SIAM J. MATH. ANAL.
Vol. 48, No. 2, pp. 821–846

A LAGRANGIAN APPROACH TO WEAKLY COUPLED HAMILTON–JACOBI SYSTEMS*

H. MITAKE†, A. SICONOLFI‡, H. V. TRAN§, AND N. YAMADA¶

Abstract. We perform a qualitative analysis of a class of weakly coupled Hamilton–Jacobi systems in the spirit of weak KAM theory. We define a family of related action functionals containing the Lagrangians associated with the Hamiltonians of the system. We use them to characterize the subsolutions of the system and to provide explicit representation formulae for subsolutions enjoying an additional maximality property. A crucial step for our analysis is to put the problem in a suitable random frame. The presentation is accessible to readers without a background in probability; only some basic knowledge of measure theory is required.

Key words. Hamilton–Jacobi equation, weakly coupled system, weak KAM theory

AMS subject classifications. 35F21, 35F50, 49L25

DOI. 10.1137/15M1010841

1. Introduction.

This paper deals with weakly coupled Hamilton–Jacobi systems of the form

$$\text{(HJ}\alpha\text{)} \qquad \begin{cases} H_1(x, Du_1) + \Lambda^1 \cdot \mathbf{u} = \alpha, \\ \cdots \\ H_M(x, Du_M) + \Lambda^M \cdot \mathbf{u} = \alpha \end{cases}$$

posed on the flat torus $\mathbb{T}^N$. Here $\mathbf{u} = (u_1, \ldots, u_M)$ is the vector-valued unknown function, Du_i is the gradient of u_i, α is a real number, and H_i are mutually unrelated convex Hamiltonians enjoying standard additional properties (see section 2). The Λ^i are the rows of the so-called $M \times M$ coupling matrix $\Lambda := (\Lambda^1 \cdots \Lambda^M)$, which constitutes the relevant item in the problem.

We are interested in the setting which should correspond in the scalar case, namely when $M = 1$ and Λ is just a constant, to taking $\Lambda = 0$. In this case the system reduces to a single equation on $\mathbb{T}^N$ not directly depending on the unknown and classified as being of Eikonal type.

In this framework a rich qualitative theory has been developed by linking PDE facts to geometrical/dynamical properties. Representation formulae for (sub)solutions have been provided through minimization of a suitable action functional. The existence of a unique value of α has been shown, called a critical value, for which (viscosity) solutions do exist. This material has found applications in a variety of related asymptotic problems, and, furthermore, connections with Hamiltonian dynamics have been

*Received by the editors March 2, 2015; accepted for publication (in revised form) December 17, 2015; published electronically March 3, 2016.
http://www.siam.org/journals/sima/48-2/M101084.html

†Institute for Sustainable Sciences and Development, Hiroshima University, Higashi-Hiroshima-shi 739-8527, Japan (hiroyoshi-mitake@hiroshima-u.ac.jp). The work of this author was partially supported by KAKENHI 15K17574, 26287024, and 23244015.

‡Dipartimento di Matematica, Università degli Studi di Roma “La Sapienza,” 00185 Rome, Italy (siconolf@mat.uniroma1.it). The work of this author was partially supported by Programma Ricerca Scientifica Sapienza 2013.

§Department of Mathematics, The University of Chicago, Chicago, IL 60637 (hung@math.uchicago.edu). The work of this author was partially supported by NSF grant DMS-1361236.

¶Department of Applied Mathematics, Faculty of Science, Fukuoka University, Fukuoka 814-0180, Japan (nyamada@math.sci.fukuoka-u.ac.jp).

established. This body of results is a part of the so-called weak KAM theory; see [1, 5, 7, 8, 9, 11] for details.

If instead $\Lambda > 0$, the corresponding equation can be uniquely solved on the whole torus for any α and the solution is the value function of a related control problem with Λ playing the role of discount factor.

To find an analogue of the Eikonal case for systems, it is convenient to start from [12], where the class of monotone systems is introduced, and existence and uniqueness results of (viscosity) solutions are established. Regarding our system, being a monotone one corresponds to the following conditions on the coupling matrix:

- any nondiagonal entry of Λ is nonpositive;
- Λ is diagonal dominant, namely $\sum_{j=1}^{M} \Lambda_{ij} \geq 0$ for any $i \in \{1, \dots, M\}$;
- strict diagonal dominance holds at least for one row.

This setting should be put in relation to strict positiveness in the scalar case, and in this perspective it is consistent to focus on the limit setup where Λ satisfies the following conditions:

- any nondiagonal entry of Λ is nonpositive;
- any row sums to 0.

It was first pointed out in [4, 17, 18, 19] that under the above assumptions some phenomena, already occurring in the Eikonal scalar case, also take place for systems and can be analyzed in the spirit of the weak KAM theory. The properties of the critical value have been investigated, namely the minimal value for which the corresponding system admits subsolutions. Applications to asymptotic problems have been considered as well. Control interpretation for the Hamilton–Jacobi system has been investigated in [18, 19]. In [10], the weak KAM theorem has been studied in connection with another type of system. In [3], the nonlinear adjoint method was introduced to study the rate of convergence to the critical value of the systems, and the solutions to the adjoint systems are related to the projected Mather measures.

A significant step forward in the qualitative analysis of systems has more recently been performed in [6]. The authors have proved the existence of a subset $\mathcal{A}$ of the torus, named after Aubry, with the property that any critical subsolution, maximal among those taking a given admissible value at a point $y \in \mathcal{A}$, is indeed a solution. Note that there is a restriction in the values that a subsolution of the system can assume at any given point. This is a further relevant property pointed out in [6], which genuinely depends on the vectorial structure of the problem and has no equivalent in the scalar case.

All the above results pertain to the PDE side of the theory and are solely obtained by means of PDE techniques. The geometric counterpart is so far missed, and the intertwining between PDE and dynamical aspects, which is at the core of the weak KAM theory, consequently still has to be understood in the framework of systems. This is actually the primary task our paper is centered upon and is above all performed by putting the problem in a suitable random frame.

As a first step we consider all the possible switchings between indices $\{1, \dots, \mathrm{M}\}$ of the system on an infinite time horizon. This gives rise to the space of $\{1, \dots, \mathrm{M}\}$-valued càdlàg paths, denoted by $\mathcal{D}$, endowed with the Skorohod metric (see Appendix B) and the corresponding Borel σ-algebra $\mathcal{F}$. The coupling matrix, under our assumptions being the generator of a semigroup of stochastic matrices, induces a linear correspondence between the simplex of probability vectors of $\mathbb{R}^M$, i.e., with nonnegative components summing to 1, and a simplex of $\mathcal{F}$-probability measures on $\mathcal{D}$; see subsection 3.1.

This construction is indeed equivalent to that of a Markov chain with rate matrix

$-\Lambda$; formula (3.1) defining the family of probability measures is nothing but the usual finite-dimensional distribution formula with given initial distribution. However, we would like to emphasize that the advantage of our approach is avoiding having to introduce an abstract probability space; we just work with concrete path spaces. We also avoid explicitly using notions such as stochastic process, conditional probability, and other probabilistic tools. This makes the presentation self-contained.

We make corresponding to elements of $\mathcal{D}$ $\mathbb{R}^N$-valued càdlàg velocity paths and get by integrating them the admissible random curves on $\mathbb{T}^N$; see subsection 3.3. Action functionals are then obtained by averaging, with respect to the above probability measures on $\mathcal{D}$, line integrals over random curves of the Lagrangians associated with the Hamiltonians of the system; see (4.1). This justifies the title of the paper.

The effectiveness of our approach is demonstrated by recovering some crucial facts of the scalar case. Namely, we fully characterize all subsolutions of the system, for any α greater than or equal to the critical value, as the functions from $\mathbb{T}^N$ to $\mathbb{R}^M$ satisfying a suitable estimate with respect to our action functionals; see section 4 and Theorem 5.7. We use the action functionals to explicitly represent critical and supercritical subsolutions enjoying an additional maximality property, through a suitable minimization procedure; see Theorem 5.2. We characterize the values that a subsolution can assume at a given point; see Theorem 5.5. By this way we also provide a representation formula for critical solutions taking a prescribed admissible value at a given point of the Aubry set, complementing the result of [6]; see Theorem 5.6.

The paper is organized as follows: In section 2 we set forth the problem and recall some known facts about critical/supercritical subsolutions and the Aubry set. Section 3 is devoted to illustrating the random frame in which our qualitative analysis takes place. The family of probability measures $\mathbb{P}_{\mathbf{a}}$ is introduced, for any probability vector $\mathbf{a}$ of $\mathbb{R}^M$, and key notions such as admissible control and stopping time are given. In section 4 we define the action functionals and prove the fundamental estimate for subsolutions. Section 5 is about representation formulae for subsolutions and related results. Finally, the two appendices gather basic material on stochastic matrices and spaces of càdlàg paths.

2. Setting of the problem.

Here we introduce the system, which is the object of investigation, as well as standing assumptions and basic preliminary facts. We refer the reader to [4, 6, 17, 18] for proofs and more details on the results stated.

As already pointed out in the introduction, we will be interested in the one-parameter family of systems of the form

$$\text{(HJ}\alpha\text{)} \qquad \begin{cases} H_1(x, Du_1) + \Lambda^1 \cdot \mathbf{u} = \alpha, \\ \cdots \\ H_M(x, Du_M) + \Lambda^M \cdot \mathbf{u} = \alpha \end{cases}$$

posed on the flat torus $\mathbb{T}^N$ identified with $\mathbb{R}^N/\mathbb{Z}^N$. Here $\mathbf{u} = (u_1, \dots, u_M)$ is the vector-valued unknown function, Λ^i are the vectors given by the rows of the $M \times M$ *coupling matrix* Λ, and α varies in $\mathbb{R}$. The following conditions will be assumed throughout the paper without any further mentioning. On the Hamiltonians H_i we require the following:

(H1) H_i is continuous in both variables;
(H2) H_i is convex in p;
(H3) H_i is superlinear in p.

The growth condition in (H3), together with (H1), (H2), allows defining the corresponding Lagrangians via the Legendre–Fenchel transform, namely

$$L_i(x,q) = \max_{p\in\mathbb{R}^n} \big(p\cdot q - H_i(x,p)\big) \quad \text{for any } i,$$

and they inherit from H_i the properties of being continuous, convex, and superlinear at infinity.

We furthermore require on coupling matrix Λ the following:

(H4) any nondiagonal entry of Λ is nonpositive;
(H5) any row of Λ sums to 0;
(H6) Λ is irreducible.

Irreducible means that, given any nonempty subset of indices $I \subsetneq \{1,\dots,\mathrm{M}\}$, there is $i\in I$, $j\notin I$ with $\Lambda_{ij}\neq 0$; loosely speaking, this condition means that the system cannot be split into separated subsystems.

As made precise in Appendix A, the key point is that (H4), (H5) are equivalent to $-\Lambda$ being the generator of a semigroup of stochastic matrices. We also recall that under (H4), (H5), (H6) the matrix Λ is singular with rank $M-1$ and kernel spanned by $\mathbf{1}$, namely the vector with all components equal to 1; moreover, $\mathrm{im}(\Lambda)$ cannot contain vectors with strictly positive or negative components. This in particular implies $\mathrm{im}(\Lambda)\cap\ker(\Lambda)=\{0\}$.

NOTATION 2.1. *The projection of $\mathbb{R}^N$ onto $\mathbb{T}^N=\mathbb{R}^N/\mathbb{Z}^N$ induces the structure of an additive group on $\mathbb{T}^N$. To ease notation we will indicate throughout the paper by the usual symbols $+$, $-$ the corresponding operations between elements of the torus.*

The notion of viscosity (sub/super)solution can be easily adapted to systems such as (HJα); we will drop in the following the term viscosity since no other kind of weak solution will be considered.

DEFINITION 2.2. *A continuous function $\mathbf{u}=(u_1,\dots,u_M)$ is a subsolution (resp., supersolution) of* (HJα) *if the inequality*

$$H_i(x, D\psi(x)) + \Lambda^i\cdot \mathbf{u}(x) \le \alpha \quad (\textit{resp.}, \ge \alpha)$$

holds for every $x\in\mathbb{T}^N$, $i\in\{1,\dots,\mathrm{M}\}$, and $\psi\in C^1(\mathbb{T}^n)$ such that $u_i-\psi$ attains a maximum (resp., minimum) at x. We call $\mathbf{u}$ a solution if it is both a subsolution and a supersolution.

