Research Article

Andrea Davini* and Maxime Zavidovique

Convergence of the solutions of discounted Hamilton–Jacobi systems

https://doi.org/10.1515/acv-2018-0037
Received June 2, 2018; revised January 16, 2019; accepted January 28, 2019

Abstract: We consider a weakly coupled system of discounted Hamilton–Jacobi equations set on a closed Riemannian manifold. We prove that the corresponding solutions converge to a specific solution of the limit system as the discount factor goes to 0. The analysis is based on a generalization of the theory of Mather minimizing measures for Hamilton–Jacobi systems and on suitable random representation formulae for the discounted solutions.

Keywords: Asymptotic behavior of solutions, Mather measures, weak KAM theory, viscosity solutions, optimal control

MSC 2010: 35B40, 37J50, 49L25

Communicated by: Frank Duzaar
Dedicated to the memory of John Mather

Introduction

In this paper, we are interested in the asymptotic behavior, as $\lambda \to 0^+$, of the solutions of the following system of weakly coupled Hamilton–Jacobi (HJ) equations:

$$\sum_{j=1}^{m} b_{ij}u_j + \lambda u_i + H_i(x, Du_i) = c \quad \text{in } M$$

for $i \in \{1, \dots, m\}$, where M is a compact, connected Riemannian manifold without boundary, c is a real number, $H_1, \dots, H_m$ are continuous functions on T^*M, convex and coercive in the gradient variable, and $B = (b_{ij})$ is an $m \times m$ irreducible and weakly diagonally dominant matrix, see Section 1.2 for the precise assumptions. The sign and degeneracy condition assumed on the coefficients of B amounts to requiring that $-B$ is the generator of a semigroup of stochastic matrices. The solution $\mathbf{u} = (u_1, \dots, u_m)^T : M \to \mathbb{R}^m$ is assumed to be continuous and to solve the above system in the viscosity sense.

It is convenient to restate the system in the following vectorial form:

$$(B + \lambda \mathrm{Id})\mathbf{u} + \mathbb{H}(x, D\mathbf{u}) = c\mathbb{1} \quad \text{in } M, \tag{1}$$

where we have used the notations $\mathbb{H}(x, D\mathbf{u}) = (H_1(x, Du_1), \dots, H_m(x, Du_m))^T$ and $\mathbb{1} = (1, \dots, 1)^T \in \mathbb{R}^m$. The conditions assumed on B imply, in particular, that $B\mathbb{1} = 0$.

***Corresponding author: Andrea Davini,** Dipartimento di Matematica, Sapienza Università di Roma, Piazzale le Aldo Moro 2, 00185 Roma, Italy, e-mail: davini@mat.uniroma1.it. https://orcid.org/0000-0001-7641-7548
Maxime Zavidovique, IMJ-PRG (projet Analyse Algébrique), UPMC, 4, place Jussieu, Case 247, 75252 Paris Cedex 5, France, e-mail: zavidovique@math.jussieu.fr

When $\lambda = 0$, there is a unique value c for which (1) admits solutions, hereafter denoted by $c(\mathbb{H})$ and termed *critical*. Furthermore, the solutions of the *critical system*

$$B\mathbf{u} + \mathbb{H}(x, D\mathbf{u}) = c(\mathbb{H})\mathbb{1} \quad \text{in } M \tag{2}$$

are not unique, not even up to addition of vectors of the form $a\mathbb{1}$, in general.

When $\lambda > 0$, on the other hand, system (1) satisfies a comparison principle, yielding the existence of a unique continuous solution $\mathbf{u}^{\lambda,c} : M \to \mathbb{R}^m$ for every fixed $c \in \mathbb{R}$. Moreover, the solutions $\{\mathbf{u}^{\lambda,c} \mid \lambda > 0\}$ are equi-Lipschitz. The peculiarity of the discounted system (1) when $c := c(\mathbb{H})$ relies on the fact that the corresponding solutions $\mathbf{u}^\lambda := \mathbf{u}^{\lambda,c(\mathbb{H})}$ are also equi-bounded. By the Ascoli–Arzelà Theorem and by the stability of the notion of viscosity solution, we infer that they uniformly converge, *along subsequences* as λ goes to 0, to viscosity solutions of the critical system (2). Since the solutions of the critical system are not unique, it is not clear at this level that the limits of the $\mathbf{u}^\lambda$ along different subsequences yield the same critical solution.

In this paper, we address this question. The main theorem we will establish is the following:

Theorem 1. *Let $\mathbf{u}^\lambda$ be the solution of system* (1) *with $c := c(\mathbb{H})$ and $\lambda > 0$. The functions $\mathbf{u}^\lambda$ uniformly converge as $\lambda \to 0^+$ to a single solution $\mathbf{u}^0$ of the critical system* (2).

We will characterize $\mathbf{u}^0$ in terms of a generalized notion of Mather minimizing measure for HJ systems.

Notice that the relationship between $\mathbf{u}^\lambda$ and $\mathbf{u}^{\lambda,c}$ when c varies is rather straightforward: it is easily verified that $\mathbf{u}^{\lambda,c} = \mathbf{u}^\lambda + \frac{c-c(\mathbb{H})}{\lambda}\mathbb{1}$. As a consequence, we derive from Theorem 1 the following fact:

Theorem 2. *Let $\mathbf{u}^{\lambda,c}$ be the solution of system* (1) *with $\lambda > 0$. Then, as $\lambda \to 0^+$, the functions $\lambda\mathbf{u}^{\lambda,c}$ uniformly converge in M to the constant vector $(c - c(\mathbb{H}))\mathbb{1}$ and the functions $\widehat{\mathbf{u}}^{\lambda,c} := \mathbf{u}^{\lambda,c} - \min_i \min_x u_i^{\lambda,c}\mathbb{1}$ uniformly converge to $\mathbf{u}^0 - \min_i \min_x u_i^0\mathbb{1}$ in M.*

Theorem 2 for $c = 0$ can be restated by saying that the *ergodic approximation* selects a specific critical solution in the limit. The ergodic approximation is a classical technique introduced in [15] for the case of a single equation (i.e. with $m = 1$ and $B = 0$). Since then, it has been extended and applied to many different settings, including the case of weakly coupled systems of Hamilton–Jacobi equations, see [3, 17]. This technique is typically employed to show the existence and uniqueness of the critical value $c(\mathbb{H})$ and the existence of a solution of the corresponding critical problem. This is achieved by renormalizing the discounted solutions so to produce a family of equi-bounded and equi-Lipschitz functions satisfying suitable perturbed discounted problems (for instance, the family $\{\widehat{\mathbf{u}}^{\lambda,0} \mid \lambda > 0\}$ in the case of HJ systems) and by taking limits, *along subsequences* as $\lambda \to 0^+$, of these renormalized functions. The fact that the limit is unique has been recently established in [6] for the case of a single equation by using tools and results issued from weak KAM theory. This selection principle was subsequently generalized in different directions, see [1, 7, 11, 13, 14, 19], testifying the interest for the issue.

The extension of the selection principle to HJ systems provided in the present work is based on a generalization of the theory of Mather minimizing measures, which is new in this setting and enriches the frame of analogies with weak KAM theory developed for scalar eikonal equations. This stream of research was initiated in [3] with the proof of the long-time convergence of the solutions to evolutive HJ systems, under hypotheses close to [20]. Other outputs in this vein can be found in a series of works including [18, 21]. The links with weak KAM theory were further made precise by the authors of the present paper in [9] where, by purely using PDE tools and viscosity solution techniques, an appropriate notion of Aubry set for systems was given and some relevant properties were generalized from the scalar case. A dynamical and variational point of view of the matter, integrating the PDE methods, was later brought in by [12, 16]. This angle allowed the authors to detect the stochastic character of the problem, displayed by the random switching nature of the dynamics and by the role of an adapted action functional. Representation formulae for viscosity (sub)solutions of the critical systems and a cycle characterization of the Aubry set were derived.

This random frame was further developed in [8] and applied to weakly coupled evolutive HJ equations. This work is the starting point of our analysis. It is exploited to provide suitable random representation formulae for the solutions of both the critical and the discounted system. A point that is crucial to our purposes consists in showing the existence of admissible minimizing curves in such formulae. This is done by making use of the results proved in [8] and by adapting the construction therein employed to the discounted system case.

We point out that our approach strongly relies on the assumptions made on the coupling matrix B. The sign and degeneracy conditions assumed on the elements of the coupling matrix B, see condition (B1) in Section 1.2, amount to requiring that $\{e^{-Bt}\}_{t\geqslant 0}$ is a semigroup of stochastic matrices, and this is at the base of the probabilistic and variational interpretation of the systems exploited in the paper. The irreducibility assumption on B, see condition (B2) in Section 1.2, is instead crucial for the extension of Aubry–Mather theory to systems provided in [9, 12, 16] and in the current paper. Yet, a generalization of Theorems 1 and 2 under a wider set of assumptions on the coupling matrix is possible. This issue will be investigated in the forthcoming paper [10].

