Nonlinear Differ. Equ. Appl. 21 (2014), 167–217

1021-9722/14/020167-51
published online August 1, 2013
DOI 10.1007/s00030-013-0243-0

Nonlinear Differential Equations and Applications NoDEA

The Mather problem for lower semicontinuous Lagrangians

Diogo A. Gomes and Gabriele Terrone

Abstract. In this paper we develop the Aubry-Mather theory for Lagrangians in which the potential energy can be discontinuous. Namely we assume that the Lagrangian is lower semicontinuous in the state variable, piecewise smooth with a (smooth) discontinuity surface, as well as coercive and convex in the velocity. We establish existence of Mather measures, various approximation results, partial regularity of viscosity solutions away from the singularity, invariance by the Euler–Lagrange flow away from the singular set, and further jump conditions that correspond to conservation of energy and tangential momentum across the discontinuity.

Mathematics Subject Classification (2010). Primary 35F21, 37J50; Secondary 35D40, 35R05, 49K20.

Keywords. Discontinuous Lagrangians, Action minimizing measures, Hamilton–Jacobi equations, Viscosity solutions.

D. Gomes was partially supported by CAMGSD-LARSys through FCT-Portugal and by grants PTDC/MAT/114397/2009, UTAustin/MAT/0057/2008, and UTA-CMU/MAT/0007/2009.
G. Terrone was supported by the UTAustin-Portugal partnership through the FCT post-doctoral fellowship SFRH/BPD/40338/2007, CAMGSD-LARSys through FCT Program POCTI - FEDER and by grants PTDC/MAT/114397/2009, UTAustin/MAT/0057/2008, and UTA-CMU/MAT/0007/2009.

List of symbols

$\{e_i\}_{i=1}^N$	The canonical basis of $\mathbb{R}^N$
x_i	The i-th component of a vector $x \in \mathbb{R}^N$
$\lvert x\rvert$	The norm of a vector $x \in \mathbb{R}^N$
$[x, y]$	The line segment $\{tx + (1-t)y,\, t \in [0,1]\}, x, y \in \mathbb{R}^N$
$\frac{\partial}{\partial x_i}\psi = \partial_{x_i}\psi$	The partial derivative of the function ψ with respect to the variable x_i
$D_x\psi$	Gradient of the function ψ with respect to x, that is $(\partial_{x_1}\psi, \ldots, \partial_{x_N}\psi)$
$\dot{\mathbf{x}}(t), \ddot{\mathbf{x}}(t)$	The first and second derivative of a function $\mathbf{x} : I \subset \mathbb{R} \to \mathbb{R}^N$
$\lvert\psi\rvert_\infty$	The L^∞-norm of a function ψ
$\bar{\Omega}$	Closure of an open set $\Omega \subset \mathbb{R}^N$
$\partial\Omega$	Boundary of an open set $\Omega \subset \mathbb{R}^N$
$B(x, r)$	The Euclidean ball in $\mathbb{R}^N$ of radius $r > 0$ around x
Ω_δ	For any open set $\Omega \subset \mathbb{R}^N$, any $\delta > 0$, the set $\{x \in \mathbb{R}^N \mid \mathrm{dist}(x, \Omega) < \delta\}$
$T_x M$	The tangent space of a smooth manifold M at the point $x \in M$
δ_{x_0}	The Dirac mass concentrated at x_0
$\mu_X(x)$	The projection on X of a measure $\mu(x, y)$ on $X \times Y$
$[\![r(t)]\!]_{t_0}$	The jump of the function $r : I \subset \mathbb{R} \to \mathbb{R}$ at t_0, that is $\lim_{t\to t_0^+} r(t) - \lim_{t\to t_0^-} r(t)$

1. Introduction

In this paper we develop the Aubry-Mather theory for Lagrangians that are discontinuous with respect to the state variable. The motivation for such problem comes from the study of singular mechanical systems in which the potential energy is discontinuous. In addition to discontinuities in the potential, low regularity Lagrangians arise in fluid mechanic problems, see for instance [16] for properties of minimizers of mechanical Lagrangians with Sobolev potentials.