Remark 2.3. One could wonder why we are considering systems with the same constant appearing in the right-hand side of any equation, while a more natural condition should be to have instead a vector of $\mathbb{R}^M$, say $\mathbf{a}$, with possibly different components. We point out that, under our assumptions, such a setting is actually no more general. In fact, if we write the vector $\mathbf{a}$ as $\mathbf{a}_1+\mathbf{a}_2$ with $\mathbf{a}_1=\alpha\,\mathbf{1}\in\ker(\Lambda)$, $\mathbf{a}_2\in\mathrm{im}(\Lambda)$, where this form is uniquely determined because $\mathrm{im}(\Lambda)\cap\ker(\Lambda)=\{0\}$, and pick $\mathbf{b}$ with $\Lambda\,\mathbf{b}=-\mathbf{a}_2$, then $\mathbf{u}$ is a (super/sub)solution to (HJα) if and only if $\mathbf{u}+\mathbf{b}$ satisfies the same properties for the system obtained from (HJα) by replacing in the right-hand side $\alpha\,\mathbf{1}$ by $\mathbf{a}$.

Remark 2.4. Due to the coercivity condition, any subsolution to (HJα) is Lipschitz continuous. Moreover, owing to the convexity of the Hamiltonians, the notion of viscosity and a.e. subsolutions are equivalent for (HJα). Furthermore, we can express the same property using generalized gradients of any component in the sense of Clarke. Namely, $\mathbf{w}$ is a subsolution to (HJα) if and only if

$$H_i(x,p) + \Lambda^i\cdot\mathbf{w}(x) \le \alpha$$

for any $x \in \mathbb{T}^N$, $p \in \partial w_i(x)$, $i \in \{1, \dots, \mathrm{M}\}$, where $\partial w_i(x)$ indicates the generalized gradient of w_i at x.

Here are two basic propositions.

Proposition 2.5. *The family of all subsolutions to* (HJα)*, if nonempty, is equi-Lipschitz continuous with Lipschitz constant denoted by ℓ_α.*

Proposition 2.6. *The family of subsolutions to* (HJα) *taking the same value at a given point, if nonempty, admits a maximal element.*

We define the *critical value* γ as

$$\gamma = \inf\{\alpha \in \mathbb{R} \mid (\mathrm{HJ}\alpha) \text{ admits subsolutions}\}.$$

The infimum in the definition of γ is actually a minimum, as made precise below.

Proposition 2.7. *The critical system* (HJγ) *is unique in the one-parameter family* (HJα)*, $\alpha \in \mathbb{R}$, for which there are solutions.*

Following [6], we give the definition of the Aubry set $\mathcal{A} \subset \mathbb{T}^N$ from the PDE point of view.

Definition 2.8. *A point y belongs to the Aubry set $\mathcal{A}$ if any maximal critical subsolution taking a given value at y is a solution to* (HJγ).

Roughly speaking, the Aubry set, which is a closed nonempty subset of $\mathbb{T}^N$, is the place where the obstruction in getting subsolutions of system below the critical level is concentrated. More specifically, there cannot be any critical subsolution which is, in addition, locally strict at a point in $\mathcal{A}$, in the sense of the above definition.

Definition 2.9. *For a given critical subsolution $\mathbf{u}$, a component u_i, for some $i \in \{1, \dots, \mathrm{M}\}$, is said to be* locally strict *at a point $y \in \mathbb{T}^N$ if there is a neighborhood U of y and a positive constant δ with*

$$H_i(x, Du_i) + \Lambda^i \cdot \mathbf{u} \le \gamma - \delta \quad \textit{a.e. } x \in U.$$

In analogy with the scalar case, we have the following property.

Proposition 2.10 (see [6, Proposition 3.9]). *A point $y \notin \mathcal{A}$ if and only if for any given index $i \in \{1, \dots, \mathrm{M}\}$ there exists a critical subsolution $\mathbf{u}$ with u_i locally strict at y.*

An interesting fact pointed out in [6] is that there is a restriction on the values that a subsolution to (HJα) can attain at a given point. This is a property due to the vectorial structure of the problem and has no counterpart in the scalar case. The authors refer to it as *rigidity property* or rigidity phenomenon. For $\alpha \ge \gamma$, we define for $x \in \mathbb{T}^N$

$$F_\alpha(x) = \{\mathbf{b} \in \mathbb{R}^M \mid \exists\, \mathbf{u} \text{ subsolution to } (\mathrm{HJ}\alpha) \text{ with } \mathbf{u}(x) = \mathbf{b}\}. \tag{2.1}$$

Notice that $F_\alpha(x)$ is convex because of the convex character of the Hamiltonians; in addition, if $\mathbf{b} \in F_\alpha(x)$, then $\mathbf{b} + \mu\,\mathbf{1}$ is still in $F_\alpha(x)$ for any $\mu \in \mathbb{R}$, being that $\mathbf{1} \in \ker(\Lambda)$. This is in a sense the equivalent of adding a constant to a subsolution in the scalar case. We have the following rigidity phenomenon on $\mathcal{A}$.

Proposition 2.11 (see [6, Theorem 5.1]). *The admissible values for critical subsolutions at a given point y in $\mathcal{A}$ are of the form*

$$\mathbf{b} + \mu\,\mathbf{1},$$

where $\mathbf{b} \in \mathbb{R}^M$ depending on y, and $\mu \in \mathbb{R}$.

3. Random setting.

3.1. A family of probability measures. To build up the random frame appropriate for systems, we introduce a family of probability measures defined on $\mathcal{D}$, namely the space of càdlàg paths taking values in $\{1,\dots,\mathrm{M}\}$ endowed with the σ-algebra $\mathcal{F}$; see Appendix B. Averaging with respect to such measures will play a crucial role in the subsequent analysis. We will more precisely show that the coupling matrix Λ induces a correspondence between the simplex $\mathcal{S}$ of probability vectors of $\mathbb{R}^M$ and a simplex of probability measures on $\mathcal{D}$.

It is convenient for later use to start by recalling that the family of cylinders of $\mathcal{F}$, or of $\mathcal{F}_t$ for any $t \geq 0$, is a *semiring* (see Appendix B for the definition of cylinder in $\mathcal{F}$). Namely, it contains the empty set, it is closed by finite intersections, and the difference of two cylinders is a finite disjoint union of cylinders. Therefore, taking into account that $\mathcal{F}$, $\mathcal{F}_t$ are generated by cylinders, we get the approximation theorem for measures; see [14, Theorem 1.65].

PROPOSITION 3.1. *Let μ be a finite measure on $\mathcal{F}$. For any $E \in \mathcal{F}$, there is a sequence E_n of multicylinders (see Terminology* B.1) *in $\mathcal{F}$ with*

$$\lim_n \mu(E_n \triangle E) = 0,$$

where $\triangle$ stands for the symmetric difference. If, in addition, $E \in \mathcal{F}_t$ for some $t \geq 0$, then the approximating multicylinders E_n can be taken in $\mathcal{F}_t$.

As a consequence we see that two finite measures on $\mathcal{D}$ coinciding on the family of cylinders are actually equal.

We go on, as announced, by performing a converse construction, namely by defining through the coupling matrix Λ, for any $\mathbf{a} \in \mathcal{S}$, a suitable function on cylinders and then uniquely extending it to a probability measure on $\mathcal{D}$.

For a probability vector $\mathbf{a} \in \mathbb{R}^M$, we define for any cylinder $\mathcal{C}(t_1,\dots,t_k;j_1,\dots,j_k)$

$$\mu_{\mathbf{a}}(\mathcal{C}(t_1,\dots,t_k;j_1,\dots,j_k)) = \left(\mathbf{a}\, e^{-t_1\Lambda}\right)_{j_1} \prod_{l=2}^{k} \left(e^{-(t_l-t_{l-1})\Lambda}\right)_{j_{l-1}\, j_l}. \tag{3.1}$$

This function enjoys the following key properties:

(i) it is, for any $k \in \mathbb{N}$, a probability measure on the family of cylinders of the form $\mathcal{C}(t_1,\dots,t_k;j_1,\dots,j_k)$ obtained by keeping $(t_1,\dots,t_k)$ fixed and varying $(j_1,\dots,j_k)$ in $\{1,\dots,\mathrm{M}\}^k$, which is actually a σ-algebra being in a one-to-one correspondence with the family of all subsets of $\{1,\dots,\mathrm{M}\}^k$;

(ii) if $(t_{i_1},\dots,t_{i_l})$ is a subsequence of $(t_1,\dots,t_k)$ with $l < k$, then for any $(j^*_{i_1},\dots,j^*_{i_l}) \in \{1,\dots,\mathrm{M}\}^l$

$$\mu_{\mathbf{a}}(\mathcal{C}(t_{i_1},\dots,t_{i_l};j^*_{i_1},\dots,j^*_{i_l}) = \sum_{(j_1,\dots,j_k)\in J} \mu_{\mathbf{a}}(\mathcal{C}(t_1,\dots,t_k;j_1,\dots,j_k)),$$

where

$$J = \{(j_1,\dots,j_k) \mid j_{i_m} = j^*_{i_m} \text{ for } m = 1,\dots,l\}.$$

The latter condition is known as the Kolmogorov consistency condition, and its validity in this context depends upon $e^{-s\Lambda}$ being a stochastic matrix for any s, which is in turn equivalent, as shown in Proposition A.5, to requiring (H4), (H5) on the coupling matrix Λ.

We are then in position to use the Kolmogorov extension theorem; see, for instance, [14, Theorem 14.36], [21, Theorem 1.2], which ensures, under the previous conditions (i), (ii), the existence of a unique probability measure, denoted by $\mathbb{P}_{\mathbf{a}}$, on $(\mathcal{D}, \mathcal{F})$ which extends $\mu_{\mathbf{a}}$ on the whole $\mathcal{F}$.

It comes from (3.1) that the map

$$\mathbf{a} \mapsto \mathbb{P}_{\mathbf{a}} \qquad \text{is linear;}$$

consequently the measures $\mathbb{P}_{\mathbf{a}}$, for $\mathbf{a} = (a_1, \dots, a_m)$ varying among probability vectors of $\mathbb{R}^M$, make up a *simplex of measures* spanned by $\mathbb{P}_i := \mathbb{P}_{\mathbf{e}_i}$, for $i \in \{1, \dots, \mathrm{M}\}$, and

$$\mathbb{P}_{\mathbf{a}} = \sum_{i=1}^{M} a_i \, \mathbb{P}_i.$$

Since by (3.1) the measures $\mathbb{P}_i$ are supported in $\mathcal{D}_i \in \mathcal{F}_0$ (see (B.4) for the definition of $\mathcal{D}_i$), we also deduce that

$$\mathbb{P}_{\mathbf{a}}(A) = \sum_{i=1}^{M} a_i \, \mathbb{P}_i(A \cap \mathcal{D}_i) \qquad \text{for any } A \in \mathcal{F}$$

and

$$a_i = \mathbb{P}_{\mathbf{a}}(\mathcal{D}_i) \qquad \text{for any } i \in \{1, \dots, \mathrm{M}\}.$$

Also notice that all measures $\mathbb{P}_{\mathbf{a}}$ corresponding to strictly positive $\mathbf{a}$ are equivalent in the sense that they have the same null sets, and these are the $E \in \mathcal{F}$ with

$$\mathbb{P}_i(E) = 0 \qquad \text{for any } i.$$

TERMINOLOGY 3.2. *By a* random variable *we mean any measurable map from $(\mathcal{D}, \mathcal{F})$ to a Polish space endowed with the Borel σ-algebra. A* simple random variable *is one that takes on finitely many values. We denote by $\mathbb{E}_{\mathbf{a}}$ the expectation operators relative to $\mathbb{P}_{\mathbf{a}}$ and put for simplicity $\mathbb{E}_i$ in place of $\mathbb{E}_{\mathbf{e}_i}$. We say that some property holds almost surely, a.s. for short, if it is valid up to a $\mathbb{P}_{\mathbf{a}}$-null set for some and consequently for all $\mathbf{a} > 0$, where $>$ must be understood componentwise.*

We consider the push-forward of the probability measure $\mathbb{P}_{\mathbf{a}}$, for any $\mathbf{a} \in \mathcal{S}$, through the flow ϕ_h on $\mathcal{D}$ defined in (B.8). For a cylinder $C := \mathcal{C}(t_1, \dots, t_k; j_1, \dots, j_k)$, we have for any $\mathbf{a} \in \mathcal{S}$

$$\begin{aligned} \phi_h \# \mathbb{P}_{\mathbf{a}}(C) &= \mathbb{P}_{\mathbf{a}}\{\omega \mid \phi_h(\omega) \in C\} = \mathbb{P}_{\mathbf{a}}(\mathcal{C}(t_1 + h, \dots, t_k + h; j_1, \dots, j_k)) \\ &= \left(\mathbf{a}\, e^{-(t_1+h)\Lambda}\right)_{j_1} \prod_{l=2}^{k-1} \left(e^{-(t_l - t_{l-1})\Lambda}\right)_{j_l\, j_{l-1}} = \mathbb{P}_{\mathbf{a}\, e^{-h\Lambda}}(C), \end{aligned}$$

which implies that

$$\phi_h \# \mathbb{P}_{\mathbf{a}}(E) = \mathbb{P}_{\mathbf{a}\, e^{-h\Lambda}}(E) \qquad \text{for any } E \in \mathcal{F}.$$

We have therefore established the following.

