|\mathbf{E}|/2}.$$

We define the *boundary operator* $\partial : \mathfrak{C}_1(\Gamma,\mathbb{R}) \to \mathfrak{C}_0(\Gamma,\mathbb{R})$ by setting for any edge

$$\partial e := \mathrm{t}(e) - \mathrm{o}(e)$$

and then extending it linearly; clearly, $\partial(-e) = -\partial e$.

The (*first*) *Homology group* of Γ with coefficients in $\mathbb{R}$ is defined by

$$H_1(\Gamma,\mathbb{R}) := \mathrm{Ker}\,\partial.$$

Some remarks:

- $H_1(\Gamma,\mathbb{R})$ is a subgroup of $\mathfrak{C}_1(\Gamma,\mathbb{R})$.
- $H_1(\Gamma,\mathbb{R})$ is a free Abelian group of finite rank. The (*first*) *Betti number* is defined to be the rank of $H_1(\Gamma,\mathbb{R})$, it is an indicator of the topological complexity of the network.
- An element of $H_1(\Gamma,\mathbb{R})$ is called a 1–*cycle*. In particular a 1–chain $\sum_{e\in\mathbf{E}^+} a_e e$ is a 1–cycle if and only if

$$\sum_{e\in\mathbf{E}^+,\,\mathrm{t}(e)=x} a_e = \sum_{e\in\mathbf{E}^+,\,\mathrm{o}(e)=x} a_e \qquad \text{for any } x\in\mathbf{V}; \tag{1}$$

This can be considered as an analogue of *Kirchhoff law* for electric circuits.

Due to (1), we can associate to any closed path $\xi = (e_i)_{i=1}^M$ in Γ an element of $H_1(\Gamma,\mathbb{R})$ via

$$[\xi] := \sum_{i=1}^{M} e_i. \tag{2}$$

We call $[\xi]$ the *homology class* of ξ. The converse is also true: every element of $H_1(\Gamma,\mathbb{Z})$ can be represented by a closed path (see [31, pp 40–41]).

Remark 2.1. The first Betti number of the graph Γ, i.e. rank of $H_1(\Gamma,\mathbb{R})$, can be also characterized in a combinatorial way via the so-called *cyclomatic number* introduced by Kirchhoff. In fact (see [31, formula (4.3)]), one can prove that it equals

$$\frac{1}{2}|\mathbf{E}| - |\mathbf{V}| + 1$$

(or, more generally, it is equal to $\frac{1}{2}|\mathbf{E}|-|\mathbf{V}|+|\mathbf{C}|$, where $\mathbf{C}$ denotes the set of connected components of the graph).

2.3. Cohomology of a graph

Let us introduce the dual entities of chains. The *0–cochain group*, denoted by $\mathfrak{C}^0(\Gamma,\mathbb{R})$, is the space of functions from $\mathbf{V}$ to $\mathbb{R}$, and the *1–cochain group*, denoted by $\mathfrak{C}^1(\Gamma,\mathbb{R})$, is the space of functions $\eta:\mathbf{E}\longrightarrow\mathbb{R}$, satisfying the compatibility condition

$$\eta(-e)=-\eta(e) \qquad \text{for any } e\in\mathbf{E}.$$

The algebraic structure of additive Abelian group is induced by the one in $(\mathbb{R},+)$.

We introduce the *differential* or *coboundary operator*

$$d:\mathfrak{C}^0(\Gamma,\mathbb{R})\longrightarrow\mathfrak{C}^1(\Gamma,\mathbb{R})$$

which is defined in the following way: for every $g\in\mathfrak{C}^0(\Gamma,\mathbb{R})$, the 1–cochain $\mathrm{d}g$ is given via

$$\mathrm{d}g(e):=g(\mathrm{t}(e))-g(\mathrm{o}(e)) \qquad \text{for all } e\in\mathbf{E};$$

it clearly satisfies the compatibility condition $\mathrm{d}g(-e)=-\mathrm{d}g(e)$.

It is easy to check that d is a group homomorphism. Hence, the (*first*) *Cohomology group* of Γ with coefficients in $\mathbb{R}$ can be defined as the quotient group

$$H^1(\Gamma,\mathbb{R}):=\mathfrak{C}^1(\Gamma,\mathbb{R})/\mathrm{Im}\,d.$$

One can show that there exists a canonical isomorphism

$$H^1(\Gamma,\mathbb{R})\simeq\mathrm{Hom}\left(H_1(\Gamma,\mathbb{R}),\mathbb{R}\right).$$

2.4. Pairings between chains and cochains, homology and cohomology

Let us introduce a *pairing* between 0–chains and 0–cochains:

$$\begin{aligned}\langle\cdot,\cdot\rangle:\mathfrak{C}^0(\Gamma,\mathbb{R})\times\mathfrak{C}_0(\Gamma,\mathbb{R})&\longrightarrow\mathbb{R}\\ \left(g,\sum_{x\in\mathbf{V}}\alpha_x x\right)&\longmapsto\sum_{x\in\mathbf{V}}\alpha_x g(x).\end{aligned}$$

Similarly, we can define the pairing between 1–chains and 1–cochains (we adopt the same notation):

$$\begin{aligned}\langle\cdot,\cdot\rangle:\mathfrak{C}^1(\Gamma,\mathbb{R})\times\mathfrak{C}_1(\Gamma,\mathbb{R})&\longrightarrow\mathbb{R}\\ \left(\eta,\sum_{e\in\mathbf{E}}\alpha_e e\right)&\longmapsto\sum_{e\in\mathbf{E}}\alpha_e\eta(e).\end{aligned}$$

The above pairings allow us to relate differential and boundary operators. Let $g\in\mathfrak{C}^0(\Gamma,\mathbb{R})$ and $\zeta=\sum_{e\in\mathbf{E}}\alpha_e e\in\mathfrak{C}_1(\Gamma,\mathbb{R})$; then we have:

$$\begin{aligned}\langle\mathrm{d}g,\zeta\rangle&=\sum_{e\in\mathbf{E}}\alpha_e\mathrm{d}g(e)=\sum_{e\in\mathbf{E}}\alpha_e\left(g(\mathrm{t}(e))-g(\mathrm{o}(e))\right)\\ &=\sum_{e\in\mathbf{E}}\alpha_e\langle g,\partial e\rangle=\left\langle g,\sum_{e\in\mathbf{E}}\alpha_e e\right\rangle=\langle g,\partial\zeta\rangle.\end{aligned} \tag{3}$$

In particular, this means that whenever $\zeta \in \mathfrak{C}_1(\Gamma,\mathbb{R})$ is such that $\partial\zeta = 0$, then $\langle \mathrm{d}g, \zeta\rangle = 0$ for all $g \in \mathfrak{C}^0(\Gamma,\mathbb{R})$. Hence, the above pairing descends to a well-defined pairing between first homology and first cohomology groups, that we continue to denote $\langle\cdot,\cdot\rangle : H^1(\Gamma,\mathbb{R}) \times H_1(\Gamma,\mathbb{R}) \longrightarrow \mathbb{R}$.

3. Hamiltonians and Lagrangians on graphs

3.1. Definitions and assumptions

We call a *Hamiltonian* on the graph $\Gamma = (\mathbf{V},\mathbf{E})$ a family of functions

$$\mathcal{H}(e,\cdot) : \mathbb{R} \to \mathbb{R}$$

labeled by the edges, such that

$$\mathcal{H}(e,p) = \mathcal{H}(-e,-p) \qquad \text{for any } e \in \mathbf{E}, p \in \mathbb{R}. \tag{4}$$

We further require that, for any $e \in \mathbf{E}$, $\mathcal{H}(e,\cdot)$ is

(H1) *strictly convex* and *differentiable*;
(H2) *superlinear* at $\pm\infty$, namely

$$\lim_{p\to\pm\infty} \frac{\mathcal{H}(e,p)}{|p|} = +\infty.$$

This implies that there exists, for any e, a unique $p_e = -p_{-e}$ global minimizer of both $\mathcal{H}(e,\cdot)$ in $\mathbb{R}$. We consider in what follows $\mathcal{H}(e,\cdot)$ mostly restricted to $[p_e,+\infty)$, (resp. $\mathcal{H}(-e,\cdot)$ restricted to $[p_{-e},+\infty)$), which is strictly increasing in this domain of definition. We set

$$a_e = \mathcal{H}(e,p_e) = \mathcal{H}(-e,p_{-e}) = a_{-e}. \tag{5}$$

We define $\sigma(e,\cdot)$ as the inverse function of $\mathcal{H}(e,\cdot)$ in $[p_e,+\infty)$. We have

$$\sigma(e,\cdot) : [a_e,+\infty) \to [p_e,+\infty) \qquad \text{for any } e \in \mathbf{E}$$

and

$$\sigma(e,a_e) = -\sigma(-e,a_e) = p_e = -p_{-e} \qquad \text{for any } e. \tag{6}$$

The properties summarized in the next statement are immediate.

Lemma 3.1. *Let $e \in \mathbf{E}$. The function $a \mapsto \sigma(e,a)$ from $[a_e,+\infty)$ to $\mathbb{R}$ is continuous, differentiable in $(a_e,+\infty)$, and strictly increasing for any e. In addition, it is strictly concave and satisfies*

$$\lim_{a\to+\infty} \frac{\sigma(e,a)}{a} = 0.$$

We define the *Lagrangian* $\mathcal{L}(e,\cdot) : \mathbb{R} \to \mathbb{R}$ as the convex conjugate of $\mathcal{H}(e,\cdot)$, namely

$$\mathcal{L}(e,q) := \max_{p\in\mathbb{R}} \left(p\,q - \mathcal{H}(e,p)\right). \tag{7}$$

Proposition 3.2. *Let $e \in \mathbf{E}$. The function $q \mapsto \mathcal{L}(e,q)$ is strictly convex and superlinear as q goes to $\pm\infty$. In addition*

$$\mathcal{L}(e,q) = \mathcal{L}(-e,-q) \qquad \textit{for any } q \in \mathbb{R}. \tag{8}$$

This is a consequence of **(H1)–(H2)** and (4) (see, for instance, [26, theorem 26.6]).

In what follows, we mostly consider $\mathcal{L}(e,\cdot)$ restricted to $[0,+\infty)$. We have

$$\mathcal{L}(e,q) = \max_{p \geqslant \sigma(e,a_e)} (pq - \mathcal{H}(e,p)) \qquad \text{for } q \geqslant 0,$$

an equivalent formula is

$$\mathcal{L}(e,q) = \max_{a \geqslant a_e} (q\sigma(e,a) - a) \qquad \text{for } q \geqslant 0, \tag{9}$$

from which it follows that $\mathcal{L}(e,0) = -a_e$.

Given $\omega \in \mathfrak{C}^1(\Gamma,\mathbb{R})$, we further consider the ω– *modified Hamiltonian*

$$\mathcal{H}^\omega(e,p) := \mathcal{H}(e,p + \langle \omega, e\rangle),$$

which clearly still satisfies assumptions **(H1)**, **(H2)**. It is therefore invertible on the right of its minimizer and the inverse is

$$\sigma^\omega(e,a) := \sigma(e,a) - \langle \omega, e\rangle. \tag{10}$$

The corresponding ω–*modified Lagrangian* is given by

$$\mathcal{L}^\omega(e,q) := \mathcal{L}(e,q) - \langle \omega, qe\rangle.$$

Remark 3.3. Note that a_e does not depend on ω, i.e. it is the same for $\mathcal{H}^\omega(e,\cdot)$. In fact by (6) a_e is characterized by the relation

$$\sigma(e,a_e) + \sigma(-e,a_e) = 0$$

and by (10)

$$\sigma(e,a_e) + \sigma(-e,a_e) = \sigma^\omega(e,a_e) + \sigma^\omega(-e,a_e) \qquad \text{for any 1–cochain } \omega.$$

4. Probability measures on edges

4.1. Preamble: parametrized paths

The notion of parametrized path is central in the paper and it will be essential to define occupation measures.

Intuitively speaking, a parametrized path is a path where it is assigned to any edge a non-negative *average speed* and a *time* needed to go through it. The time is the inverse of the speed, if the latter is positive, while it can be any possible positive number if the speed is zero. We motivate this choice in section A.2 in the case where Γ is the abstract graph associated to a network.

Definition 4.1. We say that $\xi = (e_i, q_i, T_i)_{i=1}^M$ is a *parametrized path* if

(i) $(e_i)_{i=1}^M$ is a family of concatenated edges which is called the *support* of ξ;

(ii) the q_i are non-negative numbers and

$$T_i = \begin{cases} \frac{1}{q_i} & \text{if } q_i > 0 \\ \text{a positive constant} & \text{if } q_i = 0; \end{cases}$$

we denote by $T_\xi := \sum_i T_i$ the *total time* of the parametrization of ξ;

(iii) if all the $q_i's$ vanish then $\mathrm{o}(\xi) = \mathrm{t}(\xi)$;
(iv) if $q_i = 0$ and $e_{i+1} \neq -e_i$ then $q_{i+1} \neq 0$;
(v) if $q_i \neq 0$, $i > 1$, then

$$\mathrm{o}(e_i) = \mathrm{t}(e_j) \qquad \text{with } j = \max\{k < i,\, q_k \neq 0\}.$$

We call a *parametrized cycle*, a parametrized path supported on a closed path (or cycle). We call a *parametrized circuit*, a parametrized path supported on a circuit.

Remark 4.2. Intuitively, a parametrized path can be thought as a concatenation of triples with non-zero average velocity, and pairs of triples (i.e. *equilibrium circuits*) of the form $\{(e,0,T),(-e,0,S)\}$ for some $e \in \mathbf{E}$ and $T,S > 0$. In particular, condition **(iv)** reads that there cannot be consecutive equilibrium circuits corresponding to different edges.

Equilibrium circuits represent steady states, interpreted as floating with zero average speed along an edge and its opposite. Therefore, if all speeds vanish (item **(iii)**) then initial and final position must coincide. Items **(iv)**, **(v)** further prescribe that an object possessing vanishing speed on an edge e starts floating back and forth along e and $-e$, and exits the swinging state from the same vertex it entered, only when the speed becomes positive.

Alternatively to the introduction of equilibrium circuits, one could consider the possibility that a path stops at a vertex x for some time. If this is the case, one needs to specity a 'cost per time' (or action) c_x to pay for remaining still at a vertex x; the most natural candidate for this values is $c_x := \min_{\mathrm{o}(e)=x} \mathcal{L}(e,0)$, which makes—at least for our purposes—this point of view equivalent to the one that we have adopted. Observe that in principle one could choose a different value for c_x, with the only requirement that $c_x \leqslant \min_{\mathrm{o}(e)=x} \mathcal{L}(e,0)$; this object appears in the literature with the name of *flux limiter* and becomes particularly relevant in the study of the time-dependent Hamilton–Jacobi equation in order to deal with discontinuity interfaces (which are 1-dimensional subspaces); see for instance [18, 28].

We deduce from the definition the following properties.

Proposition 4.3. *Let $\xi = (e_i,q_i,T_i)_{i=1}^M$ be a parametrized path.*

(i) *If some speed q_i is non-vanishing, and $i_1,\dots,i_K$ is the increasing sequence of indices corresponding to edges with positive speed, then*

$$\bar{\xi} := (e_{i_j},q_{i_j},T_{i_j})_{j=1}^K$$

is still a parametrized path with all average velocities different from 0 and such that $\mathrm{o}(\bar{\xi}) = \mathrm{o}(\xi)$, $\mathrm{t}(\bar{\xi}) = \mathrm{t}(\xi)$.
(ii) *If a parametrized path has all average speeds equal to zero, then it is supported on an edge and its opposite.*
(iii) *A parametrized circuit with some vanishing speed consists of an equilibrium circuit $\{(e,0,T),(-e,0,S)\}$ for some $e \in \mathbf{E}$ and $T,S > 0$.*

4.2. Basic definitions

In this section we introduce a notion of tangent cone $T\Gamma$ of Γ and define suitable sets of probability measures that we will use to build a version of Mather theory on graphs.

Definition 4.4. The *tangent cone* of Γ is defined as

$$T\Gamma := \mathbf{E} \times \mathbb{R}^+/\sim,$$

where $\mathbb{R}^+ := [0,+\infty)$ and $\sim$ is the identification $(e,0) \sim (-e,0)$.