The paper is organized as follows: in Section 1 we fix notations and the standing assumptions, and we provide some preliminary results on the critical and discounted systems. In Section 2 we present the random frame in which our analysis takes place and we prove suitable random representation formulae for the solutions of the critical and discounted systems. In Section 3 we generalize the theory of Mather minimizing measures to the case of HJ systems. Section 4 contains the proof of Theorem 1.

1 Preliminaries

1.1 Notations

In this work, we will denote by M the N-dimensional flat torus $\mathbb{T}^N$, where N is an integer number. This is done to simplify the notation and to be consistent with the references we will use. We remark however that our results and proofs keep holding, *mutatis mutandis*, whenever M is a compact connected Riemannian manifold without boundary. The associated Riemannian distance on M will be denoted by d. We denote by TM the tangent bundle and by (x, v) a point of TM, with $x \in M$ and $v \in T_xM = \mathbb{R}^N$. In the same way, a point of the cotangent bundle T^*M will be denoted by (x, p), with $x \in M$ and $p \in T_x^*M$ a linear form on the vector space T_xM. The latter will be identified with the vector $p \in \mathbb{R}^N$ such that

$$p(v) = \langle p, v\rangle \quad \text{for all } v \in T_xM = \mathbb{R}^N,$$

where $\langle\cdot,\cdot\rangle$ denotes the Euclidean scalar product in $\mathbb{R}^N$. The fibers T_xM and T_x^*M are endowed with the Euclidean norm $|\cdot|$, for every $x \in M$.

With the symbols $\mathbb{N}$ and $\mathbb{R}_+$ we will refer to the set of positive integer numbers and nonnegative real numbers, respectively. We say that a property holds *almost everywhere* (a.e. for short) in a subset E of M (respectively, of $\mathbb{R}$) if it holds up to a *negligible* subset of E, i.e. a subset of zero N-dimensional (resp., 1-dimensional) Lebesgue measure.

Given a continuous function u on M and a point $x_0 \in M$, we will denote by $D^-u(x_0)$ and $D^+u(x_0)$ the set of *subdifferential* and *superdifferential* of u at x_0, respectively. When u is locally Lipschitz in M, we will denote by $\partial^c u(x_0)$ the set of *Clarke's generalized gradient* of u at x_0, see [5] for a detailed presentation of the subject.

We will denote by $\|g\|_\infty$ the usual L^∞-norm of g, where the latter is a measurable real function defined on M. We will denote by $(\mathrm{C}(M))^m$ the Banach space of continuous functions $\mathbf{u} = (u_1, \dots, u_m)^T$ from M to $\mathbb{R}^m$, endowed with the norm

$$\|\mathbf{u}\|_\infty = \max_{1\leqslant i\leqslant m} \|u_i\|_\infty, \quad \mathbf{u} \in (\mathrm{C}(M))^m.$$

We will write $\mathbf{u}^n \rightrightarrows \mathbf{u}$ in M to mean that $\|\mathbf{u}^n - \mathbf{u}\|_\infty \to 0$. A function $\mathbf{u} \in (\mathrm{C}(M))^m$ will be termed Lipschitz continuous if each of its components is κ-Lipschitz continuous, for some $\kappa > 0$. Such a constant κ will be called a *Lipschitz constant* for $\mathbf{u}$. The space of all such functions will be denoted by $(\mathrm{Lip}(M))^m$.

We will denote by $\mathbb{1} = (1, \dots, 1)^T$ the vector of $\mathbb{R}^m$ having all components equal to 1, where the superscript symbol T stands for the transpose. We consider the following partial relations between elements $\mathbf{a}, \mathbf{b} \in \mathbb{R}^m$: $\mathbf{a} \leqslant \mathbf{b}$ if $a_i \leqslant b_i$ (resp., $<$) for every $i \in \{1, \dots, m\}$. Given two functions $\mathbf{u}, \mathbf{v} : M \to \mathbb{R}^m$, we will write $\mathbf{u} \leqslant \mathbf{v}$ in M (respectively, $<$) to mean that $\mathbf{u}(x) \leqslant \mathbf{v}(x)$ (resp., $\mathbf{u}(x) < \mathbf{v}(x)$) for every $x \in M$.

1.2 Weakly coupled systems

Throughout the paper, we will assume the Hamiltonians H_i to be continuous functions on T^*M satisfying, for every $i \in \{1, \dots, m\}$:

(H1) Convexity: $p \mapsto H_i(x, p)$ is convex on $\mathbb{R}^N$ for any $x \in M$.

(H2) Coercivity: there exist two coercive functions $\alpha, \beta : \mathbb{R}_+ \to \mathbb{R}$ such that

$$\alpha(|p|) \leqslant H_i(x, p) \leqslant \beta(|p|) \quad \text{for every } (x, p) \in T^*M.$$

For our analysis, it will be convenient and non-restrictive, see Section 2, to reinforce this coercivity condition in favor of the following:

(H2') Superlinearity: there exist two superlinear functions $\alpha, \beta : \mathbb{R}_+ \to \mathbb{R}$ such that

$$\alpha(|p|) \leqslant H_i(x, p) \leqslant \beta(|p|) \quad \text{for every } (x, p) \in T^*M.$$

We recall that a function $f : \mathbb{R}_+ \to \mathbb{R}$ is termed *coercive* if $f(h) \to +\infty$ as $h \to +\infty$, while it is termed *superlinear* if $\frac{f(h)}{h} \to +\infty$ as $h \to +\infty$.

In the sequel, we will denote by $\partial_p H_i(x, p)$ the set of subdifferentials at p of the function $p \mapsto H_i(x, p)$ in the sense of convex analysis. We recall that, due to conditions (H1)-(H2), the function $H_i(x, \cdot)$ is locally Lipschitz in $T^*_x M$, with a local Lipschitz constant that can be chosen independently of $x \in M$. In particular, the sets $\{\partial_p H_i(x, p) \mid x \in M,\ |p| \leqslant R\}$ are uniformly bounded for fixed $R > 0$.

The *coupling matrix* $B = (b_{ij})$ has dimensions $m \times m$ and satisfies

(B1) $b_{ij} \leqslant 0$ for $j \neq i$, $\sum_{j=1}^m b_{ij} = 0$,

(B2) B is irreducible, i.e. for every nonempty subset $\mathcal{I} \subsetneq \{1, \dots, m\}$ there exist $i \in \mathcal{I}$ and $j \notin \mathcal{I}$ such that $b_{ij} \neq 0$.

For $\lambda \geqslant 0$ and $c \in \mathbb{R}$, we consider the following weakly coupled system of Hamilton–Jacobi equations:

$$(B + \lambda \mathrm{Id})\mathbf{u} + \mathbb{H}(x, D\mathbf{u}) = c\mathbb{1} \quad \text{in } M, \tag{1.1}$$

where we have adopted the notation $\mathbb{H}(x, D\mathbf{u}) = (H_1(x, Du_1), \dots, H_m(x, Du_m))^T$.

Let $\mathbf{u} \in (\mathrm{C}(M))^m$. We will say that $\mathbf{u}$ is a *viscosity subsolution* of system (1.1) if the following inequality holds for every $(x, i) \in M \times \{1, \dots, m\}$:

$$H_i(x, p) + ((B + \lambda \mathrm{Id})\mathbf{u}(x))_i \leqslant c \quad \text{for every } p \in D^+u_i(x).$$

We will say that $\mathbf{u}$ is a *viscosity supersolution* of system (1.1) if the following inequality holds for every $(x, i) \in M \times \{1, \dots, m\}$:

$$H_i(x, p) + ((B + \lambda \mathrm{Id})\mathbf{u}(x))_i \geqslant c \quad \text{for every } p \in D^-u_i(x).$$

We will say that $\mathbf{u}$ is a *viscosity solution* if it is both a sub and a supersolution. In the sequel, solutions, subsolutions and supersolutions will be always meant in the viscosity sense, hence the adjective *viscosity* will be omitted.

When $\lambda = 0$, there exists a unique value c for which system (1.1) admits solutions, hereafter denoted by $c(\mathbb{H})$ and termed *critical*. In fact, $c(\mathbb{H})$ can be also characterized as

$$c(\mathbb{H}) = \min\{c \in \mathbb{R} \mid \text{system (1.1) with } \lambda = 0 \text{ admits subsolutions}\}, \tag{1.2}$$

see [9] for a detailed analysis.