A natural question one can pose when faced with a potential energy that is discontinuous is whether some of the results form Aubry-Mather theory can be established. We are interested in existence and invariance under the Euler–Lagrange flow of probability measures μ minimizing the average action

$$\mathbb{A}(\mu) := \int_{\mathbb{T}^N \times \mathbb{R}^N} L(x, v)\, \mathrm{d}\mu(x, v)$$

under the so called *holonomy constraint*:

$$\int_{\mathbb{T}^N \times \mathbb{R}^N} v\, D\varphi(x)\, \mathrm{d}\mu(x, v) = 0, \quad \text{for all } \varphi \in C^1(\mathbb{T}^N). \tag{1.1}$$

The original formulation of the problem, just as it was stated by John Mather in [19] for smooth Lagrangians, consisted in minimizing $\mathbb{A}(\mu)$ among all probability measures invariant under the Euler–Lagrange flow. Afterwards, Mañé in [18] observed that it is more convenient to consider the class of measures satisfying the constraint (1.1). The two minimization problems are actually equivalent for a wide class of Lagrangians, since any minimizing holonomic measure is invariant under the Euler–Lagrange flow.

In this paper we consider a Lagrangian $L(x, v) : \mathbb{T}^N \times \mathbb{R}^N \to \mathbb{R}$ lower semicontinuous in x, convex in the velocity v and coercive. We are interested in a situation where there exists a $(N-1)$-dimensional closed smooth and orientable hypersurface $\Sigma \subset \mathbb{T}^N$ where $L(\cdot, v)$ is discontinuous. Most of the results of the paper work for Lagrangians satisfying very general assumptions; however, to present ideas and techniques in the simplest possible setting, some of them are proved for a model problem in which

$$L(x, v) = L_0(x, v) + V(x),$$

with L_0 continuous in both variables, coercive and convex in v, and V lower semicontinuous and bounded.

In order to understand the behavior of minimizing measures in the discontinuous setting, we study first optimality conditions for trajectories minimizing the Lagrangian action

$$A(x) := \int_0^T L(\mathbf{x}(t), \dot{\mathbf{x}}(t))\ \mathrm{d}t.$$

We relate to [10] and references therein for the study of minimizing trajectories when L is not continuous. In analogy with the case of non-convex Lagrangians, piecewise smooth curves that minimize the action solve the Euler–Lagrange equations away from Σ. In addition, we show that minimizing trajectories must satisfy certain *jump conditions* across the discontinuity set.

We establish existence of Mather measures and various results concerning their approximation. Then, in order to study invariance under the Euler–Lagrange flow of such measures, we focus on the associated Hamilton–Jacobi equation, that is

$$H(x, Du) = \bar{H}, \quad x \in \mathbb{T}^N, \tag{1.2}$$

where $H(x, p)$ is the Legendre transform of L, and $\bar{H}$ is the minimal value of the average action $\mathbb{A}$. We prove existence in $\mathbb{T}^N$ and partial regularity away from the discontinuity set of viscosity solutions of (1.2). Since H is not continuous, a suitable notion of viscosity solution has been taken into account.

Global invariance of minimizing Mather measures under the Euler–Lagrange flow is clearly not possible, as the Euler–Lagrange equations are not defined where the Lagrangian is discontinuous; then only local properties can be expected. Besides local invariance away from the singular set, we proved several *jump conditions* on Σ. These properties rephrase, in terms of Mather measures, the optimality conditions we established for optimal trajectories, namely conservation of energy and conservation of tangential momentum across the discontinuity surface.

Even though the general approach and the proofs follow the same lines as in the continuous case (see, for instance, the survey paper [4]), the results presented in this paper are not a mere adaptation of those available in the literature for the continuous Lagrangians. The lack of continuity indeed required the development of several nontrivial ad hoc arguments at every stage of the work.