We denote each fibre by $\mathbb{R}^+_e := \{e\} \times \mathbb{R}^+$.

We endow $T\Gamma := \mathbf{E} \times \mathbb{R}^+$ with a distance defined as:

$$d((e_1,q_1),(e_2,q_2)) := \begin{cases} q_1+q_2+1 & \text{if } e_1 \neq \pm e_2 \\ q_1+q_2 & \text{if } e_1 = -e_2 \\ |q_1-q_2| & \text{if } e_1 = e_2. \end{cases}$$

This makes $T\Gamma$ a Polish space. A set A is open in $T\Gamma$ in the induced topology if and only if $A \cap \mathbb{R}^+_e$ is open in the natural topology of $\mathbb{R}^+$ for any e. Accordingly, F is a Borelian set on $T\Gamma$ if and only if $F \cap \mathbb{R}^+_e$ is Borelian in $\mathbb{R}^+_e$ for any edge e.

Definition 4.5. Given μ a Borel probability measure on $T\Gamma$, we define the *support of* μ as the set

$$\mathrm{supp}_{\mathbf{E}}\mu = \{e \in \mathbf{E} \mid \mu(\mathbb{R}^+_e) > 0\}.$$

Proposition 4.6. *Any Borel probability measure in* $T\Gamma$ *can be decomposed as the convex combination of Borel probability measures in each fibre, namely*

$$\mu(F) = \sum_{e\in\mathbf{E}} \lambda_e \mu_e(F \cap \mathbb{R}^+_e) \qquad \text{for any Borelian set } F \subseteq T\Gamma, \tag{11}$$

where μ_e *are Borel probability measures on* $\mathbb{R}^+_e$ *and* $\lambda_e \geqslant 0$ *such that* $\sum_{e\in\mathbf{E}} \lambda_e = 1$. *In particular,* $\mathrm{supp}_{\mathbf{E}}\mu = \{e \in \mathbf{E} \mid \lambda_e \neq 0\}$.

Proof. We distinguish two cases, according to whether $\mu(e,0) = 0$ or $\mu(e,0) > 0$. In the first case, we set $\lambda_e := (\mu(\mathbb{R}^+_e))$: if $\lambda_e = 0$ (i.e. $e \notin \mathrm{supp}_{\mathbf{E}}\mu$), then the choice of μ_e is irrelevant; otherwise we define μ_e as the restriction of μ on $\mathbb{R}^+_e$, normalized in order to be a probability measure.

If $\mu(e,0) > 0$, then μ_e is not uniquely determined since we have a degree of freedom in sharing the contribute of $\mu(e,0) = \mu(-e,0)$ between e and $-e$. For, we introduce two positive constants m_e and m_{-e}, such that $m_e + m_{-e} = 1$, and denote by $\hat{\mu}_e$ the restriction of μ to $\mathbb{R}^+_e \setminus \{0\}$. Then, we define

$$\mu_e := \frac{1}{\hat{\mu}_e(\mathbb{R}^+_e) + m_e\mu(e,0)}\, \hat{\mu}_e + m_e\,\delta(e,0)$$
$$\lambda_e := \hat{\mu}_e(\mathbb{R}^+_e) + m_e\mu(e,0),$$

where $\delta(e,0)$ denotes Dirac delta at $(e,0)$; analogously for $-e$. □

Note that a Borel probability measure $\mu = \sum_{e\in\mathbf{E}} \lambda_e \mu_e$ has finite first momentum if and only such property holds for any μ_e, namely

$$\int_0^{+\infty} q\, \mathrm{d}\mu_e < +\infty \qquad \text{for any } e \in \mathbf{E}.$$

We denote by $\mathbb{P}$ the family of Borel probability measures on $T\Gamma$ with finite first momentum and we endow it with the (first) Wasserstein distance (see, for example, [32]). The corresponding convergence of measures can be expressed in duality with continuous functions $F(e,q)$ on $T\Gamma$ possessing linear growth at infinity; namely, given a sequence $\{\mu_n\}_n$ and μ in $\mathbb{M}$

$$\mu_n \to \mu \iff \int F(e,q)\, \mathrm{d}\mu_n \to \int F(e,q)\, \mathrm{d}\mu_n$$

for any function F continuous in $T\Gamma$ such that for any $e \in \mathbf{E}$ there exist $a_e, b_e \in \mathbb{R}$ such that

$$\forall e \in \mathbf{E} \quad \exists a_e, b_e \in \mathbb{R}: \qquad |F(e,q)| \leqslant a_e\, q + b_e \qquad \forall q \geqslant 0. \tag{12}$$

4.3. Closed probability measures on $T\Gamma$

Let us observe that for any $\omega \in \mathfrak{C}^1(\Gamma,\mathbb{R})$, the function

$$(e,q) \longmapsto \langle \omega, q\,e\rangle$$

is continuous with linear growth on $T\Gamma$ (see (12)). Given $\mu = \sum_e \lambda_e \mu_e \in \mathbb{P}$, we consequently define

$$\begin{aligned}\int \omega \,\mathrm{d}\mu &:= \sum_{e\in\mathbf{E}} \lambda_e \int_0^{+\infty} \langle \omega, q\,e\rangle \,\mathrm{d}\mu_e \\ &= \left\langle \omega, \sum_{e\in\mathbf{E}} \left[\lambda_e \int_0^{+\infty} q\,\mathrm{d}\mu_e\right] e \right\rangle. \end{aligned} \tag{13}$$

This associates to μ a 1–chain

$$\rho(\mu) := \sum_{e\in\mathbf{E}} \left[\lambda_e \int_0^{+\infty} q\,\mathrm{d}\mu_e\right] e \in \mathfrak{C}_1(\Gamma,\mathbb{R}). \tag{14}$$

Definition 4.7. We say that μ is a *closed measure* if

$$\int \mathrm{d}f \mathrm{d}\mu = 0 \qquad \text{for any } f \in \mathfrak{C}^0(\Gamma,\mathbb{R}).$$

We set $\mathbb{M} := \{\mu \in \mathbb{P} : \mu \text{ is closed}\}$

Remark 4.8. **(i)** Given $\mu \in \mathbb{P}$, we have for any $g \in \mathfrak{C}^0(\Gamma,\mathbb{R})$

$$\int \mathrm{d}g \,\mathrm{d}\mu = \langle \mathrm{d}g, \rho(\mu)\rangle,$$

hence

$$\mu \text{ is closed} \iff \partial\rho(\mu) = 0 \iff \rho \in H_1(\Gamma,\mathbb{R}),$$

namely $\rho(\mu)$ is a 1–cycle. We call it *rotation vector* (or *Schwartzman asymptotic cycle*) of μ. This should be compared with the corresponding classical definitions in Aubry–Mather theory (see [10, 30]).
(ii) Given $\mu \in \mathbb{M}$ and $\omega \in \mathfrak{C}^1(\Gamma,\mathbb{R})$, it follows from the definition of closed measure and (13) that

$$\int \omega \,\mathrm{d}\mu = \langle [\omega], \rho(\mu)\rangle,$$

i.e. it only depends on the cohomology class $[\omega] \in H^1(\Gamma,\mathbb{R})$.

Proposition 4.9. *The subset $\mathbb{M} \subset \mathbb{P}$ is convex and closed in the Wasserstein topology.*

Proof. The convexity property is obvious. Let μ_n be a sequence of closed probability measures converging in the Wasserstein sense to μ. We consider $g \in \mathfrak{C}^0(\Gamma,\mathbb{R})$, then associating to $\mathrm{d}g$ the continuous function on $T\Gamma$ with linear growth $(e,q) \longmapsto \langle \mathrm{d}g, q\,e\rangle$ and taking into account (13), we get

$$\int \mathrm{d}g \,\mathrm{d}\mu_n \to \int \mathrm{d}g \,\mathrm{d}\mu.$$

This concludes the proof. □

Let us define the map $\rho : \mathbb{M} \longrightarrow H_1(\Gamma,\mathbb{R})$ that to any closed probability measure μ associates its rotation vector $\rho(\mu)$ (see remark 4.8 **(i)**). One proves the following properties.

Proposition 4.10. *The map ρ is continuous and affine (for convex combinations), i.e. for every $\lambda \in [0,1]$ and $\mu_1, \mu_2 \in \mathbb{M}$*

$$\rho(\lambda\mu_1 + (1-\lambda)\mu_2) = \lambda\rho(\mu_1) + (1-\lambda)\rho(\mu_2).$$

In particular, it is surjective.

Proof. Let us first prove continuity. If $\mu_n \to \mu$ in $\mathbb{M}$ and ω is any element of $\mathfrak{C}^1(\Gamma, \mathbb{R})$ with cohomology class c, then associating to ω the continuous function on $T\Gamma$ with linear growth $(e,q) \longmapsto \langle \omega, qe \rangle$ and taking into account (13), we have that if μ_n converges to μ in the Wasserstein sense then

$$\langle c, \rho(\mu_n) \rangle = \int \omega \, \mathrm{d}\mu_n \longrightarrow \int \omega \, \mathrm{d}\mu = \langle c, \rho(\mu) \rangle.$$

Since c has been arbitrarily chosen in $H^1(\Gamma, \mathbb{R})$, $\rho(\mu_n) \longrightarrow \rho(\mu)$ as $n \to +\infty$, which proves continuity.

The fact that the map ρ is affine (under convex combination) is an immediate consequence of the definition of the rotation vector.

Finally, let us prove surjectivity. Let $h \in H_1(\Gamma, \mathbb{R})$ given by $h = \sum_{i=1}^N a_i e_i$, with $\partial(h) = 0$; we can assume that $a_i > 0$ (otherwise we substitute e_i with $-e_i$). Then, it is sufficient to consider the measure $\mu = \sum_{i=1}^N \frac{1}{N} \delta(e_i, Na_i)$, where $\delta(e,q)$ denotes Dirac delta at (e,q); μ is closed since $\partial(h) = 0$ and one can use (14) to check that

$$\rho(\mu) = \sum_{i=1}^N \frac{Na_i}{N} e_i = \sum_{i=1}^N a_i e_i = h.$$

□

4.4. Occupation measures

Let us introduce the notion of *occupation measure*, which can be thought as a measure representation of a parametrized path.

Definition 4.11. Given a parametrized path $\xi = (e_i, q_i, T_i)_{i=1}^M$, the associated occupation measure is defined as

$$\mu_\xi := \frac{1}{T_\xi} \sum_{i=1}^M T_i \, \delta(e_i, q_i), \tag{15}$$

where $T_\xi = \sum_{i=1}^M T_i$ and $\delta(e,q)$ denotes Dirac delta concentrated on the point (e,q).

Remark 4.12. **(i)** Taking into account that an edge e can be equal to e_i for different values of the index i, we see that an occupation measure restricted to any edge is the convex combination of Dirac measures.

(ii) For any $e \in \mathbf{E}$, $\delta(e,0)$ is a closed occupation measure corresponding to the equilibrium circuit based on e with vanishing speed and any pair of positive numbers as time parametrization.

Occupation measures are not necessarily closed. However, it is possible to characterize closed ones.

Proposition 4.13. *Let μ_ξ be an occupation measure associated to a parametrized path $\xi = \{(e_i, q_i, T_i)\}_{i=1}^M$. Then, μ_ξ is closed if and only if ξ is a parametrized cycle.*

Proof. Let $g \in \mathfrak{C}^0(\Gamma, \mathbb{R})$. Observe that for every $e \in \mathbf{E}$

$$\int \mathrm{d}g \, \mathrm{d}\delta(e, 0) = 0$$

since we are integrating the function $\langle \mathrm{d}g, qe \rangle$ with respect to $\delta(e, 0)$. The statement is trivial if all q_i vanish (see proposition 4.3). Let us assume that some $q_i \neq 0$; then, recalling definition 4.1 and proposition 4.3:

$$\begin{aligned} \int \mathrm{d}g \, \mathrm{d}\mu_\xi &= \frac{1}{T_\xi} \sum_{i=1}^{M} T_i \int \mathrm{d}g \, \mathrm{d}\delta(e_i, q_i) = \frac{1}{T_\xi} \sum_{i | q_i \neq 0} T_i \langle \mathrm{d}g, q_i e_i \rangle \\ &= \frac{1}{T_\xi} \sum_{i | q_i \neq 0} (g(\mathrm{t}(e_i)) - g(\mathrm{o}(e_i))) = \frac{1}{T_\xi} (g(\mathrm{t}(\xi)) - g(\mathrm{o}(\xi))). \end{aligned}$$

Therefore, μ_ξ is closed if and only if $g(\mathrm{t}(\xi)) = g(\mathrm{o}(\xi))$ for every $g \in \mathfrak{C}^0(\Gamma, \mathbb{R})$, which is equivalent to $\mathrm{t}(\xi) = \mathrm{o}(\xi)$, i.e. ξ is a parametrized cycle. □

Remark 4.14. Given a parametrized cycle ξ, we have (see (2) for the definition of $[\xi]$)

$$\rho(\mu_\xi) = \frac{1}{T_\xi} \sum_{i=1}^{M} e_i = \frac{[\xi]}{T_\xi}.$$

We close this section with a density result. This theorem is well known for measures on the tangent bundle of a manifold, a piece of folklore according to [3]. We will not use it in the rest of the paper, however we include it for two reasons: firstly, it somehow validates our previous definition of occupation measures, secondly because the proof, which follows the same lines of [3, theorem 31], is simple and illuminating, and represents a nice application of weak KAM theory on graphs to the analysis of closed probability measures.

Theorem 4.15. *The set of closed occupation measures is dense in* $\mathbb{M}$.

The proof is in appendix B.

5. Mather's theory on graphs

Mather theory is about the minimization of the action functional

$$\mu \longmapsto \int \mathcal{L}^\omega \, \mathrm{d}\mu$$

on suitable subsets of closed probability measures. Results and definitions of this section are inspired by the corresponding ones in the classical Mather theory, see [10], [30]. We provide full details to make the text self–contained.

5.1. Existence of minimizers

We recall the main compactness criterion in the Wasserstein space $\mathbb{P}$ (see, for example, [32]).

- A subset $\mathbb{K} \subset \mathbb{P}$ is relatively compact if and only if for any $\varepsilon > 0$ there exists a compact subset K_ε of $T\Gamma$ such that

$$\int_{K_\varepsilon^c} q \, \mathrm{d}\mu < \varepsilon \qquad \text{for any } \mu \in \mathbb{K},$$

  where K_ε^c stands for the complement of K_ε in $T\Gamma$.

From the superlinearity property of $\mathcal{L}$, we derive the following property.

Proposition 5.1. *Given $a \in \mathbb{R}$, the set*

$$\mathbb{K}_a := \left\{ \mu \in \mathbb{M} \mid \int \mathcal{L}\, \mathrm{d}\mu \leqslant a \right\}$$

is compact in $\mathbb{M}$.