We recall from [9] the following result that will be crucial for our analysis:

Proposition 1.1. *Let* $\mathbf{u} = (u_1, \dots, u_m)^T \in (\mathrm{C}(M))^m$ *be a subsolution of* (1.1) *with* $\lambda = 0$ *and* $c \in \mathbb{R}$. *Then there exist constants* C_c *and* κ_c*, only depending on* c*, on the Hamiltonians* $H_1, \dots, H_m$ *and on the coupling matrix* B*, such that*

(i) $\|u_i - u_j\|_\infty \leqslant C_c$ *for every* $i, j \in \{1, \dots, m\}$,

(ii) $\mathbf{u}$ *is* κ_c*-Lipschitz continuous in* M.

We proceed presenting some basic facts about the discounted system, i.e. system (1.1) when $\lambda > 0$. The following existence and uniqueness result depends on the fact that the matrix $B + \lambda \mathrm{Id}$ is non-degenerate as soon as $\lambda > 0$.

Proposition 1.2. *Let $\lambda > 0$ and $c \in \mathbb{R}$. Let $\mathbf{v}, \mathbf{u} \in (\mathrm{C}(M))^m$ be respectively a subsolution and a supersolution to (1.1). Then $\mathbf{v} \leqslant \mathbf{u}$. In particular, there exists a unique solution $\mathbf{u}^{\lambda,c}$ in $(\mathrm{C}(M))^m$.*

Proof. The first assertion is a consequence of [9, Proposition 2.8], while the second follows via a standard application of Perron's method. □

As already mentioned in the Introduction, the relationship between those solutions when c varies is given by $\mathbf{u}^{\lambda,c} = \mathbf{u}^{\lambda,c'} + \frac{c-c'}{\lambda}\mathbb{1}$. In particular, it follows that as $\lambda \to 0^+$, the family $\mathbf{u}^{\lambda,c}$ may be bounded at most for one value c.

We now explain why this is the case for $c = c(\mathbb{H})$.

Proposition 1.3. *Let us denote by $\mathbf{u}^\lambda$ the unique solution in $(\mathrm{C}(M))^m$ of (1.1) with $c = c(\mathbb{H})$ and $\lambda > 0$. Then the functions $\{\mathbf{u}^\lambda \mid \lambda > 0\}$ are equi-Lipschitz and equi-bounded. In particular, $\|\lambda \mathbf{u}^\lambda\|_\infty \to 0$ as $\lambda \to 0^+$.*

Proof. Let $\mathbf{u} \in (\mathrm{C}(M))^m$ be a solution of (1.1) with $c = c(\mathbb{H})$ and $\lambda = 0$. By taking $A > 0$ big enough, it follows that $\overline{u}u := \mathbf{u} + A\mathbb{1}$ takes only positive values and $\underline{\mathbf{u}} := \mathbf{u} - A\mathbb{1}$ takes only negative values. Therefore, $\overline{\mathbf{u}}$ and $\underline{\mathbf{u}}$ are respectively a super- and a subsolution of (1.1) with $c = c(\mathbb{H})$ for any parameter $\lambda > 0$. By Proposition 1.2 we infer that $\underline{\mathbf{u}} \leqslant \mathbf{u}^\lambda \leqslant \overline{\mathbf{u}}$ in M for all $\lambda > 0$, thus proving the asserted equi-bounded character of the $\{\mathbf{u}^\lambda \mid \lambda > 0\}$.

Let us now prove that $\mathbf{u}^\lambda$ is Lipschitz and its Lipschitz constant can be chosen independently of $\lambda > 0$. Let us set $b = \max_{i\in\{1,\dots,m\}} \max_{x\in M} H_i(x, 0)$. The function $\mathbf{w} \equiv -\mathbb{1}(b - c(\mathbb{H}))\frac{1}{\lambda}$ is obviously a subsolution of (1.1) with $c = c(\mathbb{H})$. By Proposition 1.2, we must have $\lambda\mathbf{u}^\lambda \geqslant (-b + c(\mathbb{H}))\mathbb{1}$ in M, hence

$$B\mathbf{u}^\lambda + \mathbb{H}(x, D\mathbf{u}^\lambda) = -\lambda\mathbf{u}^\lambda + c(\mathbb{H})\mathbb{1} \leqslant b\mathbb{1} \quad \text{in } M$$

in the viscosity sense. According to Proposition 1.1 we conclude that $\mathbf{u}^\lambda$ is κ-Lipschitz, where the constant κ only depends on the constant b, on the Hamiltonians $H_1, \dots, H_m$ and on the coupling matrix B. □

Remark 1.4. Note that $b := \max_{i\in\{1,\dots,m\}} \max_{x\in M} H_i(x, 0) \geqslant c(\mathbb{H})$. This readily follows from the characterization of $c(\mathbb{H})$ given in (1.2) after noticing that the null function is a subsolution of (1.1) with $\lambda = 0$ and $c = b$.

2 Random representation formulae for solutions

In this section, we will establish suitable representation formulae for the solution of the following system:

$$(B + \lambda\mathrm{Id})\mathbf{u} + \mathbb{H}(x, D\mathbf{u}) = c(\mathbb{H})\mathbb{1} \quad \text{in } M \tag{2.1}$$

when either $\lambda > 0$ or $\lambda = 0$. This will be done by adopting the random frame introduced in [8] and by adapting the strategy therein employed to the case at issue. In the sequel, we shall refer to system (2.1) and its corresponding (sub, super) solutions as *discounted* when $\lambda > 0$, *critical* when $\lambda = 0$.

To implement this program, we need to assume that the Hamiltonians satisfy the stronger growth assumption (H2'). We want to explain here why this is not restrictive for our analysis. According to the proof of Proposition 1.3, the discounted solutions $\mathbf{u}^\lambda$ satisfy

$$B\mathbf{u}^\lambda + \mathbb{H}(x, D\mathbf{u}^\lambda) \leqslant b\mathbb{1} \quad \text{in } M$$

in the viscosity sense with $b := \max_i \max_x H_i(x, 0)$. By Remark 1.4, this is also true for the (sub-)solutions of the critical system. So all these functions are κ-Lipschitz continuous, with $\kappa = \kappa_b$ chosen according to Proposition 1.1. We can therefore modify each Hamiltonian H_i outside the compact set $K := \{(x, p) \in T^*M : |p| \leqslant \kappa\}$ to obtain a new Hamiltonian $\widetilde{H}_i$ which is still continuous and convex, and satisfies the stronger growth condition (H2'). Since $H_i \equiv \widetilde{H}_i$ on K for each $i \in \{1, \dots, m\}$, it is easily seen that $c(\mathbb{H}) = c(\widetilde{\mathbb{H}})$ and the solutions of the corresponding critical and discounted systems are the same.

In the remainder of the paper, we will therefore assume each Hamiltonian H_i to be convex and superlinear in p, i.e. hypotheses (H1) and (H2') will be in force. This allows us to introduce the associated Lagrangian $L_i : TM \to \mathbb{R}$ defined as follows:

$$L_i(x, v) := \sup_{p\in\mathbb{R}^N} \{\langle p, v\rangle - H_i(x, p)\} \quad \text{for every } (x, v) \in TM. \tag{2.2}$$

As well known, L_i satisfies properties analogous to (H1)–(H2'). By the definition of L_i we derive

$$H_i(x,p) + L_i(x,v) \geqslant \langle p, q\rangle \quad \text{for all } (x,p) \in T^*M \text{ and } (x,v) \in TM,$$

which is known as *Fenchel's inequality*.

2.1 Random frame

We briefly recall the random frame in which our analysis takes place, see [8] for more details. We take as sample space Ω the space of paths $\omega : \mathbb{R}_+ \to \{1,\dots,m\}$ that are right-continuous and possess left-hand limits (known in literature as *càdlàg paths*, a French acronym for *continu à droite, limite à gauche*, see Billingsley's book [2] for a detailed treatment of the topic). By càdlàg property and the fact that the range of $\omega \in \Omega$ is finite, the points of discontinuity of any such path are isolated and consequently finite in compact intervals of $\mathbb{R}_+$ and countable (possibly finite) in the whole of $\mathbb{R}_+$. We call them *jump times* of ω.

The space Ω is endowed with a distance, named after *Skorohod*, see [2], which turns it into a Polish space. We denote by $\mathcal{F}$ the corresponding Borel σ-algebra and, for every $t \geqslant 0$, by $\pi_t : \Omega \to \{1,\dots,m\}$ the map that evaluates each ω at t, i.e. $\pi_t(\omega) = \omega(t)$ for every $\omega \in \Omega$. It is known that $\mathcal{F}$ is the minimal σ-algebra that makes all the functions π_t measurable, i.e. $\pi_t^{-1}(i) \in \mathcal{F}$ for every $i \in \{1,\dots,m\}$ and $t \geqslant 0$.