PROPOSITION 3.3. *For any $\mathbf{a} \in \mathcal{S}$, $h \geq 0$,*

$$\phi_h \# \mathbb{P}_{\mathbf{a}} = \mathbb{P}_{\mathbf{a}\, e^{-h\Lambda}}.$$

Accordingly, for any measurable function $f : \mathcal{D} \to \mathbb{R}$, we have by the change of variables formula

$$\mathbb{E}_{\mathbf{a}} f(\phi_h) = \int_{\mathcal{D}} f(\phi_h(\omega))\, d\mathbb{P}_{\mathbf{a}} = \int_{\mathcal{D}} f(\omega)\, d\phi_h \# \mathbb{P}_{\mathbf{a}} = \mathbb{E}_{\mathbf{a}\, e^{-\Lambda h}} f. \tag{3.2}$$

We consider, for $t > 0$, the random variables with values in $\{1, \dots, \mathrm{M}\}$ given by the evaluation maps at t, i.e., $\omega \mapsto \omega(t)$. By (3.1),

$$\omega(t)\#\mathbb{P}_{\mathbf{a}}(i) = \mathbb{P}_{\mathbf{a}}(\{\omega \mid \omega(t) = i\}) = \big(\mathbf{a}\, e^{-t\Lambda}\big)_i$$

for any index $i \in \{1, \dots, \mathrm{M}\}$ so that

$$\omega(t)\#\mathbb{P}_{\mathbf{a}} = \mathbf{a}\, e^{-t\Lambda}. \tag{3.3}$$

Consequently, if we look at an M-dimensional vector, say $\mathbf{b}$, as a (measurable) function from $\{1, \dots, \mathrm{M}\}$ to $\mathbb{R}$, we have

$$\mathbb{E}_{\mathbf{a}} b_{\omega(t)} = \mathbf{a}\, e^{-t\Lambda} \cdot \mathbf{b}. \tag{3.4}$$

Formula (3.3) can be partially recovered for measures of the type $\mathbb{P}_{\mathbf{a}} \llcorner E$ ($\mathbb{P}_{\mathbf{a}}$ restricted to E), where E is any set in $\mathcal{F}$.

LEMMA 3.4. *For a given $\mathbf{a} \in \mathcal{S}$, $E \in \mathcal{F}_t$ for some $t \geq 0$, we have*

$$\omega(s)\#(\mathbb{P}_{\mathbf{a}} \llcorner E) = \big(\omega(t)\#(\mathbb{P}_{\mathbf{a}} \llcorner E)\big)\, e^{-(s-t)\Lambda} \quad \textit{for any } s \geq t.$$

Proof. We first assume E to be a cylinder

$$E = \mathcal{C}(t_1, \dots, t_k; j_1, \dots, j_k)$$

for some times and indices. Then the condition $E \in \mathcal{F}_t$ is equivalent to $t \geq t_k$. We have

$$\omega(t_k)\#(\mathbb{P}_{\mathbf{a}} \llcorner E)(i) = \mathbb{P}_{\mathbf{a}}(E \cap \mathcal{C}(t_k; i)),$$

which implies that

$$\omega(t_k)\#(\mathbb{P}_{\mathbf{a}} \llcorner E) = \mathbb{P}_a(E)\, \mathbf{e}_{j_k}$$

and, according to the definition of $\mathbb{P}_{\mathbf{a}}$ in (3.1),

$$\omega(s)\#(\mathbb{P}_{\mathbf{a}} \llcorner E) = \big(\omega(t_k)\#(\mathbb{P}_{\mathbf{a}} \llcorner E)\big)\, e^{-(s-t_k)\Lambda} \quad \text{for } s > t_k.$$

Consequently,

$$\omega(s)\#(\mathbb{P}_{\mathbf{a}} \llcorner E) = \big(\omega(t_k)\#(\mathbb{P}_{\mathbf{a}} \llcorner E)\big)\, e^{-(t-t_k)\Lambda}\, e^{-(s-t)\Lambda} = \big(\omega(t)\#(\mathbb{P}_{\mathbf{a}} \llcorner E)\big)\, e^{-(s-t)\Lambda}$$

for $s \geq t$, as claimed. The result can be extended by linearity to any multicylinder.

Finally, if E is any set in $\mathcal{F}$, then we consider a sequence of multicylinders E_n in $\mathcal{F}_t$ with $\mathbb{P}_{\mathbf{a}}(E_n \triangle E) \to 0$. By Proposition 3.1,

$$\lim_n \omega(s)\#(\mathbb{P}_{\mathbf{a}} \llcorner E_n)(i) = \lim_n \mathbb{P}_{\mathbf{a}}(E_n \cap \mathcal{C}(s; i)) = \mathbb{P}_{\mathbf{a}}(E \cap \mathcal{C}(s; i)) = \omega(s)\#(\mathbb{P}_{\mathbf{a}} \llcorner E)(i).$$

Therefore,

$$\omega(s)\#(\mathbb{P}_{\mathbf{a}} \llcorner E) = \lim_n \omega(s)\#(\mathbb{P}_{\mathbf{a}} \llcorner E_n) = \big(\omega(t)\#(\mathbb{P}_{\mathbf{a}} \llcorner E)\big)\, e^{-(s-t)\Lambda}. \qquad \square$$

3.2. Stopping times. A *stopping time*, adapted to $\mathcal{F}_t$ (see Appendix B), is a nonnegative random variable τ (see Terminology 3.2) satisfying

$$\{\tau \leq t\} \in \mathcal{F}_t \qquad \text{for any } t,$$

which also implies that $\{\tau < t\}$, $\{\tau = t\} \in \mathcal{F}_t$.

For a bounded random variable τ, we set

$$\tau_n = \sum_j \frac{j}{2^n} \, \mathbb{I}(\{\tau \in [(j-1)/2^n, j/2^n)\}), \tag{3.5}$$

where $\mathbb{I}(\cdot)$ stands for the *indicator function* of the set at the argument, namely the function equal to 1 at any element of the set and 0 in the complement. The above sum is finite, τ being bounded, so the τ_n are simple stopping times, and letting n go to infinity we get the following.

Proposition 3.5. *For a bounded stopping time τ, τ_n defined as in* (3.5)*, make up a sequence of simple stopping times with*

$$\tau_n \geq \tau, \quad \tau_n \to \tau \quad \textit{uniformly in } \mathcal{D} \textit{ as } n \to \infty.$$

We consider a simple stopping time of the form

$$\tau = \sum_{j=1}^{l} t_j \, \mathbb{I}(E_j), \tag{3.6}$$

where the sequence $t_1, \ldots, t_l$ is strictly increasing and E_j are mutually disjoint sets of $\mathcal{F}$; in addition, $E_j \in \mathcal{F}_{t_j}$ by the very definition of stopping time. The symbol $\mathbb{I}(\cdot)$ again stands for the indicator function.

We define

$$F_j = \{\tau \geq t_j\}$$

so that

$$F_j \in \mathcal{F}_{t_{j-1}} \qquad \text{for any } j.$$

It is clear that

$$E_j = \bigcap_{i=1}^{j} F_i \setminus F_{j+1}, \qquad F_1 = \mathcal{D},$$

$$F_j = \mathcal{D} \setminus \bigcup_{i=1}^{j-1} E_i \quad \text{for } j > 1, \qquad F_l = E_l. \tag{3.7}$$

We derive that τ can be equivalently expressed as

$$\tau = \sum_{j=1}^{l} (t_j - t_{j-1}) \, \mathbb{I}(F_j), \tag{3.8}$$

where we have set $t_0 = 0$ to simplify notation. The two expressions of τ given by (3.6), (3.8) are different: in (3.6) the sets E_j are mutually disjoint, while in (3.8) they are decreasing with respect to j.

For a stopping time τ, we consider the map defined as

$$\mathbf{a} \mapsto \omega(\tau)\#\mathbb{P}_{\mathbf{a}} \tag{3.9}$$

since the push-forward of $\mathbb{P}_a$ through $\omega(\tau)$ is a probability measure on $\{1, \ldots, \mathrm{M}\}$, which can be identified with an element of $\mathcal{S}$; we see that the relation in (3.9) defines a map from $\mathcal{S}$ to $\mathcal{S}$ which is, in addition, linear. Thanks to Proposition A.2, it can consequently be represented by a stochastic matrix; we will denote analogously to the case of deterministic times (see (3.3)) by $e^{-\Lambda\tau}$, acting on the right. In other terms,

$$\mathbf{a}\, e^{-\tau\Lambda} = \omega(\tau)\#\mathbb{P}_{\mathbf{a}} \qquad \text{for any } \mathbf{a} \in \mathcal{S}. \tag{3.10}$$

3.3. Admissible controls. We call *control* any random variable Ξ taking values in $\mathcal{D}\big(0, +\infty; \mathbb{R}^N\big)$ such that

(i) it is locally (in time) bounded, i.e., for any $t > 0$ there is $R > 0$ with

$$\sup_{[0,t]} |\Xi(t)| < R \qquad \text{a.s.}; \tag{3.11}$$

(ii) it is *nonanticipating*; namely, for any $t > 0$,

$$\omega_1 = \omega_2 \text{ in } [0,t] \ \Rightarrow\ \Xi(\omega_1) = \Xi(\omega_2) \text{ in } [0,t]. \tag{3.12}$$

The second condition can be equivalently rephrased requiring Ξ to be adapted to the filtration $\mathcal{F}_t$, namely requiring that $\Xi(t)$ be $\mathcal{F}_t$-measurable for any t. In fact, if (3.12) holds true, then the value of $\Xi(\omega)(t)$ just depends on the restriction of ω to $[0,t]$, which actually implies that $\Xi(t)$ is $\mathcal{F}_t$-measurable. The converse implication comes from a version of the Doob–Dynkins lemma for Polish spaces (see [13, Lemma 1.13], asserting that if the σ-algebra spanned by a random variable $\#1$ is contained in that spanned by $\#2$, then $\#1$ is a measurable function of $\#2$. In our case, $\#1$ is $\Xi(s)$ for $s \in [0,t]$ and $\#2$ is

$$\omega \mapsto \text{ restriction of } \omega \text{ to } [0,t],$$

which takes value in $\mathcal{D}(0, t; \{1, \ldots, \mathrm{M}\})$.

The paths in $\mathcal{D}\big(0, +\infty; \mathbb{R}^N\big)$ being right continuous, the condition of being adapted implies (see [21, p. 71]) that Ξ is, in addition, *progressively measurable*; namely, for any t, the map

$$(s, \omega) \mapsto \Xi(s, \omega)$$

from $[0,t] \times \mathcal{D}\big(0, +\infty; \{1, \ldots, M\}\big)$ to $\mathbb{R}^N$ is measurable with respect to the σ-algebras $\mathcal{B}[0,t] \times \mathcal{F}_t$ and $\mathcal{B}$, where $\mathcal{B}[0,t]$, $\mathcal{B}$ denote the family of Borel sets of $[0,t]$ and $\mathbb{R}^N$ with respect to the natural topology.

For a control Ξ, $\mathcal{I}(\Xi)$ is also a random variable with values in $\mathcal{C}\big(0, +\infty; \mathbb{T}^N\big)$; in addition, $\mathcal{I}(\Xi)$ is adapted and consequently progressively measurable.

For a time t, we say that a control is *piecewise constant* in $[0,t]$ if it is of the form

$$\sum_{k=1}^{m} X_k\, \mathbb{I}([s_k, s_{k+1})) \qquad \text{in } [0,t)$$

for some $\mathcal{F}_{s_k}$-measurable $\mathbb{R}^N$-valued bounded random variables X_k, where

$$s_k \text{ is an increasing finite sequence with } s_1 = 0,\ s_m = t \tag{3.13}$$

and $\mathbb{I}(\cdot)$ is as usual the indicator function. For any control Ξ and s_k as in (3.13), the $\Xi(s_k)$ are $\mathcal{F}_{s_k}$-measurable $\mathbb{R}^N$-valued bounded random variables for any k, so that

$$\Xi_0 = \begin{cases} \sum_{k=1}^m \Xi(s_k)\, \mathbb{I}([s_k, s_{k+1})) & \text{in } [0,t), \\ \Xi & \text{in } [t, +\infty) \end{cases}$$

is a control piecewise constant in $[0,t]$. We therefore directly derive from Proposition B.3 the following.

Proposition 3.6. *For any control Ξ and $t > 0$, there is a sequence of controls Ξ_n piecewise constant in $[0,t]$ and locally (in time) uniformly bounded with*

$$\Xi_n \to \Xi \qquad \text{in the Skorohod sense in } \mathcal{D}\big(0,+\infty;\mathbb{R}^N\big) \text{ for any } \omega.$$

4. An estimate for subsolutions.

For $\alpha \geq \gamma$, an initial point x in $\mathbb{T}^N$, a bounded stopping time τ, and a control Ξ, we consider in this section the action functional

$$\mathbb{E}_{\mathbf{a}}\left[\int_0^\tau L_{\omega(s)}(x + \mathcal{I}(\Xi)(s), -\Xi(s)) + \alpha\, ds\right]. \tag{4.1}$$

Notice that $I(\Xi)(\tau)$ belongs to $\mathbb{T}^N$ for any ω; see (B.9). The meaning of the sum between elements of $\mathbb{T}^N$ is made precise in Notation 2.1.