Proof. Assume that $K_a \neq \emptyset$, otherwise there is nothing to prove. According to the compactness criterion in the Wasserstein space $\mathbb{M}$ and the definition of $T\Gamma$, it is enough to prove that, given $\varepsilon > 0$, there exists $M_\varepsilon > 0$ such that

$$\int_{\mathbb{R}_e^+ \cap (M_\varepsilon, +\infty)} q\, \mathrm{d}\mu < \varepsilon \qquad \text{for any } e \in \mathbf{E}, \mu \in \mathbb{K}_a.$$

If this is not the case, we find $\varepsilon > 0$, $e_0 \in \mathbf{E}$, a sequence of numbers $M_n \to +\infty$ and a sequence of measures

$$\mu^{(n)} = \sum_{e \in \mathbf{E}} \lambda_e^{(n)}\, \mu_e^{(n)} \in \mathbb{K}_a$$

such that

$$\int_{M_n}^{+\infty} q\, \mathrm{d}\mu_{e_0}^{(n)} \geqslant \varepsilon \qquad \text{for any } n.$$

Taking into account that $\mathcal{L}(e_0, \cdot)$ is superlinear, we find another positively diverging sequence h_n satisfying

$$\mathcal{L}(e_0, q) \geqslant h_n\, q \qquad \text{for } q \geqslant M_n.$$

Since edges are finitely many, we can find a constant b such that

$$\begin{aligned} a &\geqslant \int \mathcal{L}(e,q)\, \mathrm{d}\mu^{(n)} \geqslant \int_{M_n}^{+\infty} \mathcal{L}(e_0,q)\, \mathrm{d}\mu_{e_0}^{(n)} + b \\ &\geqslant h_n \int_{M_n}^{+\infty} q\, \mathrm{d}\mu_{e_0}^{(n)} + b \geqslant h_n\, \varepsilon + b, \end{aligned}$$

which, as n goes to $+\infty$, leads to a contradiction. □

As a consequence:

Corollary 5.2. *The action functional $\mu \longmapsto \int \mathcal{L}\, \mathrm{d}\mu$ is lower semicontinuous on $\mathbb{M}$.*

This in turn implies:

Theorem 5.3. ***(i)*** *The action functional admits minimum in $\mathbb{M}$;*
(ii) *Given $h \in H_1(\Gamma, \mathbb{R})$, the action functional admits minimum in $\rho^{-1}(h)$.*

Proof. Recall that a lower-semicontinuous function admits minimum on compact sets. Therefore, **(i)** follows from proposition 5.1 and corollary 5.2. Similarly, **(ii)** follows from proposition 5.1, corollary 5.2, and the fact that $\rho^{-1}(h)$ is closed in $\mathbb{M}$ (the map $\rho : \mathbb{M} \to H_1(\Gamma, \mathbb{R})$ is continuous in force of proposition 4.10). □

5.2. *Mather's minimal average actions and Mather measures*

We define *Mather's* β*–function* as:

$$\begin{aligned} \beta : H_1(\Gamma, \mathbb{R}) &\longrightarrow \mathbb{R} \\ h &\longmapsto \min_{\mu \in \rho^{-1}(h)} \int \mathcal{L}\, \mathrm{d}\mu. \end{aligned} \tag{16}$$

The above minimum does exist in force of theorem 5.3 **(ii)**.

Definition 5.4. We say that a measure $\mu \in \mathbb{M}$ is a *Mather measure with homology h* if $\int \mathcal{L}\, \mathrm{d}\mu = \beta(h)$. We denote the subset of these measures by $\mathbb{M}^h$.

We define the *Mather set of homology h* as

$$\widetilde{\mathcal{M}}^h := \bigcup_{\mu \in \mathbb{M}^h} \operatorname{supp} \mu \subset T\Gamma, \tag{17}$$

where supp μ denotes the support of μ in $T\Gamma$. This set is closed[3].

Properties of β:

- β is convex. In fact, let $h_1, h_2 \in H_1(\Gamma, \mathbb{R})$, $\lambda \in [0, 1]$ and let us consider $\mu_i \in \mathbb{M}^{h_i}$ for $i = 1, 2$. If follows from proposition 4.10 that

  $$\rho(\lambda\mu_1 + (1-\lambda)\mu_2) = \lambda h_1 + (1-\lambda) h_2.$$

  Moreover, using the linearity of the integral and the definition of β, we obtain:

  $$\begin{aligned} \beta(\lambda h_1 + (1-\lambda)h_2) &\leqslant \int \mathcal{L}\, d(\lambda\mu_1 + (1-\lambda)\mu_2) \\ &= \lambda \int \mathcal{L}\, \mathrm{d}\mu_1 + (1-\lambda) \int \mathcal{L}\, \mathrm{d}\mu_2 \\ &= \lambda\beta(h_1) + (1-\lambda)\beta(h_2). \end{aligned}$$

- β is superlinear. This could be proved directly by using the superlinearity of $\mathcal{L}$; however, we deduce it from the finiteness of its convex conjugate α (see (18) and remark 5.5).

We consider the convex conjugate of β, that we shall call *Mather's* α*-function*:

$$\begin{aligned} \alpha : H^1(\Gamma, \mathbb{R}) &\longrightarrow \mathbb{R} \\ c &\longmapsto \max_{h \in H_1(\Gamma, \mathbb{R})} \left(\langle c, h \rangle - \beta(h) \right), \end{aligned}$$

where $\langle c, h \rangle$ denotes the pairing between $H^1(\Gamma, \mathbb{R})$ and $H_1(\Gamma, \mathbb{R})$ defined in section 2.4.

[3] Classically, one defines this set as the closure of the union of the supports of Mather measures; however, it is easy to check that it is already closed, see [30, remark 3.1.11 (i)]. In this setting, it is even easier to prove this, since, as it will follow from theorem 6.1 and the graph property (corollary 6.4), this set is a finite union of closed sets.

One can also characterize α in a variational way, which shows that it is finite everywhere:

$$\begin{aligned}\alpha(c) &= \max_{h\in H_1(\Gamma,\mathbb{R})} (\langle c,h\rangle - \beta(h)) \qquad (18)\\ &= \max_{h\in H_1(\Gamma,\mathbb{R})} \left(\langle c,h\rangle - \min_{\mu\in\rho^{-1}(h)} \int \mathcal{L}\,\mathrm{d}\mu\right)\\ &= -\min_{h\in H_1(\Gamma,\mathbb{R})} \left(\min_{\mu\in\rho^{-1}(h)} \left(\int \mathcal{L}\,\mathrm{d}\mu - \langle c,\rho(\mu)\rangle\right)\right)\\ &= -\min_{\mu\in\mathbb{M}} \int \mathcal{L}^\omega\,\mathrm{d}\mu,\end{aligned}$$

where $\omega \in \mathfrak{C}^1(\Gamma,\mathbb{R})$ has cohomology class c. Due to the superlinearity of $\mathcal{L}^\omega$, we see, arguing as in proposition 5.1, that the sublevels of $\mathcal{L}^\omega$ are compact in the Wasserstein topology, and consequently by proposition 4.9 the minimum in the above formula does exist. Therefore α is finite, convex with convex conjugate equal to β.

Remark 5.5. The fact that α is finite, convex with convex conjugate equal to β, implies that β has superlinear growth. In fact, a convex function on finite dimensional vector spaces possesses a finite convex conjugate if and only if it has superlinear growth, see [26].

Definition 5.6. Given c in $H^1(\Gamma,\mathbb{R})$ and ω in the class c, we say that a measure $\mu\in\mathbb{M}$ is a *Mather measure with cohomology c* if $\int \mathcal{L}^\omega\,\mathrm{d}\mu = -\alpha(c)$ (observe that being μ closed, this notion does not depend on the choice of the representative ω, but only on its cohomology class). We denote the subset of these measures by $\mathbb{M}_c$.

We define the *Mather set of cohomology c* as

$$\widetilde{\mathcal{M}}_c := \bigcup_{\mu\in\mathbb{M}_c} \mathrm{supp}\mu \subset T\Gamma, \qquad (19)$$

where $\mathrm{supp}\mu$ denotes the support of μ in $T\Gamma$. This set is also closed.

As a consequence of proposition 5.1, we have

Proposition 5.7. *For any $h\in H_1(\Gamma,\mathbb{R})$, $c\in H^1(\Gamma,\mathbb{R})$, the sets of Mather measures $\mathbb{M}^h$, $\mathbb{M}_c$ are compact, convex subsets of $\mathbb{M}$.*

Next proposition will help clarify the relation between the two notions of Mather measures in definitions 5.4 and 5.6. To state it, recall that, like any convex function on a finite-dimensional space, β admits a subdifferential at each point $h\in H_1(\Gamma,\mathbb{R})$, i.e. we can find $c\in H^1(\Gamma,\mathbb{R})$ such that $\beta(h')-\beta(h)\geqslant\langle c,h'-h\rangle$ for any $h\in H_1(\Gamma,\mathbb{R})$. We will denote by $\partial\beta(h)$ the set of $c\in H^1(\Gamma,\mathbb{R})$ that are subdifferentials of β at h. Similarly, we will denote by $\partial\alpha(c)$ the set of subdifferentials of α at c.

Fenchel's duality implies an easy characterization of subdifferentials (see for example [30, proposition 3.3.3]):

$$c\in\partial\beta(h) \iff h\in\partial\alpha(c) \iff \langle c,h\rangle = \alpha(c)+\beta(h). \qquad (20)$$

The next proposition can be proven as the corresponding ones in the classical Mather theory, with obvious adaptations (we omit the proof, see for example [30, proposition 3.3.4]).

Proposition 5.8.

(i) *$\mu\in\mathbb{M}$ is a Mather measure with homology h if and only if*

$$\mu\in\mathbb{M}_c \qquad \textit{for any } c\in\partial\beta(h).$$

(ii) *For every* $c \in H^1(\Gamma, \mathbb{R})$

$$\partial\alpha(c) = \{\rho(\mu) \mid \mu \in \mathbb{M}_c\}.$$

Corollary 5.9. *If* $c \in \partial\beta(h)$*, then* $\widetilde{\mathcal{M}}^h \subseteq \widetilde{\mathcal{M}}_c$*. In particular:*

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

Remark 5.10. We will say that μ is a Mather measure *tout court*, if it is a Mather measure for some cohomology c, or equivalently it is a Mather measure of homology $\rho(\mu)$.

6. Properties of Mather measures

6.1. Structural properties and Mather's graph property

Exploiting the strict convexity of $\mathcal{L}(e, \cdot)$, we can derive this first property of Mather measures, namely that they consist of a finite convex combinations of Dirac deltas, in particular each edge appears at most once.

Theorem 6.1. *The restriction of any Mather measure to an edge of its support is concentrated on a point.*

Proof. Let $\mu = \sum_{e \in \mathbf{E}} \lambda_e \mu_e$ be a Mather measure. We set

$$\nu := \sum_{e \in \mathbf{E}} \lambda_e \, \delta\left(e, \int_0^{+\infty} q \, \mathrm{d}\mu_e\right).$$

Thanks to the convexity of $\mathcal{L}(e, \cdot)$ for each $e \in \mathbf{E}$, we can apply Jensen inequality to μ_e and get

$$\begin{aligned} \int \mathcal{L}(e,q) \, \mathrm{d}\mu &= \sum_{e \in \mathbf{E}} \lambda_e \int_0^{+\infty} \mathcal{L}(e,q) \, \mathrm{d}\mu_e \geqslant \sum_{e \in \mathbf{E}} \lambda_e \, \mathcal{L}\left(e, \int_0^{+\infty} q \, \mathrm{d}\mu_e\right) \\ &= \int \mathcal{L}(e,q) \, \mathrm{d}\nu. \end{aligned}$$

Observe that $\rho(\mu) = \rho(\nu)$; hence, due to the strict convexity of $\mathcal{L}(e, \cdot)$ for each $e \in \mathbf{E}$ and the fact that μ is a Mather measure, we conclude that equality must prevail in the above formula, and this is possible if and only if $\mu = \nu$. □

Proposition 6.2. *Let* $c \in H^1(\Gamma, \mathbb{R})$ *and* $h \in H_1(\Gamma, \mathbb{R})$.

(i) *If* $(f, q_1), (f, q_2) \in \widetilde{\mathcal{M}}_c$ *(resp.* $\widetilde{\mathcal{M}}^h$*) for some* $f \in \mathbf{E}$*, then* $q_1 = q_2$.

(ii) *If* $(f, q_1), (-f, q_2) \in \widetilde{\mathcal{M}}_c$ *(resp.* $\widetilde{\mathcal{M}}^h$*) for some* $f \in \mathbf{E}$*, then* $q_1 = q_2 = 0$ *and* $\alpha(c) = \min \alpha$.

Proof. Since, by corollary 5.9, $\widetilde{\mathcal{M}}^h$ is contained in some $\widetilde{\mathcal{M}}_c$, then it suffices to prove the property for the latter.

Let (f, q_1), $(f, q_2) \in \widetilde{\mathcal{M}}_c$; then, by theorem 6.1 there are two Mather measures $\mu = \sum_{e \in \mathbf{E}} \lambda_e \mu_e$, $\nu = \sum_{e \in \mathbf{E}} \tau_e \nu_e$ in $\mathbb{M}_c$ such that

$$\lambda_f > 0, \; \tau_f > 0 \qquad \text{and} \qquad \mu_f = \delta(f, q_1), \; \nu_f = \delta(f, q_2).$$

Due to the convexity of $\mathbb{M}_c$ (see proposition 5.7), we have that $\frac{1}{2}\mu+\frac{1}{2}\nu$ is in $\mathbb{M}_c$, and the restriction of it on f is a convex combination with positive coefficients of $\delta(f,q_1)$ and $\delta(f,q_2)$. We then derive, again from theorem 6.1, that $q_1=q_2$, which concludes the proof of item **(i)**.

We proceed by proving **(ii)**. Let (f,q_1), $(-f,q_2)\in\widetilde{\mathcal{M}}_c$; then, there exists $\mu\in\mathbb{M}_c$ such that $f,-f\in\mathrm{supp}_{\mathbf{E}}\,\mu$; in fact, by definition 5.6, there exist $\mu_1,\mu_2\in\mathbb{M}_c$ such that $(f,q_1)\in\mathrm{supp}\mu_1$ and $(-f,q_2)\in\mathrm{supp}\mu_2$, hence it suffices to consider $\mu=\frac{1}{2}\mu_1+\frac{1}{2}\mu_2$, which still belongs to $\mathbb{M}_c$ (due to convexity, see proposition 5.7).

Let us define

$$\tilde{\mu}:=\frac{1}{1-(\lambda_1+\lambda_2)}\left(\mu-\lambda_1\delta(f,q_1)-\lambda_2\delta(-f,q_2)\right)$$

with $\lambda_1,\lambda_2\in(0,1)$, $q_1,q_2\geqslant 0$, so that μ can be written as

$$\mu=\lambda_1\delta(f,q_1)+\lambda_2\delta(-f,q_2)+(1-\lambda_1-\lambda_2)\tilde{\mu}.$$

Note that $\pm f\notin\mathrm{supp}_{\mathbf{E}}\tilde{\mu}$ because of theorem 6.1.

Assume, without any loss of generality, that $\lambda_1q_1\geqslant\lambda_2q_2$ (otherwise, invert the roles of f and $-f$) and define

$$\overline{q}:=\frac{\lambda_1q_1-\lambda_2q_2}{\lambda_1+\lambda_2}=\frac{\lambda_1}{\lambda_1+\lambda_2}q_1+\frac{\lambda_2}{\lambda_1+\lambda_2}(-q_2)\geqslant 0. \tag{21}$$

Consider the new measure

$$\nu:=(\lambda_1+\lambda_2)\,\delta(f,\overline{q})+(1-(\lambda_1+\lambda_2))\tilde{\mu}.$$

Clearly, ν is a probability measure and it is also closed; in fact:

$$\begin{aligned}\rho((\lambda_1+\lambda_2)\,\delta(f,\overline{q}))&=(\lambda_1+\lambda_2)\overline{q}f\;=\;(\lambda_1q_1-\lambda_2q_2)f\\&=\rho\left(\lambda_1\delta(f,q_1)+\lambda_2\delta(-f,q_2)\right),\end{aligned}$$

hence, $\rho(\nu)=\rho(\mu)$ is a 1-cycle, which implies that ν is closed (see remark 4.8 **(i)**).

In order to get a contradiction, we want to prove that the action of ν is less than the action of μ, thus contradicting minimality of μ. In fact:

$$\begin{aligned}\int\mathcal{L}\,\mathrm{d}\nu-\int\mathcal{L}\,\mathrm{d}\mu&=(\lambda_1+\lambda_2)\,\mathcal{L}(f,\overline{q})-\lambda_1\,\mathcal{L}(f,q_1)-\lambda_2\,\mathcal{L}(-f,q_2)\\&=(\lambda_1+\lambda_2)\left(\mathcal{L}(f,\overline{q})-\frac{\lambda_1}{\lambda_1+\lambda_2}\mathcal{L}(f,q_1)-\frac{\lambda_1}{\lambda_1+\lambda_2}\mathcal{L}(f,-q_2)\right)\\&\leqslant 0,\end{aligned} \tag{22}$$

where in the last inequality we have used the convexity of $\mathcal{L}(f,\cdot)$; taking into account that $\mathcal{L}(f,\cdot)$ is in addition strictly convex, we see that a strict inequality prevails in (22), leading to a contradiction, unless

$$\overline{q}=q_1=-q_2\qquad\Longleftrightarrow\qquad q_1=q_2=0.$$

The property that $\alpha(c)=\min\alpha$ follows from the fact that $\delta(f,0)$ belongs to $\mathbb{M}_c$, hence $0\in\partial\alpha(c)$ (see proposition 5.8 (ii)). Being α convex implies that $\alpha(c)$ is the minimum of α.