Let us now fix an $m \times m$ matrix B satisfying assumption (B1)-(B2). We record that e^{-tB} is a stochastic matrix for every $t \geqslant 0$, namely a matrix with nonnegative entries and with each row summing to 1. We endow Ω of a probability measure $\mathbb{P}$ defined on the σ-algebra $\mathcal{F}$ in such a way that the right-continuous process $(\pi_t)_{t\geqslant 0}$ is a *Markov chain with generator matrix* $-B$, i.e. it satisfies the Markov property

$$\mathbb{P}\big(\omega(t_k) = i_k \mid \omega(t_1) = i_1, \dots, \omega(t_{k-1}) = i_{k-1}\big) = \big(\mathrm{e}^{-B(t_k - t_{k-1})}\big)_{i_{k-1} i_k} \tag{2.3}$$

for all times $0 \leqslant t_1 < t_2 < \dots < t_k$, states $i_1,\dots,i_k \in \{1,\dots,m\}$ and $k \in \mathbb{N}$. We will denote by $\mathbb{P}_i$ the probability measure $\mathbb{P}$ conditioned to the event $\Omega_i := \{\omega \in \Omega \mid \omega(0) = i\}$ and write $\mathbb{E}_i$ for the corresponding expectation operators. It is easily seen that the Markov property (2.3) holds with $\mathbb{P}_i$ in place of $\mathbb{P}$, for every $i \in \{1,\dots,m\}$.

In the sequel, we will call *random variable* a map $X : (\Omega,\mathcal{F}) \to (\mathbb{F}, \mathscr{B}(\mathbb{F}))$, where $\mathbb{F}$ is a Polish space and $\mathscr{B}(\mathbb{F})$ its Borel σ-algebra, satisfying $X^{-1}(A) \in \mathcal{F}$ for every $A \in \mathscr{B}(\mathbb{F})$. Let us denote by $\mathrm{C}(\mathbb{R}_+; M)$ the Polish space of continuous paths taking values in M, endowed with a metric that induces the topology of local uniform convergence in $\mathbb{R}_+$.

We call *admissible curve* a random variable $\gamma : \Omega \to \mathrm{C}(\mathbb{R}_+; M)$ such that

(i) it is uniformly (in ω) locally (in t) absolutely continuous, i.e. given any bounded interval I and $\varepsilon > 0$, there is $\delta_\varepsilon > 0$ such that

$$\sum_j (b_j - a_j) < \delta_\varepsilon \implies \sum_j d(\gamma(b_j,\omega), \gamma(a_j,\omega)) < \varepsilon$$

for any finite family $\{(a_j, b_j)\}$ of pairwise disjoint intervals contained in I and for any $\omega \in \Omega$,

(ii) it is *nonanticipating*, i.e. for any $t \geqslant 0$,

$$\omega_1 \equiv \omega_2 \text{ in } [0,t] \implies \gamma(\cdot,\omega_1) \equiv \gamma(\cdot,\omega_2) \text{ in } [0,t].$$

We will say that γ is an admissible curve starting at $y \in M$ when $\gamma(0,\omega) = y$ for every $\omega \in \Omega$.

Given an admissible curve $\gamma : \Omega \to \mathrm{C}(\mathbb{R}_+; M)$ and $\omega \in \Omega$, we will denote by $\|\dot\gamma(\cdot,\omega)\|_\infty$ the L^∞-norm of the derivative of the curve $\gamma(\cdot,\omega)$.

We record for later use the following Dynkin's formula, see [8, Theorem 4.7] for a proof:

Theorem 2.1. *Let $\mathbf{g} : \mathbb{R}_+ \times M \to \mathbb{R}^m$ be a locally Lipschitz function and γ an admissible curve. Then, for every index $i \in \{1,\dots,m\}$, we have*

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}_i[g_{\omega(t)}(t,\gamma(t,\omega))]\Big|_{t=s} = \mathbb{E}_i\Big[-(B\mathbf{g})_{\omega(s)}(s,\gamma(s,\omega)) + \frac{\mathrm{d}}{\mathrm{d}t} g_{\omega(s)}(t,\gamma(t,\omega))\Big|_{t=s}\Big] \tag{2.4}$$

for a.e. $s \in \mathbb{R}_+$.

2.2 Representation formulae

In this subsection, we establish some representation formulae for solutions of system (2.1). We begin with the critical system.

Theorem 2.2. *Let* $\mathbf{u} \in (\mathrm{Lip}(M))^m$ *be a critical solution, namely a solution of system* (2.1) *with* $\lambda = 0$*, and let* $(y, \ell) \in M \times \{1, \dots, m\}$ *and* $t > 0$ *be fixed.*
(i) *The following holds:*

$$u_\ell(y) = \inf_{\gamma(0,\omega)=y} \mathbb{E}_\ell\Big[u_{\omega(t)}(\gamma(t,\omega)) + \int_0^t (L_{\omega(s)}(\gamma(s,\omega), -\dot\gamma(s,\omega)) + c(\mathbb{H}))\,\mathrm{d}s\Big],$$

where the minimization is performed over all admissible curves $\gamma : \Omega \to \mathrm{C}(\mathbb{R}_+; M)$ *starting at* y.
(ii) *There exists an admissible curve* $\eta : \Omega \to \mathrm{C}(\mathbb{R}_+; M)$ *starting at* y *for which such a minimum is attained. Moreover, for every* $\omega \in \Omega$*, the following holds:*

$$-\dot\eta(s,\omega) \in \partial_p H_{\omega(s)}(\eta(s,\omega), \partial^c u_{\omega(s)}(\eta(s,\omega))) \quad \textit{for a.e. } s \in (0,t). \tag{2.5}$$

In particular, there exists a constant k^**, only depending on* $H_1, \dots, H_m$ *and* B *such that* $\|\dot\eta(\cdot,\omega)\|_\infty \leqslant k^*$ *for every* $\omega \in \Omega$.

Proof. The assertion follows as a simple consequence of the results proved in [8]. It is easily seen that the function $\mathbf{v}(t,x) := \mathbf{u}(x)$ is a solution of the time-dependent system

$$\frac{\partial \mathbf{v}}{\partial t} + B\mathbf{v} + \mathbb{H}(x, D\mathbf{v}) - c(\mathbb{H})\mathbb{1} = 0 \quad \text{in } (0,+\infty) \times M$$

with initial datum $\mathbf{v}(0,\cdot) = \mathbf{u}$. Item (i) and the first assertion in (ii) readily follow from [8, Theorem 6.1]. Let us prove (2.5). Fix $\omega \in \Omega$. According to [8, Lemma 6.8 and Lemma 1.4], for a.e. $s \in (0,t)$ there exists $p_s \in \partial^c u_{\omega(s)}(\eta(s,\omega))$ such that

$$\langle p_s, -\dot\eta(s,\omega)\rangle = L_{\omega(s)}(\eta(s,\omega), -\dot\eta(s,\omega)) + c(\mathbb{H}) + \big(B\mathbf{u}(\eta(s,\omega))\big)_{\omega(s)},$$

hence, by Fenchel's duality we get

$$-\dot\eta(s,\omega) = \partial_p H_{\omega(s)}(\eta(s,\omega), p_s).$$

The remainder of the statement follows from Proposition 1.1 and the fact that $\partial_p H_i(x,p)$ is bounded on compact subsets of T^*M due to (H1)–(H2'). □

Let us now consider the discounted system.

Theorem 2.3. *Let* $\mathbf{u}^\lambda \in (\mathrm{Lip}(M))^m$ *be the solution of* (2.1) *with* $\lambda > 0$*. Let* $(y,\ell) \in M \times \{1,\dots,m\}$ *be fixed.*
(i) *The following holds:*

$$u^\lambda_\ell(y) = \inf_{\gamma(0,\omega)=y} \mathbb{E}_\ell\Big[\int_0^{+\infty} \mathrm{e}^{-\lambda s}\big(L_{\omega(s)}(\gamma(s,\omega), -\dot\gamma(s,\omega)) + c(\mathbb{H})\big)\,\mathrm{d}s\Big], \tag{2.6}$$

where the minimization is performed over all admissible curves $\gamma : \Omega \to \mathrm{C}(\mathbb{R}_+; M)$ *starting at* y.
(ii) *There exists an admissible curve* $\eta^\lambda : \Omega \to \mathrm{C}(\mathbb{R}_+; M)$ *starting at* y *for which such a minimum is attained. Moreover, for every* $\omega \in \Omega$*, the following holds:*

$$-\dot\eta^\lambda(s,\omega) \in \partial_p H_{\omega(s)}(\eta^\lambda(s,\omega), \partial^c u_{\omega(s)}(\eta^\lambda(s,\omega))) \quad \textit{for a.e. } s \in (0,+\infty). \tag{2.7}$$

In particular, there exists a constant k^**, only depending on* $H_1,\dots,H_m$ *and* B *such that* $\|\dot\eta^\lambda(\cdot,\omega)\|_\infty \leqslant k^*$ *for every* $\omega \in \Omega$ *and* $\lambda > 0$.