We aim at proving the following.

Theorem 4.1. *For $\alpha \geq \gamma$, let $\mathbf{u}$, τ, Ξ, $\mathbf{a}$ be a subsolution to* (HJα)*, a bounded stopping time, a control, and a probability vector in $\mathcal{S}$, respectively. For any initial point $x \in \mathbb{T}^N$, we have*

$$\mathbb{E}_{\mathbf{a}}\big[u_{\omega(0)}(x) - u_{\omega(\tau)}(x + \mathcal{I}(\Xi)(\tau))\big] \leq \mathbb{E}_{\mathbf{a}}\left[\int_0^\tau L_{\omega(s)}(x + \mathcal{I}(\Xi)(s), -\Xi(s)) + \alpha\, ds\right]. \tag{4.2}$$

The difficulty in proving Theorem 4.1 is that the two integrals appearing in (4.2) do not commute due to the presence of the random time τ. It is worthwhile to point out that this difficulty never happens in the study of evolutionary problems for weakly coupled systems (see [18, Proposition 2.5] for more details). Joint measurability properties guarantee that the Fubini theorem can be applied in regions where stopping time is constant. The idea is then to approximate τ by a sequence of simple stopping times τ_n and then exploit the subsolution property of u separately in the regions where τ_n are constant. We will take advantage of some properties about probability measures $\mathbb{P}_{\mathbf{a}}$ we have gathered in section 3.

Throughout the section we put $\alpha = 0$ to ease notation.

Lemma 4.2. *Let $\mathbf{u}$, $\mathbf{a}$ be as in the statement of Theorem* 4.1*; we further consider $t_2 > t_1 \geq 0$, $E \in \mathcal{F}_{t_1}$, $\xi_0 \in \mathcal{D}\big(0,+\infty;\mathbb{R}^N\big)$, and $z_0 \in \mathbb{T}^N$. Then*

$$\begin{aligned} &\int_E \big(u_{\omega(t_1)}(z_0) - u_{\omega(t_2)}(z_0 + \mathcal{I}(\xi_0)(t_2 - t_1))\big)\, d\mathbb{P}_{\mathbf{a}} \\ &\leq \int_E \left(\int_{t_1}^{t_2} L_{\omega(s)}(z_0 + \mathcal{I}(\xi_0)(s - t_1), -\xi_0(s))\, ds\right) d\mathbb{P}_{\mathbf{a}}. \end{aligned}$$

Proof. We can assume $z_0 = 0$ without losing generality in the proof. Since $\mathbf{u}$ is a subsolution to (HJα), we have

$$-\,p \cdot q \leq L_i(z, -q) + H_i(z,p) \leq L_i(z,-q) - \Lambda^i\, \mathbf{u}(z) \tag{4.3}$$

for any $i \in \{1,\dots,\mathrm{M}\}$, $z \in \mathbb{T}^N$, $q \in \mathbb{R}^N$, $p \in \partial u_i(z)$ (see Remark 2.4). We define

$$\mathbf{d} = \omega(t_1)\#(\mathbb{P}_{\mathbf{a}} \llcorner E),$$

and we have for a.e. $s \in (t_1,t_2)$

$$\frac{d}{ds}\left(\left(\mathbf{d}\, e^{-(s-t_1)\Lambda}\right)\cdot \mathbf{u}(\mathcal{I}(\xi_0)(s-t_1))\right)$$
$$= \left(\left(\mathbf{d}\, e^{-(s-t_1)\Lambda}\right)\cdot \left(-\Lambda\,\mathbf{u}(\mathcal{I}(\xi_0)(s-t_1)) + (p^1(s-t_1)\cdot\xi_0(s-t_1),\dots,p^M(s-t_1)\cdot\xi_0(s-t_1))\right)\right),$$

where $p^i(s-t_1)$ is a suitable element in $\partial u_i(\mathcal{I}(\xi_0))(s-t_1)$ for any i. Combining this last equality with (4.3) and setting $\mathbf{L} = (L_1,\dots,L_M)$, we deduce that

$$-\frac{d}{ds}\left(\left(\mathbf{d}\, e^{-(s-t_1)\Lambda}\right)\cdot \mathbf{u}(\mathcal{I}(\xi_0)(s))\right) \le \left(\mathbf{d}\, e^{-(s-t_1)\Lambda}\right)\cdot \mathbf{L}(\mathcal{I}(\xi_0)(s),-\xi_0(s)),$$

and consequently

$$\mathbf{d}\cdot\mathbf{u}(\mathcal{I}(\xi_0)(t_1)) - \mathbf{d}\cdot e^{-(t_2-t_1)\Lambda}\mathbf{u}(\mathcal{I}(\xi_0)(t_2)) = \int_{t_1}^{t_2} -\frac{d}{ds}\left(\left(\mathbf{d}\, e^{-(s-t_1)\Lambda}\right)\cdot \mathbf{u}(\mathcal{I}(\xi_0)(s))\right) ds$$
$$\le \int_{t_1}^{t_2}\left(\mathbf{d}\, e^{-(s-t_1)\Lambda}\right)\cdot(\mathbf{L}(\mathcal{I}(\xi_0)(s),-\xi_0(s))\, ds.$$

We have by the definition of $\mathbf{d}$, (3.4), Lemma 3.4, and $E \in \mathcal{F}_{t_1}$

$$\int_E \left(u_{\omega(t_1)}(\mathcal{I}(\xi_0)(t_1)) - u_{\omega(t_2)}(\mathcal{I}(\xi_0)(t_2))\right) d\mathbb{P}_{\mathbf{a}} = \mathbf{d}\cdot(\mathbf{u}(\mathcal{I}(\xi_0)(t_1) - e^{-(t_2-t_1)\Lambda}\mathbf{u}(\mathcal{I}(\xi_0)(t_2))$$
$$\int_E L_{\omega(s)}(\mathcal{I}(\xi_0)(s),-\xi_0(s)) = \left(\mathbf{d}\, e^{-(s-t_1)\Lambda}\right)\cdot(\mathbf{L}(\mathcal{I}(\xi_0)(s),-\xi_0(s))$$

for any s in $[t_1,t_2]$. By plugging these relations in the last inequality and using the Fubini theorem, we get

$$\int_E \left(u_{\omega(t_1)}(\mathcal{I}(\xi_0)(t_1)) - u_{\omega(t_2)}(\mathcal{I}(\xi_0)(t_2))\right) d\mathbb{P}_{\mathbf{a}} \le \int_E\left(\int_{t_1}^{t_2}(L_\omega(\mathcal{I}(\xi_0),-\xi_0)\, ds\right) d\mathbb{P}_{\mathbf{a}}.$$

∎

LEMMA 4.3. *Given a control Ξ and a bounded stopping time τ, let Ξ_n, τ_n be sequences of controls and bounded stopping times, respectively, with*

(4.4) $$\Xi_n \to \Xi \quad \text{a.s. with respect to the Skorohod metric,}$$
(4.5) $$\tau_n \to \tau \quad \text{uniformly in } \mathcal{D},$$
$$\tau_n \ge \tau \quad \text{a.s. for any } n.$$

Assume, in addition, that for any $T > 0$ there is $R = R(T) > 0$ with

(4.6) $$\sup_{s\in[0,T]} |\Xi_n(s)| < R \quad \text{a.s. for any } n.$$

Then

$$\mathbb{E}_{\mathbf{a}}\left[\int_0^{\tau_n} L_\omega(x+\mathcal{I}(\Xi_n),-\Xi_n)\, ds\right]$$

converges in $\mathbb{R}$ to

$$\mathbb{E}_{\mathbf{a}}\left[\int_0^{\tau} L_\omega(x+\mathcal{I}(\Xi),-\Xi)\, ds\right]$$

and

$$\mathbb{E}_{\mathbf{a}}\big[u_{\omega(0)}(x) - u_{\omega(\tau_n)}(x + \mathcal{I}(\Xi_n)(\tau_n)\big] \to \mathbb{E}_{\mathbf{a}}\big[u_{\omega(0)}(x) - u_{\omega(\tau)}(x + \mathcal{I}(\Xi)(\tau)\big] \tag{4.7}$$

for any $x \in \mathbb{R}^N$, $\mathbf{a} \in \mathcal{S}$.

Proof. We set $x = 0$. We know that conditions (4.4), (4.6) hold true outside a $\mathbb{P}_{\mathbf{a}}$-null set denoted by $\mathcal{D}'$. If $\omega \in \mathcal{D}\setminus\mathcal{D}'$, the $\Xi_n(\omega)$ are uniformly bounded in $[0, \tau(\omega)]$, and we derive from (B.1), (B.5) that $\Xi_n(\omega)$ converges pointwise a.e. in $[0, \tau(\omega)]$ to $\Xi(\omega)$. Also taking into account the continuity of L and $\mathcal{I}(\cdot)$ (see Proposition B.7), we get through the dominated convergence theorem

$$\int_0^\tau L_\omega(\mathcal{I}(\Xi_n), -\Xi_n)\, ds \quad\longrightarrow\quad \int_0^\tau L_\omega(\mathcal{I}(\Xi), -\Xi)\, ds \quad \text{a.s.} \tag{4.8}$$

Let T be such that $\tau \le T$ a.s.; by (4.6),

$$\max_{s\in[0,T]} |\mathcal{I}(\Xi_n)(s)| < RT \qquad \text{for any } n, \text{ outside a } \mathbb{P}_{\mathbf{a}}\text{-null set,}$$

and consequently the sequence

$$\int_0^\tau L_\omega(\mathcal{I}(\Xi_n), -\Xi_n)\, ds$$

is a.s. uniformly bounded. Also taking (4.8) into account, we can thus obtain the claimed convergence with τ in place of τ_n in the approximating sequence, again via the dominated convergence theorem. We further have

$$\left|\int_0^{\tau_n} L_\omega(\mathcal{I}(\Xi_n), -\Xi_n)\, ds - \int_0^\tau L_\omega(\mathcal{I}(\Xi_n), -\Xi_n)\, ds\right| \le \int_\tau^{\tau_n} |L_\omega(\mathcal{I}(\Xi_n), -\Xi_n)|\, ds.$$

Owing to (4.5) and the uniformly boundedness property of the integrand, the right-hand side of the above formula becomes infinitesimal, as n goes to infinity, uniformly in ω so that

$$\mathbb{E}_{\mathbf{a}}\ \left[\left|\int_0^{\tau_n} L_\omega(\mathcal{I}(\Xi_n), -\Xi_n)\, ds - \int_0^\tau L_\omega(\mathcal{I}(\Xi_n), -\Xi_n)\, ds\right|\ \right] \to 0.$$

This shows the first convergence in the statement. The limit relation (4.7) can be proved similarly using the continuity of u in $\mathbb{T}^N$. ◻

LEMMA 4.4. *Assume*

$$\tau = \sum_{j=1}^{l} t_j\, \mathbb{I}(E_j) \tag{4.9}$$

to be a simple stopping time, with the t_j *making up an increasing sequence of times, and set* $F_j = \{\tau \ge t_j\}$ *for any* $j \in \{1, \dots, l\}$. *Let* $\mathbf{u}$, Ξ, $\mathbf{a}$, x *be as in the statement of* Theorem 4.1*; then*

$$\mathbb{E}_{\mathbf{a}}\left[\int_0^\tau L_\omega(x + \mathcal{I}(\Xi), -\Xi) + \alpha\, ds\right] = \sum_{j=1}^{l} \int_{F_j} \left(\int_{t_{j-1}}^{t_j} L_\omega(x + \mathcal{I}(\Xi), -\Xi) + \alpha\, ds\right) d\mathbb{P}_{\mathbf{a}},$$

$$\mathbb{E}_{\mathbf{a}}\big[u_{\omega(0)}(x) - u_{\omega(\tau)}(x + \mathcal{I}(\Xi(\tau)))\big] = \sum_{j=1}^{l} \int_{F_j} \big(u_{\omega(t_{j-1})}(x + \mathcal{I}(\Xi(t_{j-1}))) - u_{\omega(t_j)}(x + \mathcal{I}(\Xi(t_j)))\big)\, d\mathbb{P}_{\mathbf{a}}.$$

Proof. We set $t_0 = 0$ and

$$I = \mathbb{E}_{\mathbf{a}} \left[\int_0^\tau L_\omega(x + \mathcal{I}(\Xi), -\Xi) + \alpha \, ds \right].$$