□

We can now derive a central property that can be read as an instance of the celebrated *Mather's graph theorem* (see [22, theorem 2]) in the graph setting[4].

To state it more precisely, let us introduce the *projection* $\pi_{\mathbf{E}} : T\Gamma \to \mathbf{E}$ defined as

$$\pi_{\mathbf{E}}(e,q) := \begin{cases} e & \text{if } q > 0 \\ \{e,-e\} & \text{if } q = 0. \end{cases}$$

Remark 6.3. Observe that the projection $\pi_{\mathbf{E}}$ that we have defined is multivalued at some points: this is needed in order to cope with the fact that the elements $(e,0)$, $(-e,0)$ are identified in $T\Gamma$, for any $e \in \mathbf{E}$.

Alternatively, one could consider $\pi_{\mathbf{E}^+} : T\Gamma \to \mathbf{E}^+$, denoting the projection on a given orientation $\mathbf{E}^+$ of the graph (namely, $\pi_{\mathbf{E}^+}(\pm e,q) = e$ for any $e \in \mathbf{E}^+$). In the light of proposition 6.2, the graph property in corollary 6.4 continues to hold with such a projection and all related results can be suitably restated.

Corollary 6.4 ***(Mather graph property).*** *The restriction of $\pi_{\mathbf{E}}$ to $\widetilde{\mathcal{M}}_c$ and $\widetilde{\mathcal{M}}^h$ is injective for every $c \in H^1(\Gamma,\mathbb{R})$, $h \in H_1(\Gamma,\mathbb{R})$.*

Proof. Since, by corollary 5.9, $\widetilde{\mathcal{M}}^h$ is contained in some $\widetilde{\mathcal{M}}_c$, then it suffices to prove the property for the latter. The result then follows from proposition 6.2 **(i)**. □

Remark 6.5. It follows from corollary 6.4 that for any $c \in H^1(\Gamma,\mathbb{R})$

$$\left(\pi_{\mathbf{E}}|\widetilde{\mathcal{M}}_c\right)^{-1} : \pi_{\mathbf{E}}\left(\widetilde{\mathcal{M}}_c\right) \longrightarrow \widetilde{\mathcal{M}}_c$$

is a well-defined map. In section 8 we will describe this function more explicitly (see theorem 8.5).

Next result is an important step in our analysis. It puts in relation, via theorem 6.1, Mather and occupation measures.

Theorem 6.6. *A closed probability measure, whose restriction on any edge is concentrated on a point, is a convex combination of occupation measures based on circuits.*

Proof. Let

$$\mu = \sum_{e\in\mathbf{E}} \lambda_e\, \delta(e,q_e) \tag{23}$$

with $\lambda_e \geqslant 0$ and $\sum \lambda_e = 1$, be a measure as indicated in the statement. We first assume that $q_e \neq 0$ for any e. We argue by finite induction on the cardinality of $\mathrm{supp}_{\mathbf{E}}\mu$ indicated by $|\mathrm{supp}_{\mathbf{E}}\mu|$. By taking the function which is equal to 1 at a given vertex x and 0 elsewhere, and exploiting that μ is closed, we deduce that the relation

$$\sum_{e\in\mathbf{E}_x} \lambda_e\, q_e = \sum_{e\in-\mathbf{E}_x} \lambda_e\, q_e \qquad \forall\, x \in \mathbf{V}. \tag{24}$$

If $|\mathrm{supp}_{\mathbf{E}}\mu| = 2$, set $\mathrm{supp}_{\mathbf{E}}\mu = \{e,f\}$. By applying (24) to $x = \mathrm{o}(e)$, $x = \mathrm{t}(e)$, we realize that (e,f) makes up a circuit and

$$\lambda\, q_e = (1-\lambda)\, q_f \qquad \text{for some } \lambda \in (0,1).$$

[4] Ironically, the term *graph* appearing twice in this sentence, is used with two completely distinct meanings.

This implies that

$$\lambda = \frac{q_f}{q_e + q_f} = \frac{1}{q_e}\,\frac{q_e\, q_f}{q_e + q_f} = \frac{1/q_e}{\frac{1}{q_e} + \frac{1}{q_f}} \quad \text{and} \quad 1 - \lambda = \frac{1/q_f}{\frac{1}{q_e} + \frac{1}{q_f}}.$$

This implies that μ is the occupation measure corresponding to the parametrized circuit $((e, q_e, 1/q_e), (f, q_f, , 1/q_f))$.

Let us now assume the assertion true for measures with support of cardinality less than a given M, and assume $|\mathrm{supp}_{\mathbf{E}}\mu| = M \geqslant 3$. Starting by any edge $e \in \mathrm{supp}_{\mathbf{E}}\mu$, we choose one of the edges $f \in \mathrm{supp}_{\mathbf{E}}\mu$ with

$$\mathrm{t}(e) = \mathrm{o}(f)$$

and we call it $\pi_1(e)$. This choice is possible, for any initial e, because of (24). We iterate the procedure starting from $\pi_1(e)$ to define $\pi_2(e)$. Taking again into account (24), we see that we can go on until we reach $\pi_k(e)$ with

$$\mathrm{t}(\pi_k(e)) = \mathrm{o}(\pi_h(e)) \qquad \text{for some } h \leqslant k.$$

The edges

$$\{\pi_h(e),\ \pi_{h+1}(e),\ \ldots,\ \pi_k(e)\}$$

make up a circuit contained in $\mathrm{supp}_{\mathbf{E}}\mu$. We set $M' = k + 1 - h$,

$$e_i = \pi_{h+i-1}(e), \quad \lambda_i = \lambda_{e_i} \quad q_i = q_{e_i} \qquad \text{for } i = 1, \ldots, M'$$

and consider the parametrized circuit $\xi = (e_i, q_i, 1/q_i)_{i=1}^{M'}$. The associated occupation measure is

$$\mu_\xi = \frac{1}{T_\xi} \sum_{i=1}^{M'} \frac{1}{q_i}\, \delta(e_i, q_i), \tag{25}$$

where $T_\xi = \left(\sum_{i=1}^{M'} \frac{1}{q_i} \right)$. We distinguish two cases:

- If $M = M'$ we show that $\mu = \mu_\xi$, which proves the claim. In fact, in this case for any vertex x of the graph there is an alternative: either no edge in $\mathrm{supp}_{\mathbf{E}}\mu$ is incident on it or there are exactly two incident edges, one with x as initial point and the other with x as terminal point. By applying (24) we deduce

$$\lambda_i\, q_i = \lambda_j\, q_j =: A \qquad \text{for any } i, j \in \{1, \ldots, M'\}. \tag{26}$$

  This implies that $\lambda_i = \frac{A}{q_i}$ for any i, and, since $\sum_i \lambda_i = 1$ we obtain

$$A = \left(\sum_i \frac{1}{q_i} \right)^{-1} = \frac{1}{T_\xi}.$$

  By exploiting the above relation plus (23), (25), (26) we obtain

$$\begin{aligned} \mu_\xi &= \frac{1}{T_\xi} \sum_{i=1}^{M} \frac{1}{q_i}\, \delta(e_i, q_i) = \sum_{i=1}^{M} \frac{1}{T_\xi\, q_i}\, \delta(e_i, q_i) \\ &= \sum_{i=1}^{M} \frac{A}{q_i}\, \delta(e_i, q_i) = \sum_{i=1}^{M} \lambda_i\, \delta(e_i, q_i) = \sum_{e \in \mathrm{supp}_{\mathbf{E}}\mu} \lambda_e \delta(e, q_e) = \mu. \end{aligned}$$

- Let us assume now that $M' < M$ and define

  $$\lambda = T \min_i q_i \, \lambda_i.$$

  Observe that

  $$\frac{\lambda}{T q_i} \leqslant \lambda_i \qquad \text{for any } i \in \{1, \ldots, M'\}$$

  and consequently

  $$\lambda = \lambda \sum_i \frac{1}{T q_i} \leqslant \sum_i \lambda_i < 1,$$

  where the rightmost strict inequality comes from the fact that $M' < M$. Let us define the following probability measure

  $$\nu = \frac{1}{1-\lambda} \left[\sum_{i=1}^{M'} \left(\lambda_i - \lambda \frac{1}{T q_i} \right) \delta(e_i, q_i) + \sum_{e \notin \mathrm{supp}_{\mathbf{E}} \mu_\xi} \lambda_e \, \delta(e, q_e) \right].$$

  This is actually a probability measure since

  $$\sum_i \left(\lambda_i - \lambda \frac{1}{T q_i} \right) + \sum_{e \notin \mathrm{supp}_{\mathbf{E}} \mu_\xi} \lambda_e = \sum_{e \in \mathrm{supp}_{\mathbf{E}} \mu} \lambda_e - \lambda \frac{1}{T} \sum_i \frac{1}{q_i} = 1 - \lambda.$$

  Moreover

  $$\begin{aligned} &\lambda \mu_\xi + (1-\lambda) \nu \\ &\quad = \lambda \left[\frac{1}{T} \sum_i \frac{1}{q_i} \delta(e_i, q_i) \right] + \sum_i \left(\lambda_i - \lambda \frac{1}{T q_i} \right) \delta(e_i, q_i) + \sum_{e \notin \mathrm{supp}_{\mathbf{E}} \mu_\xi} \lambda_e \, \delta(e, q_e) \\ &\quad = \sum_{e \in \mathrm{supp}_{\mathbf{E}} \mu} \lambda_e \, \delta(e, q_e) = \mu. \end{aligned} \tag{27}$$

  We see from (27) that ν is closed since both μ and μ_ξ are closed. In addition, some of the coefficients $\lambda_i - \lambda \frac{1}{T q_i}$ must vanish by the very definition of λ. The support of ν has then cardinality less than M, and by inductive assumption ν is the convex combination of occupation measures based on circuits. The same holds true for μ in force of (27).

Let us now discuss the case in which some of the q_e's vanish. Let μ be as in (23) and define

$$E = \{ e \in \mathrm{supp}_{\mathbf{E}} \mu \mid q_e > 0 \}, \quad F = \{ f \in \mathrm{supp}_{\mathbf{E}} \mu \mid q_f = 0 \}, \quad \lambda_F = \sum_{f \in F} \lambda_f.$$

If $E = \emptyset$, then $\mu = \delta(e, 0)$ for a suitable $e \in \mathbf{E}$ and this measure is supported by the equilibrium circuit based on e, so that the assertion is proved. We then assume that both E and F are nonempty. We consider the probability measure

$$\nu = \sum_{e \in E} \frac{\lambda_e}{1 - \lambda_F} \delta(e, q_e)$$

and derive

$$\mu = (1 - \lambda_F) \nu + \sum_{f \in F} \lambda_f \delta(f, 0).$$

By the first part of the proof there exist occupation measures μ_{ξ_i} corresponding to circuits ξ_i with

$$\nu = \sum_i \sigma_i \, \mu_{\xi_i} \qquad \sigma_i > 0, \sum_i \sigma_i = 1.$$

Summing, up we have

$$\mu = (1-\lambda_F) \sum_i \sigma_i \, \nu_i + \sum_{f \in F} \lambda_f \delta(f,0).$$

This concludes the proof. □

Remark 6.7. In case Γ contains loops, the proof of theorem 6.6 should include also the case $\mu = \delta(e,q_e)$ with $q_e \neq 0$; in this case, μ is closed if and only if e is a loop, and it follows easily that μ is an occupation measure on the trivial circuit e.

6.2. Irreducible Mather measures

A point in a convex set is called *extremal* if it cannot be obtained as convex combination of two distinct elements of the set.

A closed probability measure is said to be *irreducible* if it is extremal in $\mathbb{M}$.

Proposition 6.8. *A Mather measure is irreducible if and only if it is an occupation measure corresponding to a parametrized circuit.*

Proof. Let μ be a Mather measure. If it is not an occupation measure supported by a parametrized circuit, then by theorems 6.6 and 6.1 it must be the convex combination of distinct occupation measures supported on parametrized circuits. This proves that it is not irreducible.

Conversely, assume for the purpose of contradiction that μ is an occupation measure supported on a parametrized circuit and that it is not irreducible. Hence, there exist $\mu_1 \neq \mu_2$ in $\mathbb{M}$, $\lambda \in (0,1)$ such that

$$\mu = (1-\lambda)\,\mu_1 + \lambda\,\mu_2.$$

This implies by proposition 4.10 that

$$\rho(\mu) = (1-\lambda)\,\rho(\mu_1) + \lambda\,\rho(\mu_2).$$

We thus have

$$\begin{aligned} \beta(\rho(\mu)) = \int \mathcal{L}\,\mathrm{d}\mu = (1-\lambda) \int \mathcal{L}\,\mathrm{d}\mu_1 + \lambda \int \mathcal{L}\,\mathrm{d}\mu_2 \\ \geqslant (1-\lambda)\,\beta(\rho(\mu_1)) + \lambda\,\beta(\rho(\mu_2)), \end{aligned}$$

due to the convex character of β, equality must prevail in the above formula, so that both μ_1 and μ_2 are Mather measures. Taking again into account theorems 6.6 and 6.1, we find an occupation measure ν supported on a parametrized circuit with $\mathrm{supp}_{\mathbf{E}}\nu$ proper subset of $\mathrm{supp}_{\mathbf{E}}\mu$. This is in contrast with μ being supported on a circuit. □

Remark 6.9. It follows from remark 4.14 and proposition 6.8, that the rotation vector of an irreducible occupation measure μ must have a special form:

$$\lambda \sum_{e \in \mathbf{E}^+} \tau_e \, e, \tag{28}$$

where $\lambda > 0$, $\tau_e \in \{0, \pm 1\}$, and $\mathbf{E}^+$ denotes an orientation of the graph. This fact does not hold in the classical setting and it is very peculiar of the theory on graphs. This depends on the fact that irreducibile measures are supported on circuits, hence for each vertex in the support of the circuit, there is only one edge in the circuit that has it as its origin and only one that has it as its terminal point. In order to be a cycle, then, all weights on the edges must be the same (the sign only specifies the orientation of the edge, with respect to the chosen orientation of the graph).

Theorem 6.10. *For any $c \in H^1(\Gamma, \mathbb{R})$, the set of Mather measures $\mathbb{M}_c$ is the convex hull of the irreducible Mather measures with cohomology c, which are finitely many.*

Proof. We know from proposition 5.7 that $\mathbb{M}_c$ is a convex set. We claim that $\mu \in \mathbb{M}_c$ is irreducible if and only if it is an extremal point of $\mathbb{M}_c$. It is trivial that if it is irreducible then it is extremal in $\mathbb{M}_c$. Conversely, let μ be extremal in $\mathbb{M}_c$, and assume that there exist μ_1, μ_2 in $\mathbb{M}$, $\lambda \in (0,1)$ with

$$\mu = (1-\lambda)\,\mu_1 + \lambda\,\mu_2.$$

If $\omega \in \mathfrak{C}^1(\Gamma, \mathbb{R})$ is of cohomology c, we have

$$-\alpha(c) = \int \mathcal{L}^\omega \, \mathrm{d}\mu = (1-\lambda) \int \mathcal{L}^\omega \, \mathrm{d}\mu_1 + \lambda \int \mathcal{L}^\omega \, \mathrm{d}\mu_2$$

which implies, by the minimality property of $\alpha(\cdot)$ that both μ_1 and μ_2 are Mather measures of cohomology c, which is impossible. This proves the claim.