Proof. Let $\gamma : \Omega \to \mathrm{C}(\mathbb{R}_+; M)$ be an admissible curve starting at y. By applying Dynkin's formula to the function $\mathbf{g}(t,x) := \mathrm{e}^{-\lambda t}\mathbf{u}^\lambda(x)$ and by integrating (2.4) on $(0,+\infty)$ we get

$$u^\lambda_\ell(y) = \mathbb{E}_\ell\Big[\int_0^{+\infty} \mathrm{e}^{-\lambda t}((B+\lambda\mathrm{Id})\mathbf{u}^\lambda)_{\omega(s)}(\gamma(s,\omega))) + \langle Du^\lambda_{\omega(s)}(\gamma(s,\omega)), -\dot\gamma(s,\omega)\rangle\Big].$$

We now make use of Fenchel's inequality together with the fact that $\mathbf{u}^\lambda$ is a solution of the discounted system (2.1). Arguing as in the proof of [8, Proposition 5.6] we end up with

$$u_\ell^\lambda(y) \leqslant \mathbb{E}_\ell\left[\int_0^{+\infty} \mathrm{e}^{-\lambda s}\big(L_{\omega(s)}(\gamma(s,\omega), -\dot\gamma(s,\omega)) + c(\mathbb{H})\big)\,\mathrm{d}s\right]. \tag{2.8}$$

Next, we prove that there exists an admissible curve $\eta^\lambda : \Omega \to \mathrm{C}(\mathbb{R}_+; M)$ starting at y for which (2.8) holds with an equality. This will be obtained via a slight modification of the strategy employed in [8], to which we refer for more details. Let $\mathbf{v}(t,x) = \mathrm{e}^{\lambda t}\mathbf{u}^\lambda(x)$. It is readily verified that $\mathbf{v}$ verifies the following system:

$$\frac{\partial \mathbf{v}}{\partial t} + B\mathbf{v} + \mathrm{e}^{\lambda t}\big(\mathbb{H}(x, \mathrm{e}^{-\lambda t}D\mathbf{v}) - c(\mathbb{H})\mathbb{1}\big) = 0 \quad \text{in } (0,+\infty)\times M.$$

In particular, v_i is, for each fixed $i \in \{1,\dots,m\}$, a solution to the equation

$$\frac{\partial v_i}{\partial t} + G_i(t,x,Dv_i) = 0 \quad \text{in } (0,+\infty)\times M,$$

where $G_i(t,x,p) = \mathrm{e}^{\lambda t}(H_i(x,\mathrm{e}^{-\lambda t}p) - c(\mathbb{H})) + \sum_{k=1}^m b_{ik}v_k(t,x)$. As v_i is locally Lipschitz, it is standard, see for instance [8, Appendix A], that it verifies the following Lax–Oleinik formula for every $(t,z) \in (0,+\infty)\times M$:

$$v_i(t,z) = \inf_\gamma v_i(0,\gamma(-t)) + \int_{-t}^0 L_{G_i}(t+s,\gamma(s),\dot\gamma(s))\,\mathrm{d}s, \tag{2.9}$$

where L_{G_i} is the Lagrangian associated to G_i by duality and the infimum is taken amongst all absolutely continuous curves $\gamma : [-t,0] \to M$ such that $\gamma(0) = z$. By standard results in the Calculus of Variations, we know that this infimum is in fact a minimum. For any fixed $\tau > 0$, let us denote by $\gamma_{\tau,z} : [-\tau,0] \to M$ an absolutely continuous curve with $\gamma_{\tau,z}(0) = z$ and realizing the minimum in (2.9) with $t := \tau$. By the Dynamic Programming Principle, such a curve $\gamma_{\tau,z}$ is also a minimizer of (2.9) for every $t \leqslant \tau$. Arguing as in the proof of Theorem 2.2, we get

$$\dot\gamma_{\tau,z}(s) \in \partial_p G_i(t+s,\gamma_{\tau,z}(s),\partial^c v_i(t+s,\gamma_{\tau,z}(s))) = \partial_p H_i(\gamma_{\tau,z}(s),\partial^c u_i(s,\gamma_{\tau,z}(s))) \tag{2.10}$$

for a.e. $s \in (-t,0)$. Due to the equi-Lipschitz character of the functions $\{\mathbf{u}^\lambda \mid \lambda > 0\}$ established in Proposition 1.3, we infer that there exists a constant κ^*, independent of $(t,z) \in (0,+\infty)\times M$ and $\lambda > 0$, so that $\|\dot\gamma_{\tau,z}\|_\infty \leqslant \kappa^*$. Note that $L_{G_i}(t,x,v) = \mathrm{e}^{\lambda t}(L_i(x,v) + c(\mathbb{H}) - \sum_{k=1}^m b_{ik}u_k^\lambda(x))$. It follows that

$$u_i^\lambda(z) = \mathrm{e}^{-\lambda t}u_i^\lambda(\gamma_{\tau,z}(-t)) + \int_{-t}^0 \mathrm{e}^{\lambda s}\left(L_i(\gamma_{\tau,z}(s),\dot\gamma_{\tau,z}(s)) + c(\mathbb{H}) - \sum_{k=1}^m b_{ik}u_k^\lambda(\gamma_{\tau,z}(s))\right)\mathrm{d}s$$

for every $t \leqslant \tau$. Letting $\tau \to +\infty$ and extracting a subsequence, we obtain a curve $\gamma_{i,z} : (-\infty,0] \to M$ with $\gamma_{i,z}(0) = z$ and satisfying the previous equality for every $t > 0$. By sending $t \to +\infty$, we end up with

$$u_i^\lambda(z) = \int_{-\infty}^0 \mathrm{e}^{\lambda s}\left(L_i(\gamma_{i,z}(s),\dot\gamma_{i,z}(s)) + c(\mathbb{H}) - \sum_{k=1}^m b_{ik}u_k^\lambda(\gamma_{i,z}(s))\right)\mathrm{d}s. \tag{2.11}$$

Now the proof ends exactly as in [8]. For every $(z,i) \in M\times\{1,\dots,m\}$, we denote by $\Gamma(z,i)$ the set of absolutely continuous curves $\gamma : (-\infty,0] \to M$ with $\gamma(0) = z$ satisfying (2.11). The set $\Gamma(z,i)$ is nonempty, in view of the preceding discussion. Moreover, any curve in $\Gamma(z,i)$ satisfies (2.10) for a.e. $s \in (0,+\infty)$, in particular it is κ^*-Lipschitz continuous. We derive that $(z,i) \mapsto \Gamma(z,i)$ is compact-valued and upper semicontinuous as a set-valued map from $M\times\{1,\dots,m\}$ to $\mathrm{C}(\mathbb{R}_+; M)$, in particular it is measurable. By [4, Theorem III.8], there exists a measurable function $\Xi : M\times\{1,\dots,m\} \to \mathrm{C}(\mathbb{R}_+; M)$ such that

$$\Xi(z,i) \in \Gamma(z,i) \quad \text{for every } (z,i) \in M\times\{1,\dots,m\}.$$

For any fixed $\omega \in \Omega$, let $(\tau_k(\omega))_{k\geqslant 0}$ be the sequence of jump times of ω, where $\tau_0(\omega) := 0$ and $\tau_k(\omega)$ is the k-th jump time. We define inductively a sequence $(y_k(\omega))_{k\geqslant 0}$ of points in M by setting $y_0 := y$ and

$$y_k(\omega) := \Xi(y_{k-1}(\omega), \omega(\tau_{k-1}(\omega)))(\tau_k(\omega)) \quad \text{for every } k \geqslant 1.$$

The sought curve is given by

$$\eta^\lambda(t,\omega) := \Xi(y_k(\omega), \omega(\tau_k(\omega)))(-t) \quad \text{if } t \in [\tau_k(\omega), \tau_{k+1}(\omega)),$$

for every $k \geqslant 0$ and $\omega \in \Omega$. Arguing as in [8, Section 6], one can check that η^λ is an admissible curve starting at y for which (2.8) holds with an equality. The fact that η^λ satisfies (2.7) is clear by construction in view of (2.10). □

3 Mather measures for the critical system

In this section we generalize the notion of Mather minimizing measure to the case of the critical system, i.e.

$$B\mathbf{u} + \mathbb{H}(x, D\mathbf{u}) = c(\mathbb{H})\mathbb{1} \quad \text{in } M. \tag{3.1}$$

It is not so surprising that such measures will be concentrated on the support of minimizing controls associated to solutions of (3.1).