Taking into account the definition of τ and that the t_i are monotone, we have

$$I = \sum_{i=1}^{l} \int_{E_i} \int_0^{t_i} L_\omega(x + \mathcal{I}(\Xi), -\Xi) + \alpha \, ds = \sum_{i=1}^{l} \sum_{j=1}^{i} \int_{E_i} \int_{t_{j-1}}^{t_j} L_\omega(x + \mathcal{I}(\Xi), -\Xi) + \alpha \, ds$$
$$= \sum_{j=1}^{l} \sum_{i \geq j} \int_{E_i} \int_{t_{j-1}}^{t_j} L_\omega(x + \mathcal{I}(\Xi), -\Xi) + \alpha \, ds$$

and, owing to (3.7),

$$\sum_{i \geq j} \int_{E_i} \int_{t_{j-1}}^{t_j} L_\omega(x + \mathcal{I}(\Xi), -\Xi) + \alpha \, ds = \int_{F_j} \int_{t_{j-1}}^{t_j} L_\omega(x + \mathcal{I}(\Xi), -\Xi) + \alpha \, ds$$

for any $j \in \{1, \ldots, l\}$. Therefore, summing over j we get

$$I = \sum_{j=1}^{l} \int_{F_j} \left(\int_{t_{j-1}}^{t_j} L_\omega(x + \mathcal{I}(\Xi), -\Xi) + \alpha \, ds \right) d\mathbb{P}_{\mathbf{a}},$$

as desired. The second equality in the statement can be proved along the same lines; we provide some detail for the reader's convenience. We start by defining

$$J = \mathbb{E}_{\mathbf{a}} \big[u_{\omega(0)}(x) - u_{\omega(\tau)} x + \mathcal{I}(\Xi(\tau)) \big];$$

then we have

$$J = \sum_{i=1}^{l} \int_{E_i} \big(u_{\omega(0)}(x) - u_{\omega(t_j)}(x + \mathcal{I}(\Xi(t_j))) \big) \, d\mathbb{P}_{\mathbf{a}}$$
$$= \sum_{i=1}^{l} \sum_{j=1}^{i} \int_{E_i} \big(u_{\omega(t_{j-1})}(x + \mathcal{I}(\Xi(t_{j-1})) - u_{\omega(t_j)}(x + \mathcal{I}(\Xi(t_j)))) \big) \, d\mathbb{P}_{\mathbf{a}}$$
$$= \sum_{j=1}^{l} \sum_{i \geq j} \int_{E_i} \big(u_{\omega(t_{j-1})}(x + \mathcal{I}(\Xi(t_{j-1})) - u_{\omega(t_j)}(x + \mathcal{I}(\Xi(t_j)))) \big) \, d\mathbb{P}_{\mathbf{a}}$$

and, again exploiting (3.7),

$$\sum_{i \geq j} \int_{E_i} \big(u_{\omega(t_{j-1})}(x + \mathcal{I}(\Xi(t_{j-1})) - u_{\omega(t_j)}(x + \mathcal{I}(\Xi(t_j)))) \big) \, d\mathbb{P}_{\mathbf{a}}$$
$$= \int_{F_j} \big(u_{\omega(t_{j-1})}(x + \mathcal{I}(\Xi(t_{j-1})) - u_{\omega(t_j)}(x + \mathcal{I}(\Xi(t_j)))) \big) \, d\mathbb{P}_{\mathbf{a}}$$

for any $j \in \{1, \ldots, l\}$. We conclude the proof summing over j. ∎

PROPOSITION 4.5. *The assertion of Theorem* 4.1 *is true if we take the stopping time τ simple, say of the form* (4.9), *and the control Ξ piecewise constant in* $[0, T]$ *for some $T \geq t_l$.*

Proof. Since $T \geq t_l$, we can assume that Ξ has the form

$$\Xi = \sum_{k=1}^{m} X_k \, \mathbb{I}([s_{k-1}, s_k)) \qquad \text{in } [0, t_l),$$

where X_k are $\mathbb{R}^N$-valued random variables and s_k is a finite increasing sequence with $s_0 = 0$ and $s_m = t_l$; we can assume, in addition, that all the times t_j, $j = 1, \dots, l$, belong to the sequence. Consequently, an index j can be univocally associated to any interval $[s_{k-1}, s_k)$ such that $[s_{k-1}, s_k) \subset [t_{j-1}, t_j)$. Due to the nonanticipating character of Ξ,

$$X_k \text{ is } \mathcal{F}_{s_{k-1}}\text{-measurable.} \tag{4.10}$$

We fix indices k, j such that $[s_{k-1}, s_k)$ is contained in $[t_{j-1}, t_j)$; by (4.10), there is a sequence of simple $\mathcal{F}_{s_{k-1}}$-random variables

$$Y_n = \sum_r y_r^n \, \mathbb{I}(B_r^n)$$

taking values in $\mathbb{R}^N$ and converging a.s. to X_k (see [15, Theorem 1.4.4]) with $y_r^n \in \mathbb{R}^N$ and $B_r^n \in \mathcal{F}_{s_{k-1}}$ for any n. Then, slightly modifying the argument in Lemma 4.3, we get that

$$\int_{F_j} \int_{s_{k-1}}^{s_k} L_{\omega(s)}(\mathcal{I}(\Xi(s_{k-1})) + \mathcal{I}(Y_n)(s - s_{k-1}), -Y_n) \, ds \tag{4.11}$$

converges to

$$\int_{F_j} \int_{s_{k-1}}^{s_k} L_{\omega(s)}(\mathcal{I}(\Xi(s_{k-1})) + \mathcal{I}(X_k)(s - s_{k-1}), -X_k) \, ds$$

as n goes to infinity, and similarly

$$\int_{F_j} \big(u_{\omega(s_{k-1})}(\mathcal{I}(\Xi(s_{k-1})) - u_{\omega(s_k)}(\mathcal{I}(\Xi(s_{k-1}) + \mathcal{I}(Y_n)(s_k - s_{k-1}))))\big) \tag{4.12}$$

converges to

$$\int_{F_j} \big(u_{\omega(s_{k-1})}(\mathcal{I}(\Xi(s_{k-1})) - u_{\omega(s_k)}(\mathcal{I}(\Xi(s_{k-1}) + \mathcal{I}(X_k)(s_k - s_{k-1}))))\big).$$

Due to the form of Y_k, the integrals in (4.11), (4.12) can in turn be written as

$$\sum_r \int_{F_j \cap B_n^r} \int_{s_{k-1}}^{s_k} L_{\omega(s)}(\mathcal{I}(\Xi(s_{k-1})) + y_r^n \, (s - s_{k-1}), -y_r^n) \, ds \, d\mathbb{P}_{\mathbf{a}},$$
$$\sum_r \int_{F_j \cap B_n^r} \big\{u_{\omega(s_{k-1})}\big(\mathcal{I}(\Xi(s_{k-1})) - u_{\omega(s_k)}(\mathcal{I}(\Xi(s_{k-1}) + y_r^n \, (s_k - s_{k-1})))\big)\big\} \, d\mathbb{P}_{\mathbf{a}},$$

respectively. Since $F_j \in \mathcal{F}_{t_{j-1}}$, $B_n^r \in \mathcal{F}_{s_{k-1}}$, and $s_{k-1} \geq t_{j-1}$, we deduce that $F_j \cap B_n^r \in \mathcal{F}_{s_{k-1}}$, and we can apply Lemma 4.2 to any term of the previous sum. This yields

$$\begin{aligned}
&\int_{F_j \cap B_n^r} \big(u_{\omega(s_{k-1})}(\mathcal{I}(\Xi(s_{k-1})) - u_{\omega(s_k)}(\mathcal{I}(\Xi(s_{k-1}) + y_r^n \, (s_k - s_{k-1}))))\big) \, d\mathbb{P}_{\mathbf{a}} \\
&\leq \int_{F_j \cap B_n^r} \int_{s_{k-1}}^{s_k} L_{\omega(s)}(\mathcal{I}(\Xi(s_{k-1})) + y_r^n \, (s - s_{k-1}), -y_r^n) \, ds \, d\mathbb{P}_{\mathbf{a}}
\end{aligned}$$

for any r. By summing over r and passing to the limit as n goes to infinity, we further get

$$\int_{F_j} \left(u_{\omega(s_{k-1})}(\mathcal{I}(\Xi(s_{k-1})) - u_{\omega(s_k)}(\mathcal{I}(\Xi(s_{k-1}) + \mathcal{I}(X_k)(s_k - s_{k-1}))))\right) d\mathbb{P}_{\mathbf{a}}$$
$$\leq \int_{F_j} \int_{s_{k-1}}^{s_k} L_{\omega(s)}(\mathcal{I}(\Xi(s_{k-1})) + \mathcal{I}(X_{k-1})(s - s_{k-1}), -X_k)\, ds\, d\mathbb{P}_{\mathbf{a}}.$$

By summing all inequalities as above corresponding to intervals $[s_{k-1}, s_k)$ in $[t_{j-1}, t_j)$, we obtain

$$\int_{F_j} \left(u_{\omega(t_{j-1})}(\mathcal{I}(\Xi(t_{j-1})) - u_{\omega(t_j)}(\mathcal{I}(\Xi(t_j))))\right) d\mathbb{P}_{\mathbf{a}} \leq \int_{F_j} \int_{t_{j-1}}^{t_j} L_{\omega(s)}(\mathcal{I}(\Xi(s)), -\Xi(s))\, ds\, d\mathbb{P}_{\mathbf{a}}.$$

We conclude the proof summing over j and exploiting Lemma 4.4. ∎

Proof of Theorem 4.1. By Proposition 3.5, τ can be approximated uniformly in ω by a sequence of simple stopping times τ_n with $\tau_n \geq \tau$ and $\tau_n \leq T$ for some constant T; in addition, by Proposition 3.6, Ξ can be approximated a.s. with respect to the Skorohod metric by a sequence of control Ξ_n piecewise constant in $[0, T]$ and locally (in time) uniformly bounded.

Owing to Proposition 4.5, inequality (4.2) holds true if we replace τ, Ξ by τ_n, Ξ_n, respectively, for any n. We conclude by passing at the limit as n goes to infinity and exploiting Lemma 4.3. ∎

NOTATION 4.6. *For a bounded stopping time τ and a pair x, y of elements of $\mathbb{T}^N$, we set*

$$\mathcal{K}(\tau, y - x) = \{\Xi \in \mathcal{K} \mid \mathcal{I}(\Xi)(\tau) = y - x \text{ a.s.}\}.$$

Notice that both $I(\Xi)(\tau)$ and $y - x$ are elements of $\mathbb{T}^N$; see (B.9) *and refer to Notation* 2.1 *for the meaning of $y - x$. Also notice that $\mathcal{I}(\Xi)(\tau)$ is a random variable taking value in $\mathbb{R}^N$ because Ξ is progressively measurable and τ is a stopping time. We recall that the diction a.s. must be understood with respect to the family of equivalent measures $\mathbb{P}_{\mathbf{a}}$, $\mathbf{a} > 0$. We will call, with some abuse of language, the controls Ξ belonging to $\mathcal{K}(\tau, 0)$* τ-cycles.

Remark 4.7. For x, y in $\mathbb{T}^N$, the family of controls $\mathcal{K}(\tau, y - x)$ is nonempty whenever $\operatorname{ess\,inf} \tau > 0$. In fact, for such a stopping time, select $\varepsilon > 0$ with $\varepsilon < \operatorname{ess\,inf} \tau$ and define a control Ξ setting for any ω,

$$\Xi(\omega)(s) = \begin{cases} z_0 & \text{for } s \in [0, \varepsilon), \\ 0 & \text{for } s \in [\varepsilon, +\infty), \end{cases}$$

where z_0 is any vector of $\mathbb{R}^N$ with $\operatorname{proj}(\varepsilon\, z_0) = y - x$ (proj is the projection of $\mathbb{R}^N$ onto $\mathbb{T}^N$). It is indeed apparent that Ξ belongs to $\mathcal{K}(\tau, y - x)$.

Using Notation 4.6, we derive from Theorem 4.1 the following.

COROLLARY 4.8. *For any pair of points x, y in $\mathbb{R}^N$, subsolution $\mathbf{u}$ to* (HJα), *$\mathbf{a} \in \mathcal{S}$, bounded stopping time τ, and $\Xi \in \mathcal{K}(\tau, y - x)$, we have*

$$\mathbb{E}_{\mathbf{a}}\left[u_{\omega(0)}(x) - u_{\omega(\tau)}(y)\right] \leq \mathbb{E}_{\mathbf{a}}\left[\int_0^\tau L_{\omega(s)}(x + \mathcal{I}(\Xi)(s), -\Xi(s)) + \alpha\, ds\right]. \tag{4.13}$$

In the next section we will show (see Theorem 5.7) that (4.2) actually characterizes subsolutions to (HJα).