Let $\mu \in \mathbb{M}_c$ then by theorem 6.6 it is convex combination of occupation measures supported on parametrized circuits. Arguing as in the first part of the proof, we see that all the measures forming the convex combination are in $\mathbb{M}_c$, and consequently by proposition 6.8 they are irreducible Mather measures in $\mathbb{M}_c$. This shows that $\mathbb{M}_c$ is the convex hull of its extremal points. These extremal measures are finitely many since—by the graph property in corollary 6.4—a circuit identifies the Mather measures supported on it, if any, and the set of circuits in Γ is finite. □

As shown in the previous result, any $\mathbb{M}_c$, for $c \in H^1(\Gamma, \mathbb{R})$, contains some irreducible measure. The situation is rather different for the sets $\mathbb{M}^h$. In fact, we know remark 6.9 that if $\mathbb{M}^h$ contains irreducible Mather measures, then h must be as in (28); hence, not all $\mathbb{M}^h$ do contain them. We can get some information on which $\mathbb{M}^h$'s contain irreducible Mather measures by looking at the extremal points of the epigraph of β. We recall that the epigraph of β is given by

$$\mathrm{epi}(\beta) := \{(h,t) \in H_1(\Gamma, \mathbb{R}) \times \mathbb{R} : \; t \geqslant \beta(h)\}.$$

As in the classical ergodic theory, we have:

Proposition 6.11. *Let $h \in H_1(\Gamma, \mathbb{R})$. If $(h, \beta(h))$ is an extremal point of epi(β), then there exist irreducible Mather measures of rotation vector h.*

Proof. Let μ be a Mather measure with rotation vector h; then, according to theorem 6.6

$$\mu = \sum_{i=1}^{M} \lambda_i \, \mu_i$$

with $\lambda > 0$, $\sum_i \lambda_i = 1$ and μ_i occupation measures supported on parametrized circuits. Let us define

$$h_i = \rho(\mu_i) \qquad \text{for any } i = 1, \ldots, M.$$

We have

$$\beta\left(\sum_{i=1}^{M} h_i\right) = \beta(h) = \int \mathcal{L}\,\mathrm{d}\mu$$
$$= \sum_{i=1}^{M} \lambda_i \int \mathcal{L}\,\mathrm{d}\mu_i \geqslant \sum_{i=1}^{M} \lambda_i \beta(h_i).$$

Due to the convex character of β, we see that equality must prevail in the above sequence of inequalities, so that all the μ_i's must be Mather measures. In addition, thanks to proposition 6.8, they are irreducible Mather measures. We in addition have that

$$(h, \beta(h)) = \sum_{i=1}^{M} \lambda_i (h_i, \beta(h_i)).$$

Since $(h, \beta(h))$ is an extremal point of $\mathrm{epi}(\beta)$, we must necessarily have $h_i = h$ for any i. Hence, all the μ_i's are irreducible Mather measures with rotation vector h. □

7. Weak KAM facts

We pause the exposition of Aubry-Mather theory on graphs, to recall some basic results of weak KAM theory that we will use in the following section. Note that coercivity and convexity of the Hamiltonian are sufficient for these results to hold true. All the material is taken from [29], which contains a comprehensive treatment of the topic.

We consider a 1–cochain ω with cohomology class c, and the family of discrete Hamilton–Jacobi equations on Γ

$$\max_{-e \in \mathbf{E}_x} \mathcal{H}^{\omega}(e, \langle \mathrm{d}u, e\rangle) = a \qquad \text{for } x \in \mathbf{V}, a \in \mathbb{R} \tag{HJ$^{\omega}_{a}$}$$

which can be equivalently written as

$$u(x) = \min_{-e \in \mathbf{E}_x} \left(u(\mathrm{o}(e)) + \sigma^{\omega}(e,a)\right).$$

A function $u : \mathbf{V} \to \mathbb{R}$ is called *solution* if equality in (HJ$^{\omega}_{a}$) holds for every vertex x. If instead the left-hand side is less than or equal to a, we say that u is a *subsolution* of (HJ$^{\omega}_{a}$).

We set

$$a_0 := \max_{e \in \mathbf{E}} a_e.$$

Remark 7.1. If Γ contains loops, the definition of a_0 is more involved and it could happen that $a_0 > \max_{e \in \mathbf{E}} a_e$; we refer to [29, formula (7)].

Remark 7.2. It is clear that equation (HJ$^{\omega}_{a}$) does not even make sense if $a < a_0$, because in this case the a–sublevels of $\mathcal{H}(e, \cdot)$ are empty for some edge e.

Given a path $\xi = (e_i)_{i=1}^{M}$ in Γ, we define for $a \geqslant a_0$ (see (10))

$$\sigma^{\omega}(\xi, a) := \sum_{i=1}^{M} \sigma^{\omega}(e_i, a).$$

Note that this definition only depends on the concatenated edges making up ξ, no parametrization is involved. We sometimes refer to $\sigma^{\omega}(\xi, a)$ as the *intrinsic length* of the path ξ related to the Hamiltonian $\mathcal{H}^{\omega}$ and the level a.

Proposition 7.3.

(i) *Equation* (HJ^{ω}_{a}) *admits subsolutions if and only if*

$$\sigma^{\omega}(\xi,a)\geqslant 0 \qquad \text{for any closed path } \xi.$$

(ii) *A function* $u:\mathbf{V}\to\mathbb{R}$ *is a subsolution of* (HJ^{ω}_{a}) *if and only if*

$$u(x)-u(y)\leqslant\sigma^{\omega}(\xi,a) \qquad \text{for any path } \xi \text{ with } \mathrm{o}(\xi)=y,\mathrm{t}(\xi)=x.$$

(iii) *There is one and only one value of a, called critical value of* $\mathcal{H}^{\omega}$, *for which the corresponding equation has solutions on the whole* Γ. *It is given by*

$$\min\{a\in\mathbb{R}:\ (HJ^{\omega}_{a})\text{admits subsolutions}\}. \tag{29}$$

For a proof of these claims see [29, propositions 6.5, 6.8 and theorem 6.16]

Clearly the Hamiltonian $\mathcal{H}^{\omega}$ is not invariant by change of representative in the class c, however its critical value does not depend on the chosen representative, but only on the cohomology class c. If, in fact, we replace ω by $\omega'=\omega+\mathrm{d}w$, for some $w\in\mathfrak{C}^0(\Gamma,\mathbb{R})$, then, given any (sub)solution u to the equation associated to $\mathcal{H}^{\omega}$, the function $u-w$ will be a (sub)solution to the equation associated to $\mathcal{H}^{\omega'}$.

We can therefore define a function

$$\widetilde{\alpha}:H^1(\Gamma,\mathbb{R})\to\mathbb{R}$$

associating to any cohomology class the critical value of $\mathcal{H}^{\omega}$, as defined in (29) (it only depends on the cohomology class of ω). We call *critical* the equation

$$\max_{e\in-\mathbf{E}_x}\mathcal{H}^{\omega}(e,\langle\mathrm{d}u,e\rangle)=\widetilde{\alpha}(c)$$

and qualify as critical its (sub)solutions. According to remarks 3.3 and 7.2

$$\widetilde{\alpha}(c)\geqslant a_0 \qquad \text{for any } c\in H^1(\Gamma,\mathbb{R}).$$

Proposition 7.4. *Given* $c\in H^1(\Gamma,\mathbb{R})$ *and* ω *of cohomology class c, the critical value* $\widetilde{\alpha}(c)$ *is characterized by the following properties:*

(i) $\sigma^{\omega}(\xi,\widetilde{\alpha}(c))\geqslant 0$ *for all cycles* ξ *in* Γ;
(ii) *there exists a cycle* ζ *with* $\sigma^{\omega}(\zeta,\widetilde{\alpha}(c))=0$.

For a proof of these claims see [29, lemma 6.7, corollary 6.9, proposition 6.15 and theorem 6.16].

We define the Aubry sets as follows:

$$\mathcal{A}_c:=\{e\in\mathbf{E}\mid \text{ belonging to some cycle with } \sigma^{\omega}(\xi,\widetilde{\alpha}(c))=0\}. \tag{30}$$

Remark 7.5. Given an arbitrary path ξ, the intrinsic length $\sigma^{\omega}(\xi,\widetilde{\alpha}(c))$ is not invariant for the change of representative, however invariance is valid if ξ is a cycle. This is the reason why the Aubry set only depends on c and not on the representative ω.

We state in the next proposition a relevant property of the Aubry sets (see [29, lemma 7.3]).

Proposition 7.6. *Let* $c\in H^1(\Gamma,\mathbb{R})$ *and* $\omega\in\mathfrak{C}^1(\Gamma,\mathbb{R})$ *be of cohomology class c. Then, any subsolution u of* $\mathcal{H}^{\omega}=\widetilde{\alpha}(c)$ *satisfies*

$$\langle\mathrm{d}u,e\rangle=\sigma^{\omega}(e,\widetilde{\alpha}(c)) \quad \text{and} \quad \mathcal{H}^{\omega}(e,\langle\mathrm{d}u,e\rangle)=\widetilde{\alpha}(c) \qquad \text{for } e\in\mathcal{A}_c.$$

Consequently, the differentials of all such subsolutions coincide on $e\in\mathcal{A}_c$.

The value of du on the Aubry set $\mathcal{A}_c$ is clearly not invariant for change of representative in c, however the element $\frac{\partial}{\partial p}\mathcal{H}^\omega(e,\langle \mathrm{d}u,e\rangle)$, namely the element characterized by the equality

$$\frac{\partial}{\partial p}\mathcal{H}^\omega(e,\langle \mathrm{d}u,e\rangle)\,\langle \mathrm{d}u,e\rangle = \mathcal{L}^\omega(e,\frac{\partial}{\partial p}\mathcal{H}^\omega(e,\langle \mathrm{d}u,e\rangle)) + \mathcal{H}^\omega\,(e,\langle \mathrm{d}u,e\rangle) \qquad \forall\, e\in\mathcal{A}_c \quad (31)$$

possesses such an invariance, as made precise by the following result.

Lemma 7.7. *Let $\omega,\omega'\in\mathfrak{C}^1(\Gamma,\mathbb{R})$ be in the same cohomology class c, and let u, v be subsolutions to ($HJ^\omega_{\widetilde{\alpha}(c)}$) and ($HJ^{\omega'}_{\widetilde{\alpha}(c)}$), respectively; then*

$$\frac{\partial}{\partial p}\mathcal{H}^\omega(e,\langle \mathrm{d}u,e\rangle) = \frac{\partial}{\partial p}\mathcal{H}^{\omega'}(e,\langle \mathrm{d}v,e\rangle) \qquad \textit{for any } e\in\mathcal{A}_c. \quad (32)$$

Proof. We set

$$q_e := \frac{\partial}{\partial p}\mathcal{H}^\omega(e,\langle \mathrm{d}u,e\rangle) \qquad \text{for } e\in\mathcal{A}_c.$$

We have that $\omega'=\omega+\mathrm{d}w$ for some $w\in\mathfrak{C}^0(\Gamma.\mathbb{R})$, and consequently

$$\mathrm{d}v = \mathrm{d}u - \mathrm{d}w.$$

Let $e\in\mathcal{A}_c$, then keeping in mind (31) we have

$$\begin{aligned} q_e\,\langle \mathrm{d}v,e\rangle &= q_e\,\langle \mathrm{d}u,e\rangle - q_e\,\langle \mathrm{d}w,e\rangle \\ &= \mathcal{L}^\omega(e,q_e) + \mathcal{H}^\omega(e,\langle \mathrm{d}u,e\rangle) - q_e\,\langle \mathrm{d}w,e\rangle \\ &= \mathcal{L}(e,q_e) - q_e\langle\omega,e\rangle + \mathcal{H}(e,\langle \mathrm{d}u-\mathrm{d}w+\mathrm{d}w+\omega,e\rangle) - q_e\,\langle \mathrm{d}w,e\rangle \\ &= \mathcal{L}^{\omega'}(e,q_e) + \mathcal{H}^{\omega'}(e,\mathrm{d}v). \end{aligned}$$

This proves (32). □

We denote by $\mathcal{Q}_c:\mathcal{A}_c\to\mathbb{R}$ the function

$$e \longmapsto \frac{\partial}{\partial p}\mathcal{H}^\omega(e,\langle \mathrm{d}u,e\rangle). \quad (33)$$

by the monotonicity properties of $\mathcal{H}^\omega(e,\cdot)$, $\mathcal{Q}_c(e)$ is non–negative for any $e\in\mathcal{A}_c$.

8. Weak KAM and Aubry–Mather theories

In this section we put in relation weak KAM theory and Aubry-Mather theory on graphs.

8.1. Mather's α function and critical value

Theorem 8.1. *Given $c\in H^1(\Gamma,\mathbb{R})$ and $\omega\in\mathfrak{C}^1(\Gamma,\mathbb{R})$ of cohomology class c, we have:*

- ***(i)*** *$\widetilde{\alpha}(c)$ and $\alpha(c)$ coincide, i.e. the critical value of $\mathcal{H}^\omega$ and the minimal action of Mather measures of cohomology class c are the same;*
- ***(ii)*** *if an irreducible measure belongs to $\mathbb{M}_c$, then it is supported on a circuit ζ such that $\sigma^\omega(\zeta,\alpha(c))=0$;*
- ***(iii)*** *if $\zeta=(e_i)_{i=1}^N$ is a circuit such that $\sigma^\omega(\zeta,\alpha(c))=0$ and $\mathcal{Q}_c(e_i)\neq 0$ for all $i=1,\dots,N$, then there exists an irreducible Mather measure supported on a parametrization of ζ.*

We remark that Item **(iii)** in proposition 8.1 might not hold if $\mathcal{Q}_c$ vanishes on some of the edges forming the circuit ζ of vanishing intrinsic length; see also remark 8.4.

Proof. We denote by u a subsolution to $(\mathrm{HJ}^{\omega}_{\widetilde{\alpha}(c)})$. Taking into account the definition of Lagrangian, we get for any closed probability measure μ

$$\int \mathcal{L}^{\omega}(e,q)\,\mathrm{d}\mu \geqslant \int \left[q\,\langle \mathrm{d}u, e\rangle - \mathcal{H}^{\omega}(e,\langle \mathrm{d}u, e\rangle)\right]\mathrm{d}\mu = -\widetilde{\alpha}(c),$$

which shows that

$$-\alpha(c) \geqslant -\widetilde{\alpha}(c). \tag{34}$$

Let $\xi = (e_i)_{i=1}^M$ be a circuit with

$$\sigma^{\omega}(\xi,\widetilde{\alpha}(c)) = \sum_i \sigma^{\omega}(e_i,\widetilde{\alpha}(c)) = 0$$

so that ξ is contained in $\mathcal{A}_c$. We have by proposition 7.6 and (31) that

$$\widetilde{\alpha}(c) = \mathcal{H}^{\omega}(e_i,\langle \mathrm{d}u, e_i\rangle) = \sigma^{\omega}(e_i,\widetilde{\alpha}(c))\,\mathcal{Q}_c(e_i) - \mathcal{L}^{\omega}(e_i,\mathcal{Q}_c(e_i)).$$

We first assume that $\mathcal{Q}_c(e_i) \neq 0$ for every i, then we get

$$\sigma^{\omega}(e_i,\widetilde{\alpha}(c)) = \frac{1}{\mathcal{Q}_c(e_i)}\left(\widetilde{\alpha}(c) + \mathcal{L}^{\omega}(e_i,\mathcal{Q}_c(e_i))\right).$$

By summing over i, we further obtain

$$0 = \sum_{i=1}^{M} \frac{1}{\mathcal{Q}_c(e_i)}\,\mathcal{L}^{\omega}(e_i,\mathcal{Q}_c(e_i)) + \left(\sum_{i=1}^{M} \frac{1}{\mathcal{Q}_c(e_i)}\right)\widetilde{\alpha}(c). \tag{35}$$

We denote by μ_ξ the occupation measure associated with the parametrized circuit $(e_i,\mathcal{Q}_c(e_i),1/\mathcal{Q}_c(e_i))_{i=1}^M$, and deduce from (35)

$$\int \mathcal{L}^{\omega}\,\mathrm{d}\mu_\xi = -\widetilde{\alpha}(c)$$

which together with (34) proves the item **(i)**, in the case $\mathcal{Q}_c(e_i) \neq 0$ for every i; in particular, this also proves **(iii)**.