We start by adapting the notion of closed measure to this setting.

Definition 3.1. A Borel probability measure μ on $TM \times \{1,\dots,m\}$ will be termed closed if

(i) $$\int_{TM\times\{1,\dots,m\}} |v| \,\mathrm{d}\mu(x,v,i) < +\infty,$$

(ii) $$\int_{TM\times\{1,\dots,m\}} (B\boldsymbol{\phi}(x))_i + \langle D\phi_i(x), v\rangle \,\mathrm{d}\mu(x,v,i) = 0 \quad \text{for every } \boldsymbol{\phi} \in (\mathrm{C}^1(M))^m.$$

We will denote by $\mathfrak{M}$ the set of closed measures on $TM \times \{1,\dots,m\}$.

Theorem 3.2. *The following holds:*

$$-c(\mathbb{H}) = \min_{\mu\in\mathfrak{M}} \int_{TM\times\{1,\dots,m\}} L_i(x,v)\,\mathrm{d}\mu(x,v,i). \tag{3.2}$$

In particular, $\mathfrak{M}$ is nonempty.

Proof. We first observe that, for every $\varepsilon > 0$, there exists a function $\mathbf{w}^\varepsilon \in (\mathrm{C}^1(M))^m$ such that

$$B\mathbf{w}^\varepsilon + \mathbb{H}(x, D\mathbf{w}^\varepsilon) \leqslant (c(\mathbb{H}) + \varepsilon)\mathbb{1} \quad \text{for every } x \in M. \tag{3.3}$$

To see this, take a solution $\mathbf{u}$ of (3.1) and regularize it via convolution with a standard mollifier. The above inequality follows, for a proper choice of the mollifier, via a well known argument based on Jensen's inequality, the convexity of the Hamiltonians and the fact that $\mathbf{u}$ is Lipschitz, see for instance [9, Section 4].

By integrating (3.3) with respect to a measure $\mu \in \mathfrak{M}$ and by using Fenchel's inequality we get

$$\int_{TM\times\{1,\dots,m\}} (B\mathbf{w}^\varepsilon)_i + \langle Dw_i^\varepsilon(x), v\rangle - L_i(x,v)\,\mathrm{d}\mu(x,v,i) \leqslant c(\mathbb{H}) + \varepsilon.$$

Since μ is closed, the left hand side is equal to $-\int_{TM\times\{1,\dots,m\}} L_i \,\mathrm{d}\mu$. By letting $\varepsilon \to 0^+$, we obtain

$$\int_{TM\times\{1,\dots,m\}} L_i(x,v)\,\mathrm{d}\mu(x,v,i) \geqslant -c(\mathbb{H}).$$

Let us now proceed to prove the existence of a minimizing closed measure. To this aim, take a critical solution $\mathbf{u}$ and fix $(y,\ell) \in M \times \{1,\dots,m\}$. For every $k \in \mathbb{N}$, let $\eta_k : \Omega \to \mathrm{C}(\mathbb{R}_+; M)$ be an admissible curve starting at y and such that

$$u_\ell(y) = \mathbb{E}_\ell\Big[u_{\omega(k)}(\eta_k(k,\omega)) + \int_0^k (L_{\omega(s)}(\eta_k(s), -\dot\eta_k(s)) + c(\mathbb{H}))\,\mathrm{d}s\Big]. \tag{3.4}$$

We define a Borel probability measure μ_k on $TM \times \{1, \dots, m\}$ by setting

$$\int_{TM\times\{1,\dots,m\}} \mathbf{f}\, d\mu_k := \frac{1}{k}\, \mathbb{E}_\ell\Big[\int_0^k f_{\omega(s)}(\eta_k(s,\omega), -\dot\eta_k(s,\omega))\, ds\Big], \quad \mathbf{f} \in (C_c(TM))^m.$$

By Theorem 2.2, these measures have support contained in a common compact subset of $TM \times \{1, \dots, m\}$, so, up to subsequences, $(\mu_k)_k$ weakly converges to a Borel probability measure μ on $TM \times \{1, \dots, m\}$. Let us show that μ is closed. It clearly satisfies item (i) of Definition 3.1 since its support is compact. Let $\boldsymbol{\phi} \in (C^1(M))^m$. By applying Dynkin's formula to the function $\mathbf{g}(t, x) := \boldsymbol{\phi}(x)$, see Theorem 2.1, and by integrating (2.4) in $(0, k)$ we get

$$\mathbb{E}_\ell\Big[\int_0^k (B\boldsymbol{\phi})_{\omega(s)}(s, \eta_k(s,\omega)) + \langle D\phi_{\omega(s)}(\eta_k(s,\omega)), -\dot\eta_k(s,\omega)\rangle\, ds\Big] = \phi_\ell(y) - \mathbb{E}_\ell[\phi_{\omega(k)}(\eta_k(k,\omega))],$$

otherwise stated

$$\int_{TM\times\{1,\dots,m\}} (B\boldsymbol{\phi}(x))_i + \langle D\phi_i(x), v\rangle\, d\mu_k(x, v, i) = \frac{\phi_\ell(y) - \mathbb{E}_\ell[\phi_{\omega(k)}(\eta(k,\omega))]}{k}.$$

By sending $k \to +\infty$ we infer that μ satisfies item (ii) in Definition 3.1 as well. To prove that μ is minimizing, we remark that, in view of (3.4) and the fact that the measures $(\mu_k)_k$ have equi-compact support, we have

$$\begin{aligned}\int_{TM\times\{1,\dots,m\}} (L_i(x, v) + c(\mathbb{H}))\, d\mu &= \lim_{k\to+\infty} \int_{TM\times\{1,\dots,m\}} (L_i(x, v) + c(\mathbb{H}))\, d\mu_k \\ &= \lim_{k\to+\infty} \frac{1}{k}\big(u_\ell(y) - \mathbb{E}_\ell[u_{\omega(k)}(\eta_k(k,\omega))]\big) = 0.\end{aligned}$$ □

We will call *Mather measure* a closed Borel probability measure on $TM \times \{1, \dots, m\}$ which minimizes (3.2). The set of Mather measures will be denoted by $\mathfrak{M}_0$ in the sequel.

4 Convergence of the discounted solutions

This section is devoted to the proof of Theorem 1, namely that the solutions $(\mathbf{u}^\lambda)_{\lambda>0}$ of the discounted system (2.1) converge to a particular solution $\mathbf{u}^0$ of the critical system (3.1) as $\lambda \to 0^+$.

The first step consists in identifying a good candidate $\mathbf{u}^0$ for the limit of the solutions $\mathbf{u}^\lambda$. To this end, we consider the family $\mathcal{F}$ of subsolutions $\mathbf{w} \in (C(M))^m$ of the critical system (3.1) satisfying the following condition:

$$\int_{TM\times\{1,\dots,m\}} w_i(y)\, d\mu(y, v, i) \leqslant 0 \quad \text{for every } \mu \in \mathfrak{M}_0,$$

where $\mathfrak{M}_0$ denotes the set of Mather measures, see Section 3.

Note that, given any critical subsolution $\mathbf{w}$, the function $\mathbf{w} - \mathbb{1}\|\mathbf{w}\|_\infty$ is in $\mathcal{F}$. Therefore $\mathcal{F}$ is not empty.

Lemma 4.1. *The family $\mathcal{F}$ is uniformly bounded from above, i.e.*

$$\sup\{w_i(x) \mid \mathbf{w} \in \mathcal{F}\} < +\infty \quad \textit{for every } (x, i) \in M \times \{1, \dots, m\}.$$

Proof. Let us denote by κ and C the constants provided by Proposition 1.1 for $c := c(\mathbb{H})$. Pick $\mu \in \mathfrak{M}_0$. For $\mathbf{w} \in \mathcal{F}$, we have

$$\min_i \min_M w_i \leqslant \int_{TM\times\{1,\dots,m\}} w_i(y)\, d\mu(y, v, i) \leqslant 0.$$

Let $j \in \{1, \dots, m\}$ such that $\min_M w_j = \min_i \min_M w_i$. Since w_j is κ-Lipschitz, we infer

$$\max_M w_j \leqslant \max_M w_j - \min_M w_j \leqslant \kappa \operatorname{diam}(M) < +\infty.$$

On the other hand, for $i \neq j$ we have $w_i \leqslant w_j + \|w_i - w_j\|_\infty \leqslant \kappa \operatorname{diam}(M) + C$ in M. □

Therefore we can define $\mathbf{u}^0 : M \to \mathbb{R}^m$ by

$$u_i^0(x) := \sup_{\mathbf{w}\in\mathcal{F}} w_i(x) \quad \text{for every } (x,i) \in M \times \{1,\dots,m\}.$$

As the supremum of an equi-Lipschitz family of critical subsolutions, we get that $\mathbf{u}^0$ is Lipschitz continuous and a critical subsolution as well, see [9, Proposition 1.6]. The fact that $\mathbf{u}^0$ is a critical solution belonging to $\mathcal{F}$ will be a consequence of our convergence result.