5. A representation formula for subsolutions. Throughout the section we consider a constant α greater than or equal to γ. For y in $\mathbb{R}^N$, $\mathbf{b} \in \mathbb{R}^M$, we define

$$v_i(x) = \inf \mathbb{E}_i \left[\int_0^\tau L_\omega(x + \mathcal{I}(\Xi), -\Xi) + \alpha \, ds + b_{\omega(\tau)} \right] \tag{5.1}$$

for any $i \in \{1, \dots, \mathrm{M}\}$, $x \in \mathbb{R}^N$, where the infimum is taken with respect to any bounded stopping times τ and $\Xi \in \mathcal{K}(\tau, y - x)$. We have the following.

PROPOSITION 5.1. *The function* $\mathbf{v} = (v_1, \dots, v_M)$ *defined in* (5.1) *is bounded in* $\mathbb{T}^N$.

Proof. Taking into account that $\mathbf{1} \in \ker(\Lambda)$, we see that if $b_0 \in F_\alpha(y)$ (see (2.1) for the definition of F_α), then $b_0 + \mu \, \mathbf{1} \in F_\alpha(y)$ as well for any $\mu \in \mathbb{R}$. We can consequently find a subsolution $\mathbf{u}$ to (HJα) with

$$\mathbf{u}(y) \le \mathbf{b}.$$

Owing to Corollary 4.8, we then have

$$\begin{aligned} &\mathbb{E}_i \left[\int_0^\tau L_{\omega(s)}(x + \mathcal{I}(\Xi)(s), -\Xi(s)) + \alpha \, ds + b_{\omega(\tau)} \right] \\ &\qquad \ge \mathbb{E}_i \left[\int_0^\tau L_{\omega(s)}(x + \mathcal{I}(\Xi)(s), -\Xi(s)) + \alpha \, ds + u_{\omega(\tau)}(y) \right] \ge u_i(x) \end{aligned}$$

for any $i \in \{1, \dots, \mathrm{M}\}$, $x \in \mathbb{R}^N$, bounded stopping time τ, and $\Xi \in \mathcal{K}(\tau, y - x)$. This implies that

$$\mathbf{v}(x) \ge \mathbf{u}(x) \qquad \text{for any } x,$$

where $\ge$ must be understood componentwise. On the other side, by setting $\tau \equiv |x - y|$, $\Xi = \frac{x-y}{|x-y|}$ and taking into account that $\mathbf{L}$ is locally bounded, we see that $\mathbf{v}$ is also bounded from above. ∎

We aim at showing the following.

THEOREM 5.2. *The function* $\mathbf{v}$ *defined by* (5.1) *is a subsolution to* (HJα).

We postpone the proof after some preliminary material. The crucial point is to prove a dynamic programming principle–type result. We will use the flow ϕ_h defined in (B.8) in Appendix B and the change of variables formula (3.2).

PROPOSITION 5.3. *Let* h, x, ξ_0, j *be a positive time, a point in* $\mathbb{R}^N$, *a path in* $\mathcal{D}(0, +\infty; \mathbb{R}^N)$, *and an index in* $\{1, \dots, \mathrm{M}\}$, *respectively. Then*

$$v_j(x) \le \mathbb{E}_j \left[\int_0^h L_\omega(x + \mathcal{I}(\xi_0), -\xi_0) + \alpha \, ds + v_{\omega(h)}(x + \mathcal{I}(\xi_0)(h)) \right]. \tag{5.2}$$

Proof. Fix $\varepsilon > 0$, and set $\alpha = 0$, $z = x + \mathcal{I}(\xi_0)(h)$ to ease notation. Denote, for any i, by τ^i, Ξ_i bounded stopping times and controls in $\mathcal{K}(\tau^i, y - z)$ with

$$v_i(z) \ge \mathbb{E}_i \left[\int_0^{\tau^i} L_\omega(z + \mathcal{I}(\Xi_i), -\Xi_i) \, ds + b_{\omega(\tau^i)} \right] - \varepsilon. \tag{5.3}$$

We define new stopping times and controls via

$$\tau = \tau^i, \quad \Xi = \Xi_i \qquad \text{in } \mathcal{D}_i \text{ for any } i;$$

it is clear that $\Xi \in \mathcal{K}(\tau, y - z)$. We set

$$\widetilde{\tau}(\omega) = \tau(\phi_h(\omega)) + h \qquad \text{for any } \omega \in \mathcal{D};$$

this is still a stopping time since, for any $t \geq h$,

$$\{\omega \mid \widetilde{\tau}(\omega) \leq t\} = \{\omega \mid \tau(\phi_h(\omega)) \leq t - h\} = \phi_h^{-1}\big(\{\omega \mid \tau(\omega) \leq t - h\}\big),$$

which actually yields by Proposition B.5

$$\{\omega \mid \widetilde{\tau}(\omega) \leq t\} \in \mathcal{F}_t,$$

as desired. We further set

$$\widetilde{\Xi}(\omega)(s) = \begin{cases} \xi_0(s) & \text{for } \omega \in \mathcal{D},\ s \in [0, h), \\ \Xi(\phi_h(\omega))(s - h) & \text{for } \omega \in \mathcal{D},\ s \in [h, +\infty). \end{cases}$$

To justify $\widetilde{\Xi}$ being an admissible control, we define a map Ψ from $\mathcal{D}\big(0, +\infty; \mathbb{R}^N\big)$ to itself through

$$\Psi(\xi)(s) = \begin{cases} \xi_0(s) & \text{for } s \in [0, h), \\ \xi(s - h) & \text{for } s \in [h, +\infty). \end{cases}$$

According to the very definition of convergence in $\mathcal{D}\big(0, +\infty; \mathbb{R}^N\big)$, this mapping is continuous in the sense of Skorohod; in fact, if $\xi_n \to \xi$ and g_n is the corresponding time scale deformation, then we define

$$\overline{g}_n(s) = \begin{cases} s & \text{for } s \in [0, h), \\ g_n(s - h) + h & \text{for } s \in [h, +\infty), \end{cases}$$

and it is straightforward to check that $\overline{g}_n$ locally uniformly converges to the identity function in $[0, +\infty)$ and $\Psi(\xi)(\overline{g}_n(s))$ locally uniformly converges to $\Psi(\xi)(s)$. We can rephrase the definition of $\widetilde{\Xi}$ above as

$$\widetilde{\Xi}(\omega) = \Psi(\Xi(\phi_h(\omega)),$$

which shows that Ξ is a random variable as composition of continuous and measurable maps. If $\omega_1 = \omega_2$ in $[0, t]$ for some $t > h$, then

$$\phi_h(\omega_1) = \phi_h(\omega_2) \qquad \text{in } [0, t - h],$$

which implies that

$$\Xi(\phi_h(\omega_1)) = \Xi(\phi_h(\omega_2)) \qquad \text{in } [0, t - h];$$

therefore,

$$\begin{aligned} &\widetilde{\Xi}(\omega_1) = \xi_0 = \widetilde{\Xi}(\omega_2) \qquad \text{in } [0, h], \\ &\widetilde{\Xi}(\omega_1(s)) = \Xi(\phi_h(\omega_1))(s - h) = \Xi(\phi_h(\omega_2))(s - h) = \widetilde{\Xi}(\omega_2)(s) \qquad \text{in } [h, t], \end{aligned}$$

which shows that Ξ is nonanticipating. Finally, the uniformly boundedness condition is clearly fulfilled. We conclude that $\widetilde{\Xi}$ is an admissible control. To show that it belongs to $\mathcal{K}(\widetilde{\tau}, y - x)$, we consider for $\omega \in \mathcal{D}$

$$\int_0^{\widetilde{\tau}(\omega)} \widetilde{\Xi}(\omega)\, ds = \int_0^h \xi_0\, ds + \int_h^{\widetilde{\tau}(\omega)} \Xi(\phi_h(\omega))(s - h)\, ds = z - x + \int_0^{\tau(\phi_h(\omega))} \Xi(\phi_h(\omega))(s)\, ds.$$

Owing to $\Xi \in \mathcal{K}(\tau, y-z)$ and Proposition 3.3 we have for any $\mathbf{a} > 0$ in $\mathcal{S}$

$$\mathbb{P}_{\mathbf{a}}\left\{\omega \mid \int_0^{\tau(\phi_h(\omega))} \Xi(\phi_h(\omega))(s)\,ds \neq y-z\right\} = \mathbb{P}_{\mathbf{a}e^{-h\Lambda}}\left\{\omega \mid \int_0^{\tau(\omega)} \Xi(\omega)(s)\,ds \neq y-z\right\} = 0.$$

This establishes that $\widetilde{\Xi} \in \mathcal{K}(\widetilde{\tau}, y-x)$. We compute for $s > 0$

$$\begin{aligned} x + \mathcal{I}(\widetilde{\Xi})(\omega)(s+h) &= x + \int_0^h \xi_0\,dr + \int_h^{s+h} \widetilde{\Xi}(\omega)\,dr \\ &= z + \int_0^s \Xi(\phi_h(\omega))\,dr = z + \mathcal{I}(\phi_h(\omega))(s). \end{aligned} \tag{5.4}$$

According to the very definition of $\mathbf{v}$, we then have

$$\begin{aligned} v_j(x) &\le \mathbb{E}_j\left[\int_0^{\widetilde{\tau}} L_\omega(x + \mathcal{I}(\widetilde{\Xi}), -\widetilde{\Xi})\,ds + b_{\omega(\widetilde{\tau})}\right] \\ &= \mathbb{E}_j\left[\int_0^h L_\omega(x + \mathcal{I}(\xi_0), -\xi_0)\,ds + \int_h^{\widetilde{\tau}} L_\omega(x + \mathcal{I}(\widetilde{\Xi}), -\widetilde{\Xi})\,ds + b_{\omega(\widetilde{\tau})}\right]. \end{aligned} \tag{5.5}$$

Using the definitions of $\widetilde{\tau}$ and $\widetilde{\Xi}$, the change of variables formula (3.2), and (5.4), we have

$$\begin{aligned} &\mathbb{E}_j\left[\int_h^{\widetilde{\tau}} L_\omega(x + \mathcal{I}(\widetilde{\Xi}), -\widetilde{\Xi})\,ds + b_{\omega(\widetilde{\tau}(\omega))}\right] \\ =&\mathbb{E}_j\left[\int_0^{\widetilde{\tau}-h} L_{\omega(s+h)}(x + \mathcal{I}(\widetilde{\Xi})(\omega)(s+h), -\widetilde{\Xi}(\omega)(s+h))\,ds + b_{\omega(\widetilde{\tau}(\omega))}\right] \\ =&\mathbb{E}_j\left[\int_0^{\tau(\phi_h(\omega))} L_{\phi_h(\omega)}(z + \mathcal{I}(\Xi)(\phi_h(\omega)(s), -\Xi(\phi_h(\omega)(s))\,ds + b_{\phi_h(\omega)(\tau(\phi_h(\omega)))}\right] \\ =&\mathbb{E}_{\mathbf{e}_j\,e^{-h\Lambda}}\left[\int_0^{\tau(\omega)} L_\omega(z + \mathcal{I}(\Xi)(\omega)(s), -\Xi(\omega)(s))\,ds + b_{\omega(\tau)}\right]. \end{aligned}$$

Using (5.3), we further get

$$\begin{aligned} &\mathbb{E}_{\mathbf{e}_j\,e^{-h\Lambda}}\left[\int_0^{\tau(\omega)} L_\omega(z + \mathcal{I}(\Xi)(\omega)(s), -\Xi(\omega)(s))\,ds + b_{\omega(\tau)}\right] \\ &= \sum_i \left(\mathbf{e}_j e^{-h\Lambda}\cdot \mathbf{e}_i\right)\mathbb{E}_i\left[\int_0^{\tau^i} L_\omega(z + \mathcal{I}(\Xi_i)(s), -\Xi_i(s))\,ds + b_{\omega(\tau^i)}\right] \\ &\le \sum_i \left(\mathbf{e}_j\,e^{-h\Lambda}\cdot \mathbf{e}_i\right)(v_i(z)+\varepsilon) = \mathbf{e}_j\,e^{-h\Lambda}\cdot \mathbf{v}(z) + \varepsilon = \mathbb{E}_j v_{\omega(h)}(z) + \varepsilon. \end{aligned}$$

Combining the last two computations, we get

$$\mathbb{E}_j\left[\int_h^{\widetilde{\tau}} L_\omega(x + \mathcal{I}(\widetilde{\Xi}), -\widetilde{\Xi})\,ds + b_{\omega(\widetilde{\tau}(\omega))}\right] \le \mathbb{E}_j v_{\omega(h)}(z) + \varepsilon,$$

and recalling (5.5) and the definition of z we finally obtain

$$v_j(x) \le \mathbb{E}_j\left[\int_0^h L_\omega(x + \mathcal{I}(\xi_0), -\xi_0)\,ds + v_{\omega(h)}(x + \mathcal{I}(\xi_0)(h))\right] + \varepsilon.$$

Taking into account that ε is arbitrary and that we have set $\alpha = 0$, in the end we obtain the assertion. ∎

LEMMA 5.4. *The function* $\mathbf{v}$ *defined by* (5.1) *is Lipschitz continuous in* $\mathbb{T}^N$.