If some $\mathcal{Q}_c(e_i)$ vanishes, then according to proposition 4.3, ξ is an equilibrium circuit based on some edge e, namely $\xi = ((e,0,T),(-e,0,S))$ for some $T,S > 0$. In this case we have

$$\widetilde{\alpha}(c) = a_0 = a_e$$

and

$$\mathcal{L}^{\omega}(e,0) = \mathcal{L}^{\omega}(-e,0) = -a_e = -\widetilde{\alpha}(c).$$

The occupation measure related to ξ is $\delta(e,0)$, and we get

$$\int \mathcal{L}^{\omega}\,\mathrm{d}\delta(e,0) = \mathcal{L}^{\omega}(e,0) = -\widetilde{\alpha}(c).$$

This ends the proof of item **(i)**. Let $\mu \in \mathbb{M}_c$ be an irreducible Mather measure. Then, we distinguish two cases (see proposition 6.8):

- μ is the occupation measure supported on a parametrized cycle $(e_i, q_i, 1/q_i)_{i=1}^M$, with $q_i \neq 0$ for all $i = 1, \ldots, M$. Denoting by $T := \sum_{i=1}^M \frac{1}{q_i}$ and $\zeta := (e_i)_{i=1}^M$, we get:

$$\begin{aligned} -\alpha(c) &= \int \mathcal{L}^\omega \, \mathrm{d}\mu = \frac{1}{T} \sum_{i=1}^{M} \frac{1}{q_i} \mathcal{L}^\omega(e_i, q_i) \\ &\geqslant \frac{1}{T} \left[\sigma^\omega(\zeta, \alpha(c)) - T\alpha(c)\right] \geqslant -\alpha(c), \end{aligned} \tag{36}$$

  which implies $\sigma^\omega(\zeta, \alpha(c)) = 0$.
- Otherwise, $\mu = \delta(e, 0)$, for some $e \in \mathbf{E}$; in this case we must have $\alpha(c) = a_e$ and

$$\sigma^\omega(e, \alpha(c)) + \sigma^\omega(-e, \alpha(c)) = 0,$$

  hence the thesis follows with $\zeta = (e, -e)$.

This concludes the proof of **(ii)**.

□

We deduce:

Corollary 8.2. *Let $c \in H^1(\Gamma, \mathbb{R})$, for any $(e, q) \in \widetilde{M}_c$ we have*

$$\mathcal{L}^\omega(e, q) = \sigma^\omega(e, \alpha(c))\, q - \alpha(c).$$

Recalling the definition of the Aubry set $\mathcal{A}_c \subset \mathbf{E}$, we further derive:

Corollary 8.3. *Given $c \in H^1(\Gamma, \mathbb{R})$, we have*

$$\pi_{\mathbf{E}}\left(\widetilde{\mathcal{M}}_c\right) =: \mathcal{M}_c \subseteq \mathcal{A}_c.$$

In particular, equality holds if c is such that $\alpha(c) > \min \alpha$.

Remark 8.4. Note that in general $\mathcal{M}_c$ might be a strict subset of $\mathcal{A}_c$, however, this could happen only for c's corresponding to the minimum of α. In fact, the reason is that there could be circuits with vanishing intrinsic length not admitting a suitable admissible parametrization, so that we do not find an occupation measure supported on it; this phenomenon is related to the presence of minimizing measures of the form $\delta(e, 0)$ for some $e \in \mathbf{E}$.

An example is given by a graph with two vertices, say x and y, and two edges e, f connecting them. We assume that $e \neq -f$ and that $\mathrm{o}(e) = \mathrm{t}(f) = x$ and $\mathrm{t}(e) = \mathrm{o}(f) = y$. We consider the Hamiltonian defined as follows:

$$\mathcal{H}(e,p) = \mathcal{H}(-e,p) = p^2,\ \mathcal{H}(f,p) = (p+1)^2 - 1,\ \mathcal{H}(-f,p) = (-p+1)^2 - 1.$$

It is easy to check that 0 is the critical value and the vanishing function is a solution of the corresponding critical equation. We moreover have

$$\sigma(e,0) = \sigma(-e,0) = 0,\ \sigma(f,0) = 0,\ \sigma(-f,0) = 2.$$

We therefore see that $(e, -e)$ is an equilibrium circuit so that $\delta(e, 0)$ is a Mather measure and e, $-e$ belong to the Mather set. We also have that the circuit (e, f) has vanishing intrinsic length, so that f belongs to the Aubry set, however, according to the definition of parametrized path, (e, f) does not admit any admissible parametrization with vanishing speed on e, and f does not belong to the Mather set.

Notice that in this example the sets of vertices corresponding to edges in the Mather and Aubry sets coincide. However, this is not necessarily the case. A counterexample could be given by a graph consisting of three vertices, and three distinct edges e, f, g forming a cycle (in this order); the Hamiltonians on $\pm e$ and $\pm f$ are as above, and the Hamiltonians on $\pm g$ are as the ones on $\pm f$. Proceeding as above, one can check that the Mather set equals $\{e, -e\}$, while the Aubry set is $\{e, f, g\}$. Hence, the vertex $\mathrm{t}(f) = \mathrm{o}(g)$ belongs to the Aubry set, but not to the Mather set.

Next theorem refines the information provided in corollary 6.4 and remark 6.5.

Theorem 8.5. *Given* $c \in H^1(\Gamma, \mathbb{R})$,

$$\widetilde{\mathcal{M}}_c = \{(e, \mathcal{Q}_c(e)) \mid e \in \mathcal{M}_c\}.$$

Proof. Let ω be of cohomology c. We know from proposition 7.6 that the differentials of all subsolutions u to $(\mathrm{HJ}^{\omega}_{\alpha(c)})$ coincide on $\mathcal{A}_c$ and satisfy

$$\langle \mathrm{d}u, e\rangle = \sigma^{\omega}(e, \alpha(c)), \quad \mathcal{H}^{\omega}(e, \langle \mathrm{d}u, e\rangle) = \alpha(c). \tag{37}$$

Let μ be an irreducible occupation measure in $\mathbb{M}_c$, and assume that it corresponds to a parametrized circuit $\xi = (e_i, q_i, T_i)_{i=1}^{M}$. We derive from corollary 8.2 that

$$\mathcal{L}^{\omega}(e_i, q_i) = \sigma^{\omega}(e_i, \alpha(c))\, q_i - \alpha(c) \qquad \text{for } i = 1, \ldots, M.$$

This implies by (37)

$$\mathcal{L}^{\omega}(e_i, q_i) + \mathcal{H}^{\omega}(e, \langle \mathrm{d}u, e\rangle) = \langle \mathrm{d}u, e\rangle\, q_i,$$

which yields $q_i = \mathcal{Q}_c(e_i)$, for $i = 1, \ldots, M$, in view of (31). □

8.2. *Minimizers of Mather's α function*

Remark 8.6. If Γ contains loops, the material in this section must be adapted to include the case in which $a_0 > \max_{e \in \mathbf{E}} a_e$.

Proposition 8.7. *The minimum of the function α is equal to a_0.*

Proof. The function α admits minimum because of its coercive character. Assume c to be a minimizer of α and denote by $\omega \in \mathfrak{C}^1(\Gamma, \mathbb{R})$ a representative of the cohomology class c. Then there exists $\mu \in \mathbb{M}_c$ with $\rho(\mu) = 0$ in view of proposition 5.8 **(ii)**. Taking into account the definition of rotation vector, we derive that for some edge f, both f and $-f$ belong to $\mathrm{supp}_{\mathbf{E}}\mu$. This implies by proposition 6.2 **(ii)** that $\mathcal{Q}_c(f) = \mathcal{Q}_c(-f) = 0$ and $\alpha(c) = \min \alpha$. Since $\mathcal{Q}_c(f) = 0$, then:

$$\frac{\partial}{\partial p}\mathcal{H}^{\omega}(f, \langle \mathrm{d}u, f\rangle) = \mathcal{Q}_c(f) = 0,$$

where u is a subsolution to $(\mathrm{HJ}^{\omega}_{\alpha(c)})$. We deduce that $\langle \mathrm{d}u, f\rangle$ is a minimizer of $\mathcal{H}^{\omega}(f, \cdot)$ and consequently

$$\alpha(c) = \mathcal{H}^{\omega}(f, \langle \mathrm{d}u, f\rangle) = a_f \leqslant a_0 \leqslant \min \alpha,$$

which implies that $\alpha(c) = \min \alpha = a_0$. □

Corollary 8.8. *An element $c \in H^1(\Gamma, \mathbb{R})$ is a minimizer of α if and only if the function $\mathcal{Q}_c$ vanishes at some $e \in \mathcal{M}_c$.*

Proof. The fact that if $\mathcal{Q}_c(e) = 0$ for some $e \in \mathcal{M}_c$ then c is a minimizer of α, has been proved in proposition 8.7.

Conversely, if c is a minimizer of α, then $\alpha(c) = a_f$ for some $f \in \mathbf{E}$, by proposition 8.7. This implies that $f \in \mathcal{M}_c$, moreover, if u is a subsolution to (HJ$^\omega_{a_f}$), where ω is a representative of c, we get

$$\mathcal{H}^\omega(f, \langle \mathrm{d}u, f\rangle) = a_f.$$

Taking into account that a_f is the minimum of $\mathcal{H}^\omega(f, \cdot)$, we finally have

$$\mathcal{Q}_c(f) = \frac{\partial}{\partial p}\mathcal{H}^\omega(f, \langle \mathrm{d}u, f\rangle) = 0.$$

□

Data availability statement

No new data were created or analysed in this study.

Acknowledgments

The second author acknowledges the support of the University of Rome Tor Vergata's *Beyond Borders* grant '*The Hamilton–Jacobi equation: at the crossroads of Analysis, Dynamics and Geometry*', the Italian Ministry of Education and Research (MIUR)'s grants: PRIN Projects 2017S35EHN '*Regular and stochastic behaviour in dynamical systems*' and 2022FPZEES '*Stability in Hamiltonian dynamics and beyond*', as well as the Department of Excellence grant MatMod@TOV awarded to the Department of Mathematics of University of Rome Tor Vergata.

Finally, both authors with to express their gratitude to the Mathematical Sciences Research Institute in Berkeley (USA) for its kind hospitality in Fall 2018 during the trimester program '*Hamiltonian systems, from topology to applications through analysis*', where part of this project was carried out.

Appendix A. From networks to graphs

In this appendix, we describe how it is possible to develop Aubry-Mather theory on networks, by means of the discrete theory that we have developed on graphs.

Let us start by recalling the definition of network, as given in [29]. We consider a finite collection $\mathcal{E}$ of regular simple oriented curves in $\mathbb{R}^N$ parametrized over $[0,1]$. If $\gamma \in \mathcal{E}$, we denote by $-\gamma \in \mathcal{E}$ the curve

$$-\gamma(s) = \gamma(1-s) \qquad \text{for } s \in [0,1],$$

with the same support of γ and opposite orientation. We further assume

$$\gamma((0,1)) \cap \gamma'((0,1)) = \emptyset \qquad \text{whenever } \gamma \neq \pm\gamma'. \tag{38}$$

A *network* $\mathcal{G}$ is a subset of $\mathbb{R}^N$ of the form

$$\mathcal{G} = \bigcup_{\gamma \in \mathcal{E}} \gamma([0,1]) \subset \mathbb{R}^N,$$

the curves in $\mathcal{E}$ are called *arcs* of the network.

We call *vertices* the initial and terminal points of the arcs, and denote by $\mathbf{V}$ the sets of all such vertices. We assume that the network is finite and connected, namely the number of arcs and vertices is finite and there is a finite concatenation of arcs linking any pair of vertices.

Remark A.1. This setting can be naturally extended to the case in which $\mathcal{G}$ is embedded in a Riemannian manifold (M,g) (for example by means of Nash embedding theorem [24]).

We can associate to any network $\mathcal{G}$ a finite and connected abstract graph $\Gamma=(\mathbf{V},\mathbf{E})$ with the same vertices of the network and edges corresponding to the arcs. More precisely, we consider an abstract set $\mathbf{E}$ with a bijection

$$\Psi:\mathbf{E}\longrightarrow\mathcal{E}. \tag{39}$$

This induces maps $o:\mathbf{E}\longrightarrow\mathbf{V}$, $-\ :\mathbf{E}\longrightarrow\mathbf{E}$ via

$$\mathrm{o}(e)=\Psi(e)(0) \quad \text{and} \quad -e=\Psi^{-1}(-\Psi(e)),$$

satisfying the properties in the definition of the graph, see section 2.

A.1. Hamiltonians and Lagrangians on networks

A Hamiltonian on $\mathcal{G}$ is a collection of Hamiltonians

$$H_\gamma:[0,1]\times\mathbb{R}\to\mathbb{R}; \qquad (s,p)\mapsto H_\gamma(s,p)$$

labeled by the arcs. We assume the compatibility conditions

$$H_{-\gamma}(s,p)=H_\gamma(1-s,-p) \qquad \text{for any } \gamma\in\mathcal{E}. \tag{40}$$

As we will discuss with more detail hereafter, we can associate to the family H_γ a Hamiltonian $\mathcal{H}(e,\cdot)$ on Γ. Exploiting the results of [29], we see that the corresponding Hamilton–Jacobi equations

$$H_\gamma(s,(u\circ\gamma)')=a \qquad \text{on } (0,1) \text{ for } \gamma\in\mathcal{E},$$

and

$$\max_{-e\in\mathbf{E}_x}\mathcal{H}(e,\langle \mathrm{d}u,e\rangle)=a \qquad \text{for } x\in\mathbf{V}, a\in\mathbb{R}$$

are equivalent, in the sense that if $u:\mathcal{G}\to\mathbb{R}$ is a (sub)solution of the former then its trace on $\mathbf{V}$ solves the latter, and, conversely, any function $w:\mathbf{V}\to\mathbb{R}$ solution of the latter can be uniquely extended on $\mathcal{G}$ in such a way that the extended function is solution of the former equation. In addition, in [29] we developed in parallel weak KAM results for the two equations, proved that the two critical values coincide, define the corresponding Aubry sets, etc….

The aim of this appendix to determine a set of rather natural assumptions on the H_γ's such that the corresponding Hamiltonian on the graph Γ satisfies **(H1)**, **(H2)**. This will allow to take advantage of the output of this paper to provide an Aubry–Mather theory on networks.

We require the H_γ's to satisfy the following properties:

(H1′) H_γ is *continuous* in (s,p), *differentiable* in p for any fixed s, and such that the function

$$(s,p)\mapsto\frac{\partial}{\partial p}H_\gamma(s,p)$$

is continuous;

(H1$'$) H_γ is *superlinear* in p, uniformly in $[0,1]$, namely

$$\lim_{r\to+\infty} \min\left\{\frac{H_\gamma(s,p)}{p} \mid p>r,\, s\in[0,1]\right\} = +\infty; \tag{41}$$

(H3$'$) H_γ is *strictly convex* in p;
(H4$'$) the map $s \longmapsto \min_{p\in\mathbb{R}} H_\gamma(s,p)$ is constant in $[0,1]$, for any given $\gamma\in\mathcal{E}$.

We define $a_\gamma = a_{-\gamma}$ as the value of the constant function appearing in the assumption **(H4')**, in other terms the sublevel of the Hamiltonian H_γ corresponding to a_γ is a singleton for any s; we further denote by p_s^γ the minimizer of $H_\gamma(s,\cdot)$. Therefore **(H4')** reads

$$H_\gamma(s,p_s^\gamma) = a_\gamma \qquad \text{for any } s\in[0,1].$$

Remark A.2. Actually condition **(H4')** is required only for $\gamma\in\mathcal{E}$ such that $a_\gamma = \max\{a_\lambda : \lambda\in\mathcal{E}\}$. We refer to [29, remark 3.3] for an explanation of the role of this condition.

We fix $\gamma\in\mathcal{E}$, $e\in\mathbf{E}$ with $\gamma=\Psi(e)$. The procedure to pass from H_γ to $\mathcal{H}(e,\cdot)$ consists in the following three steps:

- consider, for any s, the inverse, with respect to the composition, of $H_\gamma(s,\cdot)$ in $[p_s^\gamma,+\infty)$, denoted by $\sigma_\gamma^+(s,\cdot)$;
- for any fixed $a\geqslant a_\gamma$, integrate $\sigma_\gamma^+(\cdot,a)$ in $[0,1]$ obtaining $\sigma(e,a)$, where

$$\sigma_\gamma^+(s,a) := \max\{p \mid H_\gamma(s,p) = a\}$$
$$\sigma(e,a) := \int_0^1 \sigma_\gamma^+(s,a)\,\mathrm{d}s;$$

- define

$$\mathcal{H}(e,p) := \begin{cases} \sigma^{-1}(e,p) & \text{for } p\geqslant\sigma(e,a_\gamma) \\ \sigma^{-1}(-e,-p) & \text{for } p\leqslant\sigma(e,a_\gamma) \end{cases}, \tag{42}$$

  where the inverse is with respect the composition.