We proceed by studying the asymptotic behavior of the discounted solutions $\mathbf{u}^\lambda$ as $\lambda \to 0^+$ and the relation with $\mathbf{u}^0$. Let us denote by

$$\mathfrak{S} := \Big\{ \mathbf{u} \in (\mathrm{Lip}(M))^m \;\Big|\; \mathbf{u} = \lim_{k\to+\infty} \mathbf{u}^{\lambda_k} \text{ for some sequence } \lambda_k \to 0 \Big\}.$$

Note that any function in $\mathfrak{S}$ is a solution to the critical system (3.1) by the stability of the notion of viscosity solution.

We begin with the following result:

Proposition 4.2. *Let* $\mathbf{u} \in \mathfrak{S}$. *Then*

$$\int_{TM\times\{1,\dots,m\}} u_i(x)\,\mathrm{d}\mu(x,v,i) \leqslant 0 \quad \textit{for every } \mu \in \mathfrak{M}_0.$$

In particular, $\mathbf{u} \leqslant \mathbf{u}^0$.

Proof. Fix $\mu \in \mathfrak{M}_0$. The assertion will be a direct consequence of the following fact:

$$\int_{TM\times\{1,\dots,m\}} u_i^\lambda(x)\,\mathrm{d}\mu(x,v,i) \leqslant 0 \quad \text{for every } \lambda > 0. \tag{4.1}$$

Indeed, let us fix $\lambda > 0$. Regularizing $\mathbf{u}^\lambda$ by convolution, we find a sequence of smooth functions $\mathbf{w}^n : M \to \mathbb{R}^m$ such that $\mathbf{w}^n \rightrightarrows \mathbf{u}^\lambda$ and

$$(B + \lambda \mathrm{Id})\mathbf{w}^n(x) + \mathbb{H}(x, D\mathbf{w}^n(x)) \leqslant \Big(c(\mathbb{H}) + \frac{1}{n}\Big)\mathbb{1} \quad \text{for every } x \in M.$$

By integrating this inequality with respect to μ and by using Fenchel's inequality we get

$$\begin{aligned} c(\mathbb{H}) + \frac{1}{n} &\geqslant \int_{TM\times\{1,\dots,m\}} \big((B+\lambda\mathrm{Id})\mathbf{w}^n(x)\big)_i + H_i(x, Dw_i^n)\,\mathrm{d}\mu \\ &\geqslant \int_{TM\times\{1,\dots,m\}} \big((B+\lambda\mathrm{Id})\mathbf{w}^n(x)^\varepsilon\big)_i + \langle Dw_i^n(x), v\rangle - L_i(x,v)\,\mathrm{d}\mu \\ &= c(\mathbb{H}) + \int_{TM\times\{1,\dots,m\}} \lambda w_i^n\,\mathrm{d}\mu, \end{aligned}$$

where for the last equality we have used the fact that μ is closed and minimizing. Inequality (4.1) follows after sending $n \to +\infty$ and dividing by $\lambda > 0$. □

The next (and final) step is to show that $\mathbf{u} \geqslant \mathbf{u}^0$ in M whenever $\mathbf{u} \in \mathfrak{S}$. This will be obtained by defining a special family of Borel probability measures on $TM \times \{1,\dots,m\}$ for the discounted systems (2.1). The construction is the following: fix $(y,\ell) \in M \times \{1,\dots,m\}$ and, for every $\lambda > 0$, let $\eta^\lambda : \Omega \to \mathrm{C}(\mathbb{R}_+; M)$ be an admissible curve starting at y that realizes the infimum in (2.6). We define a Borel probability measure μ_y^λ on $TM \times \{1,\dots,m\}$ by setting

$$\int_{TM\times\{1,\dots,m\}} f_i\,\mathrm{d}\mu_y^\lambda := \lambda \mathbb{E}_\ell\Bigg[\int_0^{+\infty} \mathrm{e}^{-\lambda s} f_{\omega(s)}(\eta^\lambda(s,\omega), -\dot\eta^\lambda(s,\omega))\,\mathrm{d}s\Bigg] \tag{4.2}$$

for every $\mathbf{f} \in (\mathrm{C}_c(TM))^m$. The following proposition holds.

Proposition 4.3. *The measures* $\{\mu_y^\lambda \mid \lambda > 0\}$ *defined above are Borel probability measures on* $TM \times \{1, \dots, m\}$*, whose supports are all contained in a common compact subset of* $TM \times \{1, \dots, m\}$*. In particular, they are relatively compact in the space of Borel probability measures on* $TM \times \{1, \dots, m\}$ *with respect to the weak convergence. Furthermore, if* $(\mu_y^{\lambda_n})_n$ *is weakly converging to* μ_y *for some sequence* $\lambda_n \to 0^+$*, then* μ_y *is a minimizing Mather measure.*

Proof. According to Theorem 2.3, there exists a constant κ^* such that $\|\dot\eta^\lambda(\cdot, \omega)\|_\infty \leqslant k^*$ for every $\omega \in \Omega$ and $\lambda > 0$. Set $K := \{(x, v) \in TM \mid |v| \leqslant \kappa^*\}$. Then the measures μ_y^λ are all supported in the compact set $K \times \{1, \dots, m\}$ and are Borel probability measures, as it can be easily checked by their definition. This readily implies the asserted relative compactness of $\{\mu_y^\lambda \mid \lambda > 0\}$. Let now assume that $(\mu_y^{\lambda_n})_n$ is weakly converging to μ_y for some $\lambda_n \to 0$. Then μ_y is a Borel probability measure with support in $K \times \{1, \dots, m\}$, in particular it satisfies item (i) in Definition 3.1. Moreover, if $\boldsymbol{\phi} \in (\mathrm{C}^1(M))^m$, by Dynkin's formula applied to the function $\mathbf{g}(t, x) := \mathrm{e}^{-\lambda t}\boldsymbol{\phi}(x)$, see Theorem 2.1, we get

$$\mathbb{E}_\ell\Bigg[\int_0^{+\infty} \mathrm{e}^{-\lambda s}\big(\langle D\phi_{\omega(s)}(\eta^\lambda(s,\omega)), -\dot\eta^\lambda(s,\omega)\rangle + (B\boldsymbol{\phi})_{\omega(s)}(\eta^\lambda(s,\omega)) + \lambda\phi_{\omega(s)}(\eta^\lambda(s,\omega))\big)\,\mathrm{d}s\Bigg] = \phi_\ell(y),$$

yielding

$$\int_{TM\times\{1,\dots,m\}} (B\boldsymbol{\phi}(x))_i + \langle D\phi_i(x), v\rangle \,\mathrm{d}\mu_y^\lambda = \lambda\phi_\ell(y) - \lambda \int_{TM\times\{1,\dots,m\}} \phi_i \,\mathrm{d}\mu_y^\lambda.$$

By setting $\lambda := \lambda_n$ in the previous equality and sending $n \to +\infty$ we infer

$$\int_{TM\times\{1,\dots,m\}} (B\boldsymbol{\phi}(x))_i + \langle D\phi_i(x), v\rangle \,\mathrm{d}\mu_y = 0,$$

thus proving that μ_y is closed.

To prove that μ_y is minimizing, we recall that, by definition,

$$\lambda u_\ell^\lambda(y) = \int_{TM\times\{1,\dots,m\}} (L_i(x, v) + c(\mathbb{H}))\,\mathrm{d}\mu_y^\lambda \quad \text{for every } \lambda > 0.$$

The assertion follows by setting $\lambda := \lambda_n$ and sending $n \to +\infty$. □

We proceed by proving a lemma that will be crucial for the proof of Theorem 1.

Lemma 4.4. *Let* $\mathbf{w}$ *be any critical subsolution. For every* $\lambda > 0$ *and* $(y, \ell) \in M \times \{1, \dots, m\}$ *we have*

$$u_\ell^\lambda(y) \geqslant w_\ell(y) - \int_{TM\times\{1,\dots,m\}} w_i \,\mathrm{d}\mu_y^\lambda, \tag{4.3}$$

where μ_y^λ *is the Borel probability measure defined by* (4.2).