Proof. We consider two points $z \neq x$ and set $\tau_0 \equiv |z-x|$, $\Xi_0 = \frac{z-x}{|z-x|} =: q$. Then, according to (5.2),

$$v_i(x) - \mathbf{e}_i\, e^{-|x-z|\Lambda} \cdot \mathbf{v}(z) \leq \mathbb{E}_i \left[\int_0^{|x-z|} L_{\omega(s)}(x + s\,q, -q) + \alpha\, ds \right],$$

from which we derive

$$v_i(x) - v_i(z) + \mathbf{e}_i \left(I - e^{-|x-z|\Lambda} \right) \cdot \mathbf{v}(z) \leq \mathbb{E}_i \left[\int_0^{|x-z|} L_{\omega(s)}(x + s\,q, -q) + \alpha\, ds \right]. \tag{5.6}$$

We take a constant R which is at the same time an upper bound of both $\mathbf{L}(x,q)$ in $\mathbb{T}^N \times B(0,1)$ and $|\mathbf{v}(x)|$ in $\mathbb{T}^N$ (see Proposition 5.1), and in addition a Lipschitz constant of

$$t \mapsto \mathbf{e}_i\, e^{-t\Lambda} \quad \text{in } [0, +\infty)$$

for any i; see Proposition A.6. We deduce from (5.6) that

$$v_i(x) - v_i(z) \leq (R + \alpha + R^2)\, |x-z|.$$

This completes the proof. ∎

Proof of Theorem 5.2. We consider a point $x \in \mathbb{R}^N$ where all components of $\mathbf{v}(x)$ are differentiable and fix a nonvanishing vector $q \in \mathbb{R}^N$; further, we take $\xi_0 \equiv q$, and accordingly

$$x + \mathcal{I}(\xi_0)(s) = x + s\,q \qquad \text{for any } s \geq 0.$$

Formula (5.2) then reads as

$$v_j(x) - \mathbf{e}_j\, e^{-h\Lambda} \cdot \mathbf{v}(x + h\,q) \leq \int_0^h \mathbf{e}_j\, e^{-s\Lambda} \cdot \mathbf{L}(x + s\,q, -q) + \alpha\, ds,$$

which implies that

$$\frac{v_j(x) - \mathbf{e}_j\, e^{-h\Lambda} \cdot \mathbf{v}(x + h\,q)}{h} \leq \frac{1}{h} \int_0^h \mathbf{e}_j\, e^{-s\Lambda} \cdot \mathbf{L}(x + s\,q, -q) + \alpha\, ds.$$

Passing to the limit as h goes to 0, and taking into account that all the v_j are differentiable at x, we get

$$\Lambda^j \cdot \mathbf{v}(x) - Dv_j(x) \cdot q \leq L_j(x, -q) + \alpha.$$

q being arbitrary, we further obtain

$$\Lambda^j \cdot \mathbf{v}(x) + H_j(x, Dv_j(x)) = \Lambda^j\, \mathbf{v}(x) + \sup_q \{-Dv_j(x) \cdot q - L_j(x, -q)\} \leq \alpha.$$

This shows that $\mathbf{v}(x)$ is a.e. and so a viscosity subsolution of the system (HJα). ∎

THEOREM 5.5. *For* $y \in \mathbb{T}^N$, $\mathbf{b} \in F_\alpha(y)$ *if and only if*

$$\mathbb{E}_i \left[\int_0^\tau L_{\omega(s)}(y + \mathcal{I}(\Xi)(s), -\Xi(s)) + \alpha\, ds - b_i + b_{\omega(\tau)} \right] \geq 0 \tag{5.7}$$

for any $i \in \{1, \dots, \mathrm{M}\}$*, bounded stopping times* τ*, and* τ*-cycles* Ξ*.*

Proof. We denote as usual by $\mathbf{v}$ the function defined in (5.1). By taking the stopping time $\tau \equiv 0$ and the control $\Xi \equiv 0$, we see that

$$\mathbf{v}(y) \leq \mathbf{b},$$

where $\leq$ must be understood componentwise. If (5.7) holds, then we also get the converse inequality so that $\mathbf{v}(y) = \mathbf{b}$, which proves $\mathbf{b} \in F_\alpha(y), \mathbf{v}$ being a subsolution to (HJα).

Conversely, if there is a subsolution $\mathbf{u}$ of (HJα) with $u(y) = \mathbf{b}$, then (5.7) is a direct consequence of Corollary 4.8. ∎

We give a characterization of the Aubry set from the Lagrangian point of view.

THEOREM 5.6. *Assume the element* $\mathbf{b}$ *appearing in* (5.1) *to be in* $F_\alpha(y)$*; then*
(i) $\mathbf{v}(y) = \mathbf{b}$;
(ii) $\mathbf{v}$ *is the maximal subsolution to* (HJα) *taking the value* $\mathbf{b}$ *at* y;
(iii) *if* $\alpha = \gamma$ *and* $y \in \mathcal{A}$*, then* $\mathbf{v}$ *is a critical solution.*

Proof. Item (i) has already been proved in Theorem 5.5. If $\mathbf{u}$ is a subsolution to (HJα) with $\mathbf{u}(y) = \mathbf{b}$, then by Corollary 4.8 we get

$$u_i(y) \leq \mathbb{E}_i \left[\int_0^\tau L_{\omega(s)}(x + \mathcal{I}(\Xi)(s), -\Xi(s)) + \alpha\, ds + b_{\omega(\tau)} \right]$$

for any $i \in \{1, \dots, \mathrm{M}\}$, bounded stopping time τ, and τ-cycle Ξ. This shows that

$$\mathbf{v} \geq \mathbf{u}.$$

Item (iii) comes directly from the definition of the Aubry set. ∎

We finish the section by showing that for any $\alpha \geq \gamma$ inequality (4.13) actually characterizes subsolutions to (HJα).

THEOREM 5.7. *A function* $\mathbf{u} : \mathbb{T}^N \to \mathbb{R}^M$ *is a subsolution to* (HJα) *if and only if inequality* (4.13) *holds true for any pair of points* x*,* y *in* $\mathbb{T}^N$*,* $\mathbf{a} \in \mathcal{S}$*, any bounded stopping time* τ*, and* $\Xi \in \mathcal{K}(\tau, y - x)$*.*

In view of Corollary 4.8 , it is enough to show the following.

PROPOSITION 5.8. *If a function* $\mathbf{u} : \mathbb{T}^N \to \mathbb{R}^M$ *satisfies inequality* (4.13) *for any pair of points* x*,* y *in* $\mathbb{T}^N$*,* $\mathbf{a} \in \mathcal{S}$*, any bounded stopping time* τ*, and* $\Xi \in \mathcal{K}(\tau, y - x)$*, then* $\mathbf{u}$ *is a subsolution to* (HJα)*.*

Proof. By using the same argument of Lemma 5.4, we see that $\mathbf{u}$ is Lipschitz continuous. Fix $i \in \{1, \dots, \mathrm{M}\}$, and take a differentiability point y of u_i; define $\mathbf{v}$ as in (5.1) with $\mathbf{u}(y)$ in place of $\mathbf{b}$, and then, owing to Theorem 5.6,

$$\mathbf{v} \geq \mathbf{u} \quad \text{in } \mathbb{T}^N \quad \text{and} \quad \mathbf{v}(y) = \mathbf{u}(y).$$

Hence u_i is subtangent to v_i at y, which implies that $Du_i(y) \in \partial v_i(y)$, and, $\mathbf{v}$ being a subsolution to (HJα), by Theorem 5.2 and Remark 2.4 we get

$$H_i(y, Du_i(y)) + \Lambda^i\, \mathbf{u}(y) = H_i(y, Du_i(y)) + \Lambda^i\, \mathbf{v}(y) \leq \alpha.$$

This concludes the proof. ∎

Appendix A. Stochastic matrices. In this appendix we collect some basic material on stochastic matrices. All matrices appearing below are square matrices. We refer the reader to [16, 20] for the results stated without proof.

We denote by $\mathcal{S} \subset \mathbb{R}^M$ the simplex of *probability vectors* of $\mathbb{R}^M$, namely with nonnegative components summing to 1.

Definition A.1. *A (right) stochastic matrix is a matrix possessing nonnegative entries and with each row summing to* 1*.*

Proposition A.2. *A matrix B is stochastic if and only if*

$$\mathbf{a}\, B \in \mathcal{S} \quad \textit{whenever } \mathbf{a} \in \mathcal{S}. \tag{A.1}$$

Proof. B is stochastic if and only if all its rows are probability vectors or, in other terms, if and only if

$$\mathbf{e}_i \cdot B \in \mathcal{S} \qquad \text{for any } i.$$

This is in turn equivalent to (A.1). ∎

By the Perron–Frobenius theorem for nonnegative matrices, we have the following.

Proposition A.3. *Let B be a stochastic matrix; then its maximal eigenvalue is* 1*, and there is a corresponding left eigenvector in $\mathcal{S}$.*

By the Perron–Frobenius theorem for positive matrices, we have the following.

Proposition A.4. *Let B be a positive stochastic matrix; then its maximal eigenvalue is* 1 *and is simple. In addition, the unique corresponding left eigenvector belonging to $\mathcal{S}$ is positive.*

Even if it is an elementary fact, we give for completeness the proof of the key property that the coupling matrix of the Hamilton–Jacobi system under investigation spans a semigroup of stochastic matrices.

Proposition A.5. *For a matrix A, e^{-tA} is stochastic for any t if and only if* (H4)*,* (H5) *hold with A in place of Λ.*

Proof. Assume that A satisfies (H4), (H5). Then, given $t > 0$, $I - \frac{tA}{n}$ is stochastic for n suitably large, and consequently $\left(I - \frac{tA}{n}\right)^n$ is stochastic because the product of stochastic matrices is still stochastic, and

$$e^{-tA} = \lim_{n \to \infty} \left(I - \frac{tA}{n}\right)^n$$

is stochastic because stochastic matrices make up a compact subset in the space of square matrices. Conversely, if e^{-tA} is stochastic, then the relation

$$A = \lim_{t \to 0} \frac{I - e^{-tA}}{t}$$

implies that A satisfies (H4), (H5). ∎

Proposition A.6. *The function*

$$t \mapsto \mathbf{e}_i\, e^{-t\Lambda}$$

is Lipschitz continuous in $[0, +\infty)$ for any $i \in \{1, \dots, \mathrm{M}\}$.

Proof. We have

$$\frac{d}{dt}\mathbf{e}_i\, e^{-t\Lambda} = -\mathbf{e}_i\,\Lambda\, e^{-t\Lambda},$$

which is bounded in $t \in [0,+\infty)$ because the matrices $e^{-t\Lambda}$, being stochastic, vary in a compact subset of the space of $M \times M$ matrices. ∎

Appendix B. Path spaces. We refer the reader to [2] for more details about this section. The term *càdlàg* corresponds to the French acronym *continue à droite limite à gauche*, namely continuous on the right and with left limit. We consider the space of càdlàg paths defined in $[0,+\infty)$, with value in $\{1,\dots,\mathrm{M}\}$ and $\mathbb{R}^N$, denoted by $\mathcal{D} := \mathcal{D}\big(0,+\infty;\{1,\dots,M\}\big)$ and $\mathcal{D}\big(0,+\infty;\mathbb{R}^N\big)$, respectively. For any $t>0$, we also indicate by $\mathcal{D}(0,t;\{1,\dots,\mathrm{M}\})$ the space of càdlàg paths defined in $[0,t]$ with values in $\{1,\dots,\mathrm{M}\}$. The following can be proved:

(B.1) Any càdlàg path has at most countably many discontinuities.

(B.2) Any càdlàg path is locally (in time) bounded.

Terminology B.1. *To any finite increasing sequence of times $t_1,\dots,t_k$, with $k\in\mathbb{N}$, and indices $j_1,\dots,j_k$ in $\{1,\dots,\mathrm{M}\}$ we associate a (thin)* cylinder *defined as*

$$\mathcal{C}(t_1,\dots,t_k;j_1,\dots,j_k) = \{\omega \mid \omega(t_1)=j_1,\dots,\omega(t_k)=j_k\} \subset \mathcal{D}. \tag{B.3}$$

To ease notation, we set

$$\mathcal{D}_i = \mathcal{C}(0;i) \qquad \text{for any } i \in \{1,\dots,\mathrm{M}\}. \tag{B.4}$$

We call multicylinders *the sets made up by finite unions of mutually disjoint cylinders.*

We endow $\mathcal{D}$ with the σ-algebra $\mathcal{F}$ spanned by cylinders; those of the type $\mathcal{C}(s;j)$ for $s\ge 0$, $j\in\{1,\dots,\mathrm{M}\}$ are indeed enough. A natural related filtration $\mathcal{F}_t$ is obtained by picking, as generating sets, just the cylinders $\mathcal{C}(t_1,\dots,t_k;j_1,\dots,j_k)$ with $t_k\le t$ for any fixed $t\ge 0$.