For example, techniques in Aubry-Mather theory are valid for smooth enough Lagrangians (typically C^3) and make use of viscosity solutions of Hamilton–Jacobi equations. In contrast, in our setting we need to work with discontinuous Hamiltonians, which creates several challenges in adapting standard arguments. A key tool in our analysis is an approximation procedure based on the use of inf-convolution.

In the continuous case, viscosity solution theory and Aubry-Mather theory revealed to share many aspects as pointed out, for example, in the papers [14,15] by Fathi and Siconolfi. Also in the discontinuous setting considered in this paper, viscosity solutions turn out to be an essential tool in the analysis of minimizing measures. Viscosity solutions of Hamilton–Jacobi equations with discontinuous Hamiltonians have been studied extensively by many authors, in different settings; we refer to the books by Barles [2] and Bardi and Capuzzo-Dolcetta [1] for a general treatment. They have been used in the analysis of geodesic distances and in the study of some discontinuous control problems, combustion phenomena in nonhomogeneous media, and geometric optic propagation in the presence of layers; see [6,20,22,23]. Measurable Hamiltonians have been considered in [7–9,11]. The notion of viscosity solution has been also adapted in a recent paper by Barles et al. [3] to study Bellman equations related to deterministic control problems in which dynamics and costs are different in complementary domains, and consequently discontinuities may arise at the boundary of these domains. However, to the best of our knowledge, this is the first time that viscosity solutions to discontinuous Hamilton–Jacobi equations are applied to the study of Mather measures.

To study regularity properties of viscosity solutions we adapt to the present context some techniques of [13]. We also developed a simple calculus of variation argument that permits to define a class of variations of an holonomic measure that preserves the constraint (1.1). This type of variations were introduced in [4] to prove invariance of minimizing measure, in the case in which L is smooth. The same, argument adapted to the present discontinuous setting, permits to establish local invariance of Mather measures outside the discontinuity surface.

The paper is organized as follows. After describing the main assumptions in Sect. 2, we outline the main results of the paper in Sect. 3. Our analysis begins in Sect. 4 with a brief discussion of necessary conditions for trajectories minimizing the Lagrangian action. In Sect. 5 we prove existence of Mather measures and some duality results, we discuss a one-dimensional example and introduce holonomy preserving variations. In subsequent Sect. 6 various approximation results are presented. Existence, main properties and partial regularity of viscosity solutions are investigated in Sect. 7. Finally, in Sect. 8, after discussing some examples, we establish further properties of Mather measures and main results of the paper. More precisely, local invariance away from Σ is proved in Sect. 8.2 whereas properties and jump conditions for minimizing measures across the discontinuity surface are studied in Sect. 8.3.

2. Setting and assumptions

We consider a Lagrangian $L : \mathbb{T}^N \times \mathbb{R}^N \to \mathbb{R}$, where $\mathbb{T}^N$ is the N-dimensional torus, which satisfy the following hypotheses.
Coercivity. For every fixed $x \in \mathbb{T}^N$ the mapping $v \mapsto L(x, v)$ is continuous, uniformly convex and coercive, that is

$$\lim_{|v|\to+\infty} \frac{L(x, v)}{|v|} = +\infty. \tag{H1}$$

We also assume that L is bounded by below, say $L \geq 1$, which from the coercivity entails no loss of generality.
Lower semicontinuity. We do not assume continuity in x. However, in order to use standard methods in the calculus of variations, we suppose that, in addition to continuity in v for fixed x we have that

$$\begin{aligned}&L(\cdot, v) \text{ is lower semicontinuous, that is,}\\ &\text{for any } (x, v) \in \mathbb{T}^N \times \mathbb{R}^N \quad \liminf_{(y,w)\to(x,v)} L(y, w) \geq L(x, v).\end{aligned} \tag{2.1}$$

More precisely, we are interested in a situation where there exists a $(N-1)$-dimensional closed smooth and orientable hypersurface $\Sigma \subset \mathbb{T}^N$, which we call the *discontinuity locus*, where $L(\cdot, v)$ is discontinuous. Locally, around the hypersurface $\Sigma, \mathbb{T}^N$ is divided into two connected subsets Ω^+ and Ω^-, that is, there exists a neighborhood U of Σ such that $U = \Omega^+ \cup \Sigma \cup \Omega^-$. Notice that we are not assuming Σ connected, thus the discontinuity set could be also the finite union of closed smooth disjoint hypersurfaces.