Proof. Let $\mathbf{w}$ be a critical subsolution. By convolution with a regularizing kernel, we construct a family of smooth function $\mathbf{w}^n : M \to \mathbb{R}^m$ uniformly converging $\mathbf{w}$ such that

$$B\mathbf{w}^n(x) + \mathbb{H}(x, D\mathbf{w}^n(x)) \leqslant \Big(c(\mathbb{H}) + \frac{1}{n}\Big)\mathbb{1} \quad \text{for every } x \in M.$$

Starting again from the definition of $\mathbf{u}^\lambda$ and by exploiting Fenchel's inequality, we obtain

$$\begin{aligned}
u_\ell^\lambda(y) &= \frac{1}{\lambda}\int_{TM\times\{1,\dots,m\}} (L_i(x, v) + c(\mathbb{H}))\,\mathrm{d}\mu_y^\lambda \\
&\geqslant \frac{1}{\lambda}\int_{TM\times\{1,\dots,m\}} (\langle Dw_i^n(x), v\rangle - H_i(x, Dw_i^n(x)) + c(\mathbb{H}))\,\mathrm{d}\mu_y^\lambda \\
&\geqslant \frac{1}{\lambda}\int_{TM\times\{1,\dots,m\}} \Big(\langle Dw_i^n(x), v\rangle + (B\mathbf{w}^n(x))_i - \frac{1}{n}\Big)\,\mathrm{d}\mu_y^\lambda.
\end{aligned}$$

Using the definition of μ_y^λ and Dynkin's formula with $\mathbf{g}(t,x) = \mathrm{e}^{-\lambda t}\mathbf{w}^n(x)$, see Theorem 2.1, we get

$$\begin{aligned} u_\ell^\lambda(y) &\geqslant \mathbb{E}_\ell\bigg[\int_0^{+\infty} \mathrm{e}^{-\lambda s}\Big(\langle Dw^n_{\omega(s)}(\eta^\lambda(s,\omega)), -\dot\eta^\lambda(s,\omega)\rangle + (B\mathbf{w}^n)_{\omega(s)}(\eta^\lambda(s,\omega)) - \frac{1}{n}\Big)\,\mathrm{d}s\bigg] \\ &= w_\ell^n(y) + \mathbb{E}_\ell\bigg[\int_0^{+\infty}\Big(-\lambda \mathrm{e}^{-\lambda s} w^n_{\omega(s)}(\eta^\lambda(s,\omega)) - \frac{\mathrm{e}^{-\lambda s}}{n}\Big)\,\mathrm{d}s\bigg] \\ &= w_\ell^n(y) - \int_{TM\times\{1,\dots,m\}} w_i^n\,\mathrm{d}\mu_y^\lambda - \frac{1}{\lambda n}. \end{aligned}$$

The desired inequality follows by sending $n \to +\infty$. □

We have now all the ingredients to prove our main result.

Proof of Theorem 1. Let $\mathbf{u} \in \mathfrak{S}$. By Proposition 4.2, we already know that $\mathbf{u} \leqslant \mathbf{u}^0$. Let us prove the opposite inequality. By definition, there exists a sequence $\lambda_n \to 0^+$ such that $\mathbf{u}^{\lambda_n} \rightrightarrows \mathbf{u}$ as $n \to +\infty$. Pick $\mathbf{w} \in \mathcal{F}$ and fix $(y,\ell) \in TM \times \{1,\dots,m\}$. By setting $\lambda := \lambda_n$ in (4.3) and by sending $n \to +\infty$, we infer, thanks to Proposition 4.3, that there exists a Mather measure $\mu_y \in \mathfrak{M}_0$ such that

$$u_\ell(y) \geqslant w_\ell(y) - \int_{TM\times\{1,\dots,m\}} w_i\,\mathrm{d}\mu_y \geqslant w_\ell(y),$$

where, for the last inequality, we have used the fact that $\mathbf{w} \in \mathcal{F}$. As this is true for any $\mathbf{w} \in \mathcal{F}$ and arbitrary $(y,\ell) \in M \times \{1,\dots,m\}$, we infer that $\mathbf{u} \geqslant \mathbf{u}^0$. This concludes the proof. □

Acknowledgment: This research was initiated in May 2015, while the first author was visiting, as Professeur Invité, the Institut de Mathématiques de Jussieu, Université Pierre et Marie Curie (Paris), that he gratefully acknowledges for the financial support and hospitality.

Funding: Work supported by ANR-07-BLAN-0361-02 KAM faible and ANR-12-BS01-0020 WKBHJ. The second author was also supported by a PEPS fellowship.

References

[1] E. S. Al-Aidarous, E. O. Alzahrani, H. Ishii and A. M. M. Younas, A convergence result for the ergodic problem for Hamilton–Jacobi equations with Neumann-type boundary conditions, *Proc. Roy. Soc. Edinburgh Sect. A* **146** (2016), no. 2, 225–242.
[2] P. Billingsley, *Convergence of Probability Measures*, 2nd ed., Wiley Ser. Probab. Stat., John Wiley & Sons, New York, 1999.
[3] F. Camilli, O. Ley, P. Loreti and V. D. Nguyen, Large time behavior of weakly coupled systems of first-order Hamilton–Jacobi equations, *NoDEA Nonlinear Differential Equations Appl.* **19** (2012), no. 6, 719–749.
[4] C. Castaing and M. Valadier, *Convex Analysis and Measurable Multifunctions*, Lecture Notes in Math. 580, Springer, Berlin, 1977.
[5] F. H. Clarke, *Optimization and Nonsmooth Analysis*, Canad. Math. Soc. Ser. Monogr. Adv. Texts, John Wiley & Sons, New York, 1983.
[6] A. Davini, A. Fathi, R. Iturriaga and M. Zavidovique, Convergence of the solutions of the discounted Hamilton–Jacobi equation: Convergence of the discounted solutions, *Invent. Math.* **206** (2016), no. 1, 29–55.
[7] A. Davini, A. Fathi, R. Iturriaga and M. Zavidovique, Convergence of the solutions of the discounted equation: The discrete case, *Math. Z.* **284** (2016), no. 3–4, 1021–1034.
[8] A. Davini, A. Siconolfi and M. Zavidovique, Random Lax–Oleinik semigroups for Hamilton–Jacobi systems, *J. Math. Pures Appl.* **120** (2018), 294–333.
[9] A. Davini and M. Zavidovique, Aubry sets for weakly coupled systems of Hamilton–Jacobi equations, *SIAM J. Math. Anal.* **46** (2014), no. 5, 3361–3389.
[10] A. Davini and M. Zavidovique, A general weak KAM theorem for HJ systems and applications, in preparation.

[11] D. A. Gomes, H. Mitake and H. V. Tran, The selection problem for discounted Hamilton–Jacobi equations: Some non-convex cases, *J. Math. Soc. Japan* **70** (2018), no. 1, 345–364.
[12] H. Ibrahim, A. Siconolfi and S. Zabad, Cycle characterization of the Aubry set for weakly coupled Hamilton–Jacobi systems, *Commun. Contemp. Math.* **20** (2018), no. 6, Article ID 1750095.
[13] H. Ishii, H. Mitake and H. V. Tran, The vanishing discount problem and viscosity Mather measures. Part 1: The problem on a torus, *J. Math. Pures Appl. (9)* **108** (2017), no. 2, 125–149.
[14] H. Ishii, H. Mitake and H. V. Tran, The vanishing discount problem and viscosity Mather measures. Part 2: Boundary value problems, *J. Math. Pures Appl. (9)* **108** (2017), no. 3, 261–305.
[15] P.-L. Lions, G. Papanicolaou and S. R. S. Varadhan, Homogenization of Hamilton–Jacobi equation, unpublished preprint (1987).
[16] H. Mitake, A. Siconolfi, H. V. Tran and N. Yamada, A Lagrangian approach to weakly coupled Hamilton–Jacobi systems, *SIAM J. Math. Anal.* **48** (2016), no. 2, 821–846.
[17] H. Mitake and H. V. Tran, Remarks on the large time behavior of viscosity solutions of quasi-monotone weakly coupled systems of Hamilton–Jacobi equations, *Asymptot. Anal.* **77** (2012), no. 1–2, 43–70.
[18] H. Mitake and H. V. Tran, Homogenization of weakly coupled systems of Hamilton–Jacobi equations with fast switching rates, *Arch. Ration. Mech. Anal.* **211** (2014), no. 3, 733–769.
[19] H. Mitake and H. V. Tran, Selection problems for a discount degenerate viscous Hamilton–Jacobi equation, *Adv. Math.* **306** (2017), 684–703.
[20] G. Namah and J.-M. Roquejoffre, Convergence to periodic fronts in a class of semilinear parabolic equations, *NoDEA Nonlinear Differential Equations Appl.* **4** (1997), no. 4, 521–536.
[21] V. D. Nguyen, Some results on the large-time behavior of weakly coupled systems of first-order Hamilton–Jacobi equations, *J. Evol. Equ.* **14** (2014), no. 2, 299–331.