Using the same construction, mutatis mutandis can be performed in $\mathcal{D}\big(0,+\infty;\mathbb{R}^N\big)$; in this case, the σ-algebra is spanned by cylinders of the type

$$\{\xi \in \mathcal{D}\big(0,+\infty;\mathbb{R}^N\big) \mid \xi(s)\in E\}$$

for s, E varying in $[0,+\infty)$ and in the Borel σ-algebra related to the natural topology of $\mathbb{R}^N$, respectively.

Both $\mathcal{D}$ and $\mathcal{D}\big(0,+\infty;\mathbb{R}^N\big)$ can be endowed with a metric, named after Skorohod, which make them Polish spaces, namely complete and separable, and such that the aforementioned σ-algebras are the corresponding Borel σ-algebras. See [2] for a comprehensive treatment of the topic.

Remark B.2. A consequence of the previous definitions is that $\mathcal{F}$ is the minimal σ-algebra for which the evaluation maps

$$t\mapsto \omega(t), \qquad t\in[0,+\infty),$$

are measurable, and the same holds true for the σ-algebra in $\mathcal{D}\big(0,+\infty;\mathbb{R}^N\big)$ with respect to the evaluation maps

$$\xi \mapsto \xi(t).$$

A map $\Xi : \mathcal{D} \to \mathcal{D}\big(0,+\infty;\mathbb{R}^N\big)$ (resp., $\phi : \mathcal{D}\to\mathcal{D}$) is accordingly measurable if and only if the maps $\omega\mapsto \Xi(\omega)(t)$ from $\mathcal{D}$ to $\mathbb{R}^N$ (resp., $\omega \mapsto \phi(\omega)(t)$ from $\mathcal{D}$ to $\{1,\dots,\mathrm{M}\}$) are measurable for any t.

The convergence induced by the Skorohod metric can be defined, say in $\mathcal{D}\big(0,+\infty;\mathbb{R}^N\big)$ to fix ideas, requiring that there exists a sequence g_n of increasing continuous functions from $[0,+\infty)$ onto itself (then $g_n(0)=0$ for any n) such that

$$g_n(s) \to s \quad \text{uniformly in } [0,+\infty),$$
$$\xi_n(g_n(s)) \to \xi(s) \quad \text{locally uniformly in } [0,+\infty).$$

This is basically locally uniform convergence, up to a uniformly small deformation of the time scale given by the g_n. We infer from the previous definition that

$$\text{(B.5)} \quad \xi_n \to \xi \text{ in the Skorohod sense } \Rightarrow \xi_n(t) \to \xi(t) \text{ at any continuity point of } \xi,$$

which in particular implies that

$$\text{(B.6)} \qquad \xi_n \to \xi \text{ in the Skorohod sense } \Rightarrow \xi_n(0) \to \xi(0).$$

Moreover, we have the following:

(B.7)

Any sequence convergent in the Skorohod sense is locally uniformly bounded.

For $t>0$, we say that a path in $\mathcal{D}\big(0,+\infty;\mathbb{R}^N\big)$ is *piecewise constant* in $[0,t]$ if it is of the form

$$\sum_{k=1}^{l-1} x_k \,\mathbb{I}([s_k,s_{k+1})) \qquad \text{for } s \in [0,t),$$

where $x_k \in \mathbb{R}^N$ and s_k is an increasing sequence of times with $s_1=0$, $s_l=t$. We will use the following approximation result (see [2, section 12, Lemma 3]) in a version slightly accommodated to our needs.

Proposition B.3. *For $t>0$ and $\xi \in \mathcal{D}\big(0,+\infty;\mathbb{R}^N\big)$, let s_k^n, $k=1,\dots,l_n$, be a family of strictly increasing finite sequences with $s_1^n=0$, $s_{l_n}^n=t$, and*

$$\sup_k s_k^n - s_{k-1}^n \to 0 \qquad \textit{as } n \textit{ goes to infinity;}$$

then the sequence of (piecewise constant in $[0,t]$) paths

$$\xi_n = \begin{cases} \sum_k \xi(s_k^n)\,\mathbb{I}([s_{k-1}^n,s_k^n)) & \textit{in } [0,t), \\ \xi & \textit{in } [t,+\infty) \end{cases}$$

converges to ξ in $\mathcal{D}\big(0,+\infty;\mathbb{R}^N\big)$.

For any $h>0$, we consider the shift flow ϕ_h on $\mathcal{D}$ defined by

$$\text{(B.8)} \qquad \phi_h(\omega)(s) = \omega(s+h) \qquad \text{for any } s \in [0,+\infty),\ \omega \in \mathcal{D}.$$

Notice that ϕ_h is not in general continuous since the fact that $\omega_n \to \omega$ in the Skorohod metric does not in general imply that $\phi_h(\omega_n)(0) = \omega_n(h) \to \phi_h(\omega)(0) = \omega(h)$, unless of course h is a continuity point for ω, and hence does not in turn imply, by (B.6), that $\phi_h(\omega_n)$ converges to $\phi_h(\omega)$. However, we directly derive from Remark B.2 the following.

Proposition B.4. *The shift flow $\phi_h : \mathcal{D} \to \mathcal{D}$ is measurable for any $h>0$.*

Proposition B.5. *For nonnegative constants h, t, we have*

$$\phi_h^{-1}(\mathcal{F}_t) \subset \mathcal{F}_{t+h}.$$

Proof. For any $t_1 \geq 0$, $j_1 \in \{1, \dots, \mathrm{M}\}$ we have

$$\phi_h^{-1}(\mathcal{C}(t_1; j_1)) = \mathcal{C}(t_1 + h, j_1).$$

The assertion thus comes from the fact that $\mathcal{F}_t$ is spanned by cylinders of the form $\mathcal{C}(t_1; j_1)$, with $t_1 \leq t$, and in this case $\mathcal{C}(t_1 + h; j_1) \in \mathcal{F}_{t+h}$. ∎

We also consider the space $\mathcal{C}\big(0, +\infty; \mathbb{T}^N\big)$ of continuous paths defined in $[0, +\infty)$ taking values in $\mathbb{T}^N$. It is endowed with a metric giving it the structure of a Polish space, which induces the local uniform convergence.

We define a map

$$\mathcal{I} : \mathcal{D}\big(0, +\infty; \mathbb{R}^N\big) \to \mathcal{C}\big(0, +\infty; \mathbb{T}^N\big)$$

via

$$\mathcal{I}(\xi)(t) = \mathrm{proj}\left(\int_0^t \xi \, ds\right), \tag{B.9}$$

where proj indicates the projection from $\mathbb{R}^N$ onto $\mathbb{T}^N$.

Proposition B.6. *The map $\mathcal{I}(\cdot)$ is continuous.*

Proof. Let us consider a sequence ξ_n in $\mathcal{D}\big(0, +\infty; \mathbb{R}^N\big)$ converging to some ξ; then by (B.7) it is locally (in time) uniformly bounded, and by (B.1), (B.5)

$$\xi_n(s) \to \xi(s) \qquad \text{a.e. in } [0, +\infty).$$

Then, by the dominated convergence theorem and the continuity of proj,

$$\mathcal{I}(\xi_n)(t) \to \mathcal{I}(\xi)(t) \qquad \text{for any } t. \tag{B.10}$$

Furthermore, from the uniformly boundedness of ξ_n and the fact that proj is nonexpansive, we derive that the $\mathcal{I}(\xi_n)$ are locally equi-Lipschitz continuous and locally uniformly bounded. By the Arzelà–Ascoli theorem with (B.10), we get

$$\mathcal{I}(\xi_n) \to \mathcal{I}(\xi) \qquad \text{locally uniformly in time,}$$

as desired. ∎

For $\omega \in \mathcal{D}$, $t > 0$, $x \in \mathbb{R}^N$, we consider the function

$$\xi \mapsto \int_0^t L_{\omega(s)}(x + \mathcal{I}(\xi)(s), -\xi(s)) \, ds \tag{B.11}$$

from $\mathcal{D}\big(0, +\infty; \mathbb{R}^N\big)$ to $\mathbb{R}$.

Proposition B.7. *The function defined in* (B.11) *is continuous.*

Proof. Let ξ_n be a sequence converging to some ξ in $\mathcal{D}\big(0, +\infty; \mathbb{R}^N\big)$; then the ξ_n are uniformly bounded in $[0, t]$ and converge pointwise to ξ a.e. by (B.1), (B.5), (B.7). Furthermore, bearing in mind Proposition B.6, we know that $\mathcal{I}(\xi_n)$ converges to $\mathcal{I}(\xi)$ in $\mathcal{C}\big(0, +\infty; \mathbb{T}^N\big)$. Using the continuity of L_i, for any i, we derive that

$$L_{\omega(s)}(x + \mathcal{I}(\xi_n), -\xi_n) \to L_{\omega(s)}(x + \mathcal{I}(\xi), -\xi) \quad \text{a.e. in } [0, t]$$

and, in addition, that the $L_{\omega(s)}(x + \mathcal{I}(\xi_n), -\xi_n)$ are uniformly bounded. We thus get the assertion through the dominated convergence theorem. ∎

REFERENCES

[1] P. Bernard, *Existence of $C^{1,1}$ critical sub-solutions of the Hamilton-Jacobi equation on compact manifolds*, Ann. Sci. Éc. Norm. Supér. (4), 40 (2007), pp. 445–452, doi:10.1016/j.ansens.2007.01.004.

[2] P. Billingsley, *Convergence of Probability Measures*, John Wiley, New York, 1999, doi:10.1002/9780470316962.

[3] F. Cagnetti, D. A. Gomes, and H. V. Tran, *Adjoint methods for obstacle problems and weakly coupled systems of PDE*, ESAIM Control Optim. Calc. Var., 19 (2013), pp. 754–779, doi:10.1051/cocv/2012032.

[4] 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), pp. 719–749, doi:10.1007/s00030-011-0149-7.

[5] A. Davini and A. Siconolfi, *A generalized dynamical approach to the large time behavior of solutions of Hamilton–Jacobi equations*, SIAM J. Math. Anal., 38 (2006), pp. 478–502, doi:10.1137/050621955.

[6] A. Davini and M. Zavidovique, *Aubry sets for weakly coupled systems of Hamilton–Jacobi equations*, SIAM J. Math. Anal., 46 (2014), pp. 3361–3389, doi:10.1137/120899960.

[7] A. Fathi, *Weak KAM Theorem in Lagrangian Dynamics*, ghost book, http://www.crm.sns.it/media/person/1235/fathi.pdf.

[8] A. Fathi and A. Siconolfi, *Existence of C^1 critical subsolutions of the Hamilton-Jacobi equation*, Invent. Math., 155 (2004), pp. 363–388, doi:10.1007/s00222-003-0323-6.

[9] A. Fathi and A. Siconolfi, *PDE aspects of Aubry-Mather theory for quasiconvex Hamiltonians*, Calc. Var. Partial Differential Equations, 22 (2005), pp. 185–228, doi:10.1007/s00526-004-0271-z.

[10] D. A. Gomes and A. Serra, *Systems of weakly coupled Hamilton-Jacobi equations with implicit obstacles*, Canad. J. Math., 64 (2012), pp. 1289–1309, doi:10.4153/cjm-2011-085-7.

[11] H. Ishii, *Asymptotic solutions for large time of Hamilton-Jacobi equations in Euclidean n space*, Ann. Inst. H. Poincaré Anal. Non Linéaire, 25 (2008), pp. 231–266, doi:10.1016/j.anihpc.2006.09.002.

[12] H. Ishii and S. Koike, *Viscosity solutions for monotone systems of second-order elliptic PDEs*, Comm. Partial Differential Equations, 16 (1991), pp. 1095–1128.

[13] O. Kallenberg, *Foundations of Modern Probability*, Springer, New York, 2002, doi:10.1007/978-1-4757-4015-8.

[14] A. Klenke, *Probability Theory*, Springer, Berlin, 2008, doi:10.1007/978-1-84800-048-3.

[15] K. Kuttler, *Notes for Seminar*, Mimeo, www.math.byu.edu/~klkuttle/, 2013.

[16] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*, SIAM, Philadelphia, 2000.

[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), pp. 43–70.

[18] H. Mitake and H. V. Tran, *A dynamical approach to the large-time behavior of solutions to weakly coupled systems of Hamilton–Jacobi equations*, J. Math. Pures Appl. (9), 101 (2014), pp. 76–93.

[19] 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), pp. 733–769, doi:10.1007/s00205-013-0685-x.

[20] J. R. Norris, *Markov Chains*, Cambridge University Press, Cambridge, UK, 1997, doi:10.1017/CBO9780511810633.

[21] J. Swart And A. Winter, *Markov Processes: Theory and Examples*, Mimeo, https://www.uni-due.de/~hm0110/Markovprocesses/sw20.pdf, 2013.