Since Σ is smooth, for any $x \in \Sigma$ there exists $\nu(x)$, the unit normal to Σ at x, that we agree to point towards Ω^-. We assume then the following:

(H2)

1. for any $v \in \mathbb{R}^N$ (2.1) holds;
2. the mapping $(x, v) \mapsto L(x, v)$ is continuous and smooth in $\Omega^+ \times \mathbb{R}^N$ and $\Omega^- \times \mathbb{R}^N$;
3. the mapping $(x, v) \mapsto L(x, v)$ is continuous in $(\Omega^- \cup \Sigma) \times \mathbb{R}^N$ and can be extended by continuity from $\Omega^+ \times \mathbb{R}^N$ to $(\Omega^+ \cup \Sigma) \times \mathbb{R}^N$, since L is not continuous, these extensions will not agree on Σ;
4. the mappings $(x, v) \mapsto D_x L(x, v)$, $(x, v) \mapsto D_v L(x, v)$ are continuous in $\Omega^- \times \mathbb{R}^N$ and $\Omega^+ \times \mathbb{R}^N$, and can be extended by continuity to $(\Omega^- \cup \Sigma) \times \mathbb{R}^N$ and $(\Omega^+ \cup \Sigma) \times \mathbb{R}^N$, but, of course, these extensions may not agree on Σ;
5. for any $x \in \Sigma$ and $v \in \mathbb{R}^N$ there exist
$$\lim_{\substack{(y,w)\to(x,v)\\ y\in\Omega^+}} L(y, w) =: L^+(x, v), \qquad \lim_{\substack{(y,w)\to(x,v)\\ y\in\Omega^-}} L(y, w) =: L^-(x, v),$$
and $L^+(x, v) \geq L^-(x, v)$.

Remark 2.1. In consequence of the previous assumptions it follows that the values L are also comparable in the neighborhood U of Σ:

$$\text{there exists } C > 0 \text{ such that for any } x \in \Omega^+ \cap U, \text{ and any } y \in \Omega^- \cap U,$$
$$L(y, v) \leq L(x, v) + C|x - y|.$$

□

We now introduce the Hamiltonian $H : \mathbb{T}^N \times \mathbb{R}^N \to \mathbb{R}$ associated to L through the Legendre–Fenchel transform:

$$H(x, p) := \sup_{v \in \mathbb{R}^N} \{p \cdot v - L(x, v)\}.$$

Note that $H(\cdot, p)$ is upper semicontinuous. Furthermore, by property 5 in (**H2**), the values H^+ and H^- defined analogously as L^+ and L^- respectively satisfy

$$H^+(x, p) \leq H^-(x, p) \quad \text{for any } x \in \Sigma \text{ and } p \in \mathbb{R}^N.$$

Model problem. Results in Sects. 4 and 5 are valid under the general assumptions (**H1**)–(**H2**). In Sects. 6, 7 and 8, to simplify the setting, we will work with a model problem where

$$L(x, v) = L_0(x, v) + V(x), \tag{H3}$$

with L_0 continuous in both variables, coercive and convex in v, and V lower semicontinuous and bounded.

Lagrangians of this type arise, in classical mechanics, for the study of the motion of a particle with kinetic energy $L_0(x, v) = k(v)$ and potential energy $V(x)$. A particular case is

$$L(x, v) = \frac{|v|^2}{2} + V(x),$$

with V as before.

Most of the results of the paper work under more general Lagrangians satisfying (**H1**)–(**H2**) with additional technical hypothesis. However, working under a general setting would make the exposition harder to follow but the key techniques of the paper would not change.