It is easy to see that if H_γ is independent of s, then $H_\gamma(\cdot)$ and $\mathcal{H}(e,\cdot)$ coincide.

Proposition A.3. *If assumptions **(H1$'$)–(H4$'$)** hold, then $\mathcal{H}(e,\cdot):\mathbb{R}\to\mathbb{R}$ satisfies **(H1)–(H2)**. Moreover, $a_e = a_\gamma$ and $p_e = \sigma(e,a_\gamma)$, as defined in (5).*

We need a preliminary result.

Lemma A.4. *The function $a\longmapsto\sigma(e,a)$ from $[a_\gamma,+\infty)$ to $\mathbb{R}$ is:*

(i) *continuous and strictly increasing;*
(ii) *strictly concave with $\lim_{a\to+\infty}\frac{\sigma(e,a)}{a}=0$;*
(iii) *differentiable in $(a_\gamma,+\infty)$ with $\lim_{a\to a_\gamma}\frac{\partial}{\partial a}\sigma(e,a)=+\infty$.*

Proof. The claimed continuity and monotonicity properties in item **(i)** have been already proved in [29, lemma 5.15]. Exploiting the strict convexity assumption on H_γ, we deduce that, for any $s\in[0,1]$, $\lambda\in(0,1)$, a, b in $[a_\gamma,+\infty)$

$$\begin{aligned} H_\gamma\left(s,\sigma_\gamma^+(s,(1-\lambda)a+\lambda b)\right) &= (1-\lambda)\,a+\lambda b \\ &= (1-\lambda)\,H_\gamma(s,\sigma_\gamma^+(s,a))+\lambda H_\gamma(s,\sigma_\gamma^+(s,b)) \\ &> H_\gamma(s,(1-\lambda)\,\sigma_\gamma^+(s,a)+\lambda\sigma_\gamma^+(s,b)). \end{aligned} \tag{43}$$

Since $H_\gamma(s,\cdot)$ is increasing in the interval $(p_s,+\infty)$, the inequality in (43) yields

$$\sigma_\gamma^+(s,(1-\lambda)a+\lambda b) > (1-\lambda)\,\sigma_\gamma^+(s,a)+\lambda\,\sigma_\gamma^+(s,b).$$

By integrating the above relation over $[0,1]$, we finally get

$$\sigma(e,(1-\lambda)a+\lambda b) > (1-\lambda)\,\sigma(e,a)+\lambda\,\sigma(e,b),$$

which shows the strictly concave character of $\sigma(e,\cdot)$.

To prove the limit relation in **(ii)**, we exploit the uniform superlinearity assumption **(H2′)** on H_γ. Assume by contradiction that there is a sequence $a_n \to \infty$ and a positive M such that

$$\lim_{n\to+\infty} \frac{\sigma(e,a_n)}{a_n} > M.$$

It follows from the definition of $\sigma(e,a_n)$ that there exist, for any n, $s_n \in [0,1]$, $p_n \in \mathbb{R}$ such that

$$H_\gamma(s_n,p_n) = a_n \quad \text{and} \quad \frac{p_n}{a_n} > M.$$

Hence, we derive

$$p_n \to +\infty \quad \text{and} \quad \frac{H_\gamma(s_n,p_n)}{p_n} < \frac{1}{M},$$

which is in contrast with (41). We deduce from **(H1′)** that the inverse function $a \mapsto \sigma_\gamma^+(s,a)$ is differentiable in $(a_\gamma,+\infty)$. By differentiating under the integral sign, we further get that $a \mapsto \sigma(e,a)$ is differentiable in $(a_\gamma,+\infty)$ and

$$\frac{\partial}{\partial a}\sigma(e,a) = \int_0^1 \frac{\partial}{\partial a}\sigma_\gamma^+(s,a)\,\mathrm{d}s.$$

We denote by $\omega(\cdot)$ a uniform continuity modulus of $(s,a) \mapsto \sigma_\gamma^+(s,a)$ in $[0,1]\times[a_\gamma,a_\gamma+1]$ and of $(s,p) \mapsto \frac{\partial}{\partial p}H_\gamma(s,p)$ in K (see assumption **(H1′)**), where

$$K = \{(s,p) \mid s\in[0,1],\, p\in[p_s,+\infty),\, H_\gamma(s,p) \leqslant a_\gamma+1\}$$

is compact by the superlinearity assumption **(H2′)**. Then

$$0 \leqslant \frac{\partial}{\partial p}H_\gamma(s,p) \leqslant \omega(p-p_s) \qquad \text{for } (s,p)\in K. \tag{44}$$

Observe that

$$a = H_\gamma(s,\sigma_a^+(s,a)) \qquad \Longrightarrow \qquad 1 = \frac{\partial}{\partial p}H_\gamma(s,\sigma_a^+(s,a))\,\frac{\partial}{\partial a}\sigma_a^+(s,a).$$

This and (44) imply that for $a \in (a_e,a_e+1)$ we have

$$\frac{\partial}{\partial a}\sigma(e,a) = \int_0^1 \frac{1}{\frac{\partial}{\partial p}H_\gamma(s,\sigma_\gamma^+(s,a))}\,\mathrm{d}s$$
$$\geqslant \int_0^1 \frac{1}{\omega(\sigma_\gamma^+(s,a)-p_s)}\,\mathrm{d}s \geqslant \frac{1}{\omega\circ\omega(a-a_\gamma)}.$$

From this we derive item **(iii)**, and conclude the proof. □

Proof of proposition A.3. We derive from (43) and lemma A.4 that $\mathcal{H}(e,\cdot)$ is continuous in $\mathbb{R}$ and differentiable in $\mathbb{R}\setminus\{\sigma(e,a_\gamma)\}$. Taking into account that

$$\frac{\partial}{\partial p}\mathcal{H}(e,p) = \frac{1}{\frac{\partial}{\partial a}\sigma(e,\sigma^{-1}(e,p))} \qquad \text{for } p > \sigma(e,a_\gamma)$$
$$\frac{\partial}{\partial p}\mathcal{H}(e,p) = -\frac{1}{\frac{\partial}{\partial a}\sigma(-e,\sigma^{-1}(-e,p))} \qquad \text{for } p < \sigma(e,a_\gamma)$$

we derive from lemma A.4 **(iii)** that

$$\lim_{p\to\sigma(e,a_\gamma)} \frac{\partial}{\partial p}\mathcal{H}(e,p) = 0,$$

which implies that $\mathcal{H}(e,\cdot)$ is differentiable in $\sigma(e,a_\gamma)$ with vanishing derivative. Strict convexity is straightforward from the previous discussion. Let us prove **(H2)**, namely that $\lim_{p\to\pm\infty}\frac{\mathcal{H}(e,p)}{|p|} = +\infty$.

Recalling (43) and using **(ii)** in lemma A.4:

$$\lim_{p\to+\infty}\frac{\mathcal{H}(e,p)}{p} = \lim_{p\to+\infty}\frac{\sigma^{-1}(e,p)}{p} = \lim_{a\to+\infty}\frac{a}{\sigma(e,a)} = +\infty. \tag{45}$$

Similarly for $p\to-\infty$, considering $\sigma(-e,a)$.

Easily follows that $a_e = a_\gamma$ and $p_e = \sigma(e,a_\gamma)$ (see (5)).

□

For every $\gamma\in\mathcal{E}$, consider the Lagrangian associated to H_γ, namely its *convex conjugate* $L_\gamma : [0,1]\times\mathbb{R}\longrightarrow\mathbb{R}$ defined as

$$L_\gamma(s,q) := \sup_{p\in\mathbb{R}}(pq - H_\gamma(s,p)), \tag{46}$$

where equality is achieved for p such that $\frac{\partial H_\gamma}{\partial p}(s,p) = q$.

Since H_γ satisfies **(H1′)–(H3′)**, then it follows (see for example [26]) that L_γ is *continuous* in (s,q), *differentiable*, *superlinear* and *strictly convex* in q.

Using (40) we see that the L_γ's satisfy the following compatibility condition:

$$L_{-\gamma}(s,q) = L_\gamma(1-s,-q) \qquad \forall\, s\in[0,1],\ q\in\mathbb{R}.$$

A.2. How to develop Aubry-Mather theory on networks

In this section we look, from the point of view of networks, at some of the notions that we have introduced in the previous sections. This will help clarify and validate the setting that we have proposed, and it will outline the ideas and the tools that are needed in order to transfer the previous construction to the network setting.

In this section we assume the Hamiltonian $\{H_\gamma\}_{\gamma\in\mathcal{E}}$ to be *Tonelli*, namely, besides **(H1′)**–**(H4′)**, we further require that for any $\gamma \in \mathcal{E}$

(H5′) $L_\gamma(s,q)$ is of class C^2 in (s,q) and $\frac{\partial^2}{\partial q^2}L_\gamma$ is positive definite.

We consider the network $\mathcal{G}$ and its corresponding abstract graph Γ. We fix an arc γ and an edge e with $\Psi(e) = \gamma$.

Given a parametrization (q_e, T_e) of the edge $e \in \mathbf{E}$, we provide an interpretation of it on the corresponding arc γ. We first assume $q > 0$, so that, according to the definition of parametrized path, $T_e = \frac{1}{q_e}$. Then, due to the strict convexity of $\mathcal{L}(e,\cdot)$, there exists a unique $p_{q_e} \geqslant p_e$ such that

$$\mathcal{L}(e,q) = p_{q_e} q_e - \mathcal{H}(e,p_{q_e}) = q\,\sigma(e,a_{q_e}) - a_{q_e}, \tag{47}$$

where $a_{q_e} > a_e$ is such that $p_{q_e} = \sigma(e,a_{q_e})$ (it is uniquely defined because of the continuity and strict monotonicity of $\sigma(e,\cdot)$ stated in lemma A.4). This also implies the relation

$$q_e = \frac{\partial}{\partial p}\mathcal{H}(e,p_{q_e}) = \frac{\partial}{\partial p}\mathcal{H}(e,\sigma(e,a_{q_e})).$$

We consider the solution to $H_\gamma(s,w'(s)) = a_{q_e}$ in $(0,1)$ given by

$$w(s) = \int_0^s \sigma_\gamma^+(t,a_{q_e})\,\mathrm{d}t,$$

and the orbit of the Hamiltonian flow related to H_γ in $[0,1]\times\mathbb{R}$ with initial datum $(0,\sigma_\gamma^+(0,a_{q_e})) = (0,w'(0))$, contained in the energy level a_{q_e}. This orbit has as first component the curve ξ_{q_e} with $\xi_{q_e}(0) = 0$ and

$$\dot{\xi}_{q_e} = \frac{\partial}{\partial p}H_\gamma(\xi_{q_e}(t),w'(\xi_{q_e}(t))),$$

while the second component is given by $w'(\xi_{q_e}(t))$. We have in fact

$$\begin{aligned} 0 &= \frac{d}{\mathrm{d}t}H_\gamma(\xi_{q_e}(t),w'(\xi_{q_e}(t))) \\ &= \frac{\partial}{\partial s}H_\gamma(\xi_{q_e}(t),w'(\xi_{q_e}(t)))\,\dot{\xi}_{q_e}(t) + \dot{\xi}_{q_e}(t)\,\frac{d}{\mathrm{d}t}w'(\xi_{q_e}(t)), \end{aligned}$$

and accordingly

$$\frac{d}{\mathrm{d}t}w'(\xi_{q_e}(t)) = -\frac{\partial}{\partial s}H_\gamma(\xi_{q_e}(t),w'(\xi_{q_e}(t))).$$

The orbit is defined in $[0,T_{q_e}]$, where T_{q_e} is the time in which ξ_{q_e} reaches the boundary point $s = 1$.

Proposition A.5. *Let $q_e > 0$ and let ξ_{q_e} and T_{q_e} be defined as above. Then:*

(i) *The time T_{q_e} is equal to $\frac{1}{q_e}$;*
(ii) *q_e is the average speed of ξ_{q_e} in the time interval $[0,T_{q_e}]$;*
(iii) *$\mathcal{L}(e,q_e) = \frac{1}{T_{q_e}}\int_0^{T_{q_e}} L_\gamma(\xi_{q_e},\dot{\xi}_{q_e})\,\mathrm{d}t$;*
(iv) *$\mathcal{L}(e,q_e) = \frac{1}{T_{q_e}}\min\left\{\int_0^{T_{q_e}} L_\gamma(\zeta(t),\dot{\zeta}(t))\,\mathrm{d}t\right\}$, where the minimum is taken in the family of absolutely continuous curves $\zeta : [0,T_{q_e}] \longrightarrow [0,1]$ with $\zeta(0) = 0$, $\zeta(T_{q_e}) = 1$.*

Proof. We have that $\dot{\xi}_{q_e}(t)$ and $w'(\xi_{q_e}(t))$ are conjugate in $[0,T_{q_e}]$, in the sense that

$$\dot{\xi}_{q_e}(t)\,w'(\xi_{q_e}(t)) = L_\gamma(\xi_{q_e}(t),\dot{\xi}_{q_e}(t)) + H_\gamma(\xi_{q_e}(t),w'(\xi_{q_e}(t))).$$

which implies

$$L_\gamma(\xi_{q_e}(t),\dot{\xi}_{q_e}(t)) = \dot{\xi}_{q_e}(t)\,\sigma_\gamma^+(\xi_{q_e}(t),a_{q_e}) - a_{q_e}. \tag{48}$$

In addition, it follows from the definition of L_γ that

$$L_\gamma(\xi_{q_e}(t),\dot{\xi}_{q_e}(t)) \geqslant \dot{\xi}_{q_e}(t))\,\sigma_\gamma^+(\xi_{q_e}(t),b) - b \quad \text{for any } b \geqslant a_e. \tag{49}$$

By integrating (48), (49) over $[0,T_{q_e}]$ we further get

$$\int_0^{T_{q_e}} L_\gamma(\xi_{q_e}(t),\dot{\xi}_{q_e}(t)) = \sigma(e,a_{q_e}) - T_{q_e}\,a_{q_e}$$

$$\int_0^{T_{q_e}} L_\gamma(\xi_{q_e}(t),\dot{\xi}_{q_e}(t)) \geqslant \sigma(e,b) - T_{q_e}\,b \quad \text{for any } b \geqslant a_e. \tag{50}$$

Taking into account (9), we derive

$$\mathcal{L}(e,1/T_{q_e}) = \frac{1}{T_{q_e}}\,\sigma(e,a_{q_e}) - a_{q_e}. \tag{51}$$

This implies by (47) and the strict convexity of $\mathcal{L}(e,\cdot)$

$$T_{q_e} = \frac{1}{q_e} \qquad \text{and} \qquad q_e = \frac{1}{T_{q_e}}\int_0^{T_{q_e}} \dot{\xi}_{q_e}(t)\,\mathrm{d}t,$$

showing items **(i)** and **(ii)**. By combining (50) and (51), we get **(iii)**.

Finally, to obtain item **(iv)**, it is enough to observe that for any absolutely continuous curve ζ in $[0,1]$ with $\zeta(0)=0$ and $\zeta(T_{q_e})=1$, one has

$$\int_0^{T_{q_e}} L_\gamma(\zeta(t),\dot{\zeta}(t)) \geqslant \sigma(e,b) - T_{q_e}\,b.$$

□

Remark A.6. The equality in item **(iii)** of proposition A.5 can be interpreted by saying that the action functional on the graph computed in $\delta(e,q)$ equals the action functional on the network computed in the occupation measure corresponding to the speed curve $(\xi_{q_e}(t),\dot{\xi}_{q_e}(t))$ in $[0,T_{q_e}]$. The latter measure is obtained by pushing forward through $(\xi_{q_e}(t),\dot{\xi}_{q_e}(t))$ the 1–dimensional Lebesgue measure restricted to $[0,T_{q_e}]$ and normalize it.

In particular, item **(iv)** of proposition A.5 reads that the curve ξ_{q_e} defined on $[0,T_{q_e}]$ is action minimizing for L_γ. This sheds light on the reason why Mather measures on the graph consist of convex combinations of Dirac deltas (see theorem 6.1), in analogy with what happens in the classical theory, in which Mather measures are supported on action-minimizing curves (see [22, 30]).

Remark A.7. In the case where $e \in \mathcal{M}_0$ and $q_e = \mathcal{Q}_0(e) > 0$—we have chosen the cohomology class 0 just for simplicity –, the above construction acquires a global significance, in the sense that $\sigma_\gamma^+(s,\alpha(0))$ is not just the derivative of a local (in $(0,1)$) solution of $H_\gamma = \alpha(0)$, but we also have that

$$\sigma_\gamma^+(s,\alpha(0)) = \frac{d}{ds} u\circ\gamma(s)$$

for any critical subsolution u of the Hamilton–Jacobi equation on the network, see [29].

Remark A.8. To discuss the case when the speed q_e vanishes for some $e\in\mathbf{E}$, the equilibrium circuit $(e,-e)$ with the parametrization $((e,0,T_1),(-e,0,T_2))$, with T_1, T_2 positive constants. We set $\gamma=\Psi(e)$ and consequently $-\gamma=\Psi(-e)$. We have

$$\mathcal{L}(e,0)=\mathcal{L}(-e,0)=-\mathcal{H}(e,p_e)=\mathcal{H}(-e,p_{-e})=-a_e=a_{-e}.$$

In addition we have by assumption **(H4′)**

$$L_\gamma(s,0)=L_{-\gamma}(s,0)=a_e=a_{-e} \qquad \text{for every } s\in[0,1]$$

so that

$$0=\frac{\partial}{\partial s}L_\gamma(s,0)=-\frac{\partial}{\partial s}H_\gamma(s,\sigma_\gamma^+(s,a_e))$$
$$0=\frac{\partial}{\partial s}L_{-\gamma}(s,0)=-\frac{\partial}{\partial s}H_{-\gamma}(s,\sigma_{-\gamma}^+(s,a_{-e})).$$

This implies that all the points $(s,\sigma_{a_e}^+(\gamma,s))$, $(s,\sigma_{a_{-e}}^+(-\gamma,s))$ are equilibria of the Hamiltonian flows related to H_γ, $H_{-\gamma}$, respectively. We can put in relation the measures $\delta(e,0)=\delta(-e,0)$ with the Dirac measures concentrated at all points of the arcs γ,$-\gamma$, which—in analogy with what we did in the graph—can be identified.

Appendix B. Proof of theorem 4.15

We need a preliminary result, see [3, proposition 42]. We refer to (12) for a precise definition of having linear growth.

Lemma B.1. *Let $\mathbb{K}$ be a closed convex subset of $\mathbb{P}$, we set*

$$C^+=\left\{F:T\Gamma\to\mathbb{R} \text{ continuous with linear growth} \mid \int F\,d\mu\geqslant 0\ \forall\,\mu\in\mathbb{K}\right\}.$$

Then:

$$\mathbb{K}=\left\{\nu\in\mathbb{P}\mid \int F\,d\nu\geqslant 0\ \forall\,F\in C^+\right\}.$$

The proof of this lemma is based on a separation result in Wasserstein spaces that we take from [19]. We state it below with slight changes to adapt it to our notation and setting.

Lemma B.2 ([19, theorem 2.9]). *Let $\mathbb{K}$ be a closed convex subset of $\mathbb{P}$, and $\nu\notin\mathbb{K}$. Then, there exists $F:T\Gamma\to\mathbb{R}$ with linear growth such that*

$$\int F\,d\mu>\int F\,d\nu \qquad \textit{for any } \mu\in\mathbb{K}.$$

Proof of lemma B.1. Given $\nu\notin\mathbb{K}$, we fix $\mu_0\in\mathbb{K}$ and define

$$\nu_\lambda=(1-\lambda)\nu+\lambda\mu_0 \qquad \text{for } \lambda\in[0,1].$$

Since $\mathbb{K}$ is closed, there exists $\lambda_0 \in (0,1)$ with $\nu_{\lambda_0} \notin \mathbb{K}$. We denote by F a function satisfying the statement of lemma B.2 with respect to ν_{λ_0}; we can in addition assume, without losing generality, that

$$\int F \, d\nu_{\lambda_0} = 0. \tag{52}$$

Therefore $F \in C^+$ and

$$\int F \, d\mu > 0 \qquad \text{for any } \mu \in \mathbb{K} \tag{53}$$

It follows from the definition of ν_{λ_0}, (52), (53) that

$$\int F \, d\nu < 0.$$

Summing up, we have found that for any $\nu \notin \mathbb{K}$, there exists $F \in C^+$ whose integral with respect to ν is strictly negative. This proves the assertion. □

Lemma B.3. *The closure in $\mathbb{P}$ of the space of closed occupation measures is convex.*

The fact that the closure of the space of closed occupation measures is convex, stems from the property that a closed occupation measure stays unchanged under any finite repetition of the corresponding cycle. Therefore, we can connect a finite number of cycles through simple paths in order to make a unique cycle. We can then repeat n times the cycles leaving unaffected the connecting paths and obtain a sequence of closed occupation measures indexed by n converging, as $n \to +\infty$, to a measure which does not 'see' the connecting simple paths and is a convex combination of the occupation measures corresponding to the cycles with repetitions.

A formal argument can be found [3, lemma 30] for measures on the tangent bundle of a compact manifold. It can be adapted with minor modifications to our setting.

We can now prove the main result of this appendix.

Proof of theorem 4.15. In view of lemma B.1, it is enough to show that if a continuous function F with linear growth in $T\Gamma$ satisfies

$$\int F \, d\mu \geqslant 0 \qquad \text{for any closed occupation measure } \mu, \tag{54}$$

then it also satisfies

$$\int F \, d\nu \geqslant 0 \qquad \text{for any measure } \nu \in \mathbb{M}. \tag{55}$$

Let F satisfy (54). By integration with respect to the closed occupation measures $\delta(e,0)$, for any $e \in \mathbf{E}$, we get

$$F(e,0) \geqslant 0 \qquad \text{for any } e \in \mathbf{E}.$$

Thanks to the above inequality, we can modify F in $[0,1/n] \cup [n,+\infty) \subset \mathbb{R}^+_e$, for any $e \in \mathbf{E}$, constructing a sequence of continuous functions F_n defined on $T\Gamma$ such that

$$F_n(e,0) > 0, \quad F_n(e,\cdot) \text{ has superlinear growth at } +\infty \qquad \text{for any } e \in \mathbf{E} \tag{56}$$

and in such a way that for any n, for each $e \in \mathbf{E}$, $q \geqslant 0$

$$F_{n+1}(e,q) \leqslant F_n(e,q) \tag{57}$$

$$F_n(e,q) \geqslant F(e,q) \tag{58}$$

$$F_n(e,q) \to F(e,q) \qquad \text{as } n \to +\infty. \tag{59}$$

We define

$$G_n(e,p) := \max_{q \geqslant 0} (p\,q - F_n(e,q))\,;$$

the function $G_n(e,\cdot)$ is finite by the superlinear growth of F_n, convex and superlinear at $+\infty$; in addition $G_n(e,\cdot)$ is increasing in p and by (56)

$$\inf_{p\in\mathbb{R}} G_n(e,p) = \lim_{p\to-\infty} G_n(e,p) = -F_n(e,0) < 0.$$

Therefore, the value 0 is attained by $G_n(e,\cdot)$ and is above the infimum. We denote by φ_e^n, for any $e \in \mathbf{E}$, the unique element such that

$$G_n(e,\varphi_e^n) = 0.$$

The quantity φ_e^n must be understood as an intrinsic length of the edge e related to the Hamiltonian G_n and the value 0. We have

$$\varphi_e^n\, q \leqslant F_n(e,q) \qquad \text{for any } q \geqslant 0 \tag{60}$$

and there exists $q_e > 0$ with

$$0 = G_n(e,\varphi_e) = \varphi_e^n\, q_e - F_n(e,q_e). \tag{61}$$

We consider the discrete Hamilton–Jacobi equation on Γ

$$\max_{-e\in\mathbf{E}_x} G_n(e,\langle \mathrm{d}u,e\rangle) = 0 \qquad \text{for } x \in \mathbf{V}. \tag{62}$$

We know from proposition 7.3 **(i)** that in order (62) to have subsolutions it is necessary and sufficient that for any cycle $\xi = (e_i)_{i=1}^M$ in Γ the intrinsic length

$$\varphi^n(\xi) := \sum_{i=1}^M \varphi_{e_i}^n \geqslant 0.$$

We deduce from (61) that

$$\varphi_{e_i}^n = \frac{1}{q_i} F_n(e_i,q_i), \tag{63}$$

where $q_i := q_{e_i}$ (see (60), (61)). We consider the parametrized version of ξ given by $(e_i,q_i,1/q_i)_{i=1}^M$ and denote by μ_ξ the corresponding closed occupation measure. We have by (58) and the assumption that

$$0 \leqslant \int F_n\,\mathrm{d}\mu_\xi = \frac{1}{\sum_{i=1}^M \frac{1}{q_i}} \sum_{i=1}^M \frac{1}{q_i} F_n(e_i,q_i),$$

which finally implies, using (63), that $\varphi^n(\xi) \geqslant 0$. If $u : \mathbf{V} \to \mathbb{R}$ is a subsolution of (62), we have

$$\langle \mathrm{d}u,e\rangle \leqslant \varphi_e^n \qquad \text{for any } e \in \mathbf{E}. \tag{64}$$

Let $\nu = \sum_{e\in\mathbf{E}} \lambda_e \nu_e$ be a closed measure on $T\Gamma$, then by (60), (64)

$$\begin{aligned} 0 &= \sum_{e\in\mathbf{E}} \lambda_e \int q\,\langle \mathrm{d}u, e\rangle \mathrm{d}\nu_e \\ &\leqslant \sum_{e\in\mathbf{E}} \lambda_e \int \varphi_e^n\, q\, \mathrm{d}\nu_e \\ &\leqslant \sum_{e\in\mathbf{E}} \lambda_e \int F_n(e,q)\,\mathrm{d}\nu_e \;=\; \int F_n\,\mathrm{d}\nu. \end{aligned}$$

Taking into account (57), (59) and passing to the limit as $n \to +\infty$ in the above inequality, we obtain

$$\int F\,\mathrm{d}\nu \geqslant 0,$$

which shows that F satisfies (55). This concludes the proof. □

ORCID iDs

Antonio Siconolfi https://orcid.org/0000-0002-6111-6109
Alfonso Sorrentino https://orcid.org/0000-0002-5680-2999

References

[1] Achdou Y, Dao M, Ley O and Tchou N 2019 A class of infinite horizon mean field games on networks *Netw. Heterog. Media* **14** 537–66
[2] Aubry S and Le Daeron P Y 1983 The discrete Frenkel-Kontorova model and its extensions. I. Exact results for the ground-states *Physica* D **8** 381–422
[3] Bernard P 2008 Young measure, superposition and transport *Indiana Univ. Math. J.* **57** 247–76
[4] Bernard P and Buffoni B 2006 The Monge problem for supercritical Mañé potentials on compact manifolds *Adv. Math.* **207** 691–706
[5] Bernard P and Buffoni B 2007 Optimal mass transportation and Mather theory *J. Eur. Math. Soc.* **9** 85–121
[6] Bernard P and Buffoni B 2007 Weak KAM pairs and Monge-Kantorovich duality *Asymptotic analysis and singularities—elliptic and parabolic PDEs and related problems* (*Advanced Studies in Pure Mathematics* vol 2) (Mathematics Society Japan) pp 397–420
[7] Camilli F and Marchi C 2016 Stationary Mean Field Games systems defined on networks *SIAM J. Control Optim.* **54** 1085–103
[8] Camilli F, De Maio R and Tosin A 2017 Transport of measures on networks *Netw. Heterog. Media* **12** 191–215
[9] Candel A and de la Llave R 1998 On the Aubry-Mather theory in statistical mechanics *Commun. Math. Phys.* **192** 649–69
[10] Contreras G and Iturriaga R 1999 Global minimizers of autonomous Lagrangians *22o Colóquio Brasileiro de Matemática (22nd Brazilian Mathematics Colloquium)* (Instituto de Matemática Pura e Aplicada (IMPA)) p 148
[11] Contreras G, Iturriaga R and Siconolfi A 2015 Homogenization on arbitrary manifolds *Calc. Var. PDE* **52** 237–52
[12] Fathi A 1997 Théorème KAM faible et théorie de Mather sur les systèmes lagrangiens *C. R. Acad. Sci., Paris* I **324** 1043–6
[13] Fathi A 2014 Weak KAM theory: the connection between Aubry-Mather theory and viscosity solutions of the Hamilton-Jacobi equation *Proc. Int. Congress of Mathematicians (ICM)* (*Seoul, Korea, 2014*) vol III pp 597–62
[14] Forcadel N and Salazar W 2020 Homogenization of a discrete model for a bifurcation and application to traffic flow *J. Math. Pures Appl.* **136** 356–414

[15] Galise G, Imbert C and Monneau R 2015 A junction condition by specified homogenization and application to traffic lights *Anal. PDE* **8** 1891–929
[16] Gangbo W, Li W and Mou C 2019 Geodesics of minimal length in the set of probability measures on graphs *ESAIM Control Optim. Calc. Var.* **25** 36
[17] Guéant O 2015 Existence and uniqueness result for mean field games with congestion effect on graphs *Appl. Math. Optim.* **72** 291–303
[18] Imbert C and Monneau R 2017 Flux-limited solutions for quasi-convex Hamilton-Jacobi equations on networks *Ann. Sci. Ec. Norm. Super.* **50** 357–448
[19] Laschos V, Obermayer K, Shen Y and Stannat W 2019 A Fenchel-Moreau-Rockafellar type theorem on the Kantorovich-Wasserstein space with applications in partially observable Markov decision processes *J. Math. Anal. Appl.* **477** 1133–56
[20] Mañé R 1996 Generic properties and problems of minimizing measures of Lagrangian systems *Nonlinearity* **9** 273–310
[21] Mather J N 1982 Existence of quasiperiodic orbits for twist homeomorphisms of the annulus *Topology* **21** 457–67
[22] Mather J N 1991 Action minimizing invariant measures for positive definite Lagrangian systems *Math.* Z **207** 169–207
[23] Mazon J, Rossi J D and Toledo J 2015 Optimal mass transport on metric graphs *SIAM J. Control Optim.* **25** 1609–32
[24] Nash J F 1956 The imbedding problem for Riemannian manifolds *Ann. Math.* **63** 20–63
[25] Pozza M and Siconolfi A 2021 Discounted Hamilton-Jacobi equations on networks and asymptotic analysis *Indiana Univ. Math. J.* **70** 1103–29
[26] Rockafellar R T 1970 *Convex Analysis* (*Princeton Mathematical Series* vol 28) (Princeton University Press)
[27] Siburg K F 2004 *The Principle of Least Action in Geometry and Dynamics* (*Lecture Notes in Math* vol 1844) (Springer) p xii+128
[28] Siconolfi A 2022 Time-dependent Hamilton-Jacobi equations on networks *J. Math. Pures Appl.* **163** 702–38
[29] Siconolfi A and Sorrentino A 2018 Global results for Eikonal Hamilton-Jacobi equations on networks *Anal. PDE* **1** 171–211
[30] Sorrentino A 2015 *Action-Minimizing Methods in Hamiltonian Dynamics: An Introduction to Aubry-Mather Theory* (*Mathematical Notes* vol 50) (Princeton University Press) p xii+115
[31] Sunada T 2013 *Topological Crystallography: With a View Towards Discrete Geometric Analysis* (*Surveys and Tutorials in the Applied Mathematical Sciences* vol 6) (Springer) p xii+229
[32] Villani C 2009 *Optimal Transport: Old and New* (*Grundlehren der mathematischen Wissenschaften* vol 338) (Springer) p xxii+976