Remark 2.2. For L satisfying (**H3**) the corresponding Hamiltonian is

$$H(x, p) = H_0(x, p) - V(x), \tag{2.2}$$

where H_0 is the Legendre transform of L_0. In particular, if $L_0(x, v) = \frac{|v|^2}{2}$, then

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

□

3. Outline of main results

We describe next the main results of the paper.
Necessary conditions for optimal trajectories. In Sect. 4 we discuss the main properties of minimizing trajectories of discontinuous Lagrangians. We use a classical calculus of variation approach to determine first order optimality conditions. More precisely, we show that if a trajectory $\mathbf{x} : [0, T] \to \mathbb{T}^N$ minimizes the action functional

$$A(\mathbf{x}) := \int_0^T L(\mathbf{x}(t), \dot{\mathbf{x}}(t)) \, \mathrm{d}t$$

then:

1. if $\mathbf{x}(t)$ *crosses* the discontinuity locus Σ at $t = t_0$ then the tangential component of the momentum is conserved, that is

$$\text{for any } \xi \in T_{\mathbf{x}(t_0)}\Sigma, \quad [\![\xi \cdot D_v L(\mathbf{x}(t), \dot{\mathbf{x}}(t))]\!]_{t_0} = 0$$

(see Proposition 4.3). The same condition holds when $\mathbf{x}(t)$ *enters* or *exits* Σ at $t = t_0$ (see Proposition 4.6);
2. if $\mathbf{x}(t) \in \Sigma$ for $t \in (t_1, t_2)$ $(t_1 < t_2)$ then the tangential component and a unilateral condition for the normal component of the Euler–Lagrange equation are satisfied (see Proposition 4.4);
3. the energy $E_{\mathbf{x}}(t) = \dot{\mathbf{x}}(t) \cdot D_v L(\mathbf{x}(t), \dot{\mathbf{x}}(t)) + L(\mathbf{x}(t), \dot{\mathbf{x}}(t))$ of $\mathbf{x}$ is conserved on $[0, T]$ (see Proposition 4.7).

Mather measures. In Sect. 5 we prove existence of holonomic probability measures minimizing

$$\int L(x, v) \, \mathrm{d}\mu. \tag{3.1}$$

This easily follows by the lower semicontinuity of the Lagrangian (see Proposition 5.2); furthermore, as in the continuous case, minimizing holonomic measures are supported on a graph of a measurable function $\mathbf{v}(x)$ (see Proposition 5.5). For continuous Lagrangians, $\mathbf{v}$ turns out to be Lipschitz (see [4]); in the discontinuous setting we will establish partial regularity results in Sects. 7 and 8.

In Sect. 5 we also define the value of the Mather problem:

$$\inf \left\{ \int L(x, v) \, \mathrm{d}\mu \,\middle|\, \mu \text{ holonomic prob. measure} \right\} =: -\bar{H},$$

and in Corollary 5.11 we prove, using a duality argument, that

$$\bar{H} := \inf_{\varphi \in C^1(\mathbb{T}^N)} \sup_{x \in \mathbb{T}^N} H(x, D\varphi(x)),$$

where H is the Legendre transform of L.
Approximation. We start considering in Sect. 6 the model problem

$$L(x, v) = L_0(x, v) + V(x)$$

described in the previous Section, and we prove the following result (Theorem 6.1): let V^ϵ be the inf-convolution of V, and μ^ϵ an holonomic probability measure minimizing (3.1), with L replaced by $L^\epsilon(x,v) = L_0(x,v) + V^\epsilon(x)$. If μ^ϵ converges weak-star to some $\bar{\mu}$ as $\epsilon \to 0$, then $\bar{\mu}$ minimizes (3.1).

Viscosity solutions. In order to establish properties for the minimizing measure, we study in Sect. 7 the Hamilton–Jacobi equation with Hamiltonian given by the Legendre transform of L. For the model problem we are considering, it is given by

$$H(x,p) = H_0(x,p) - V(x),$$

where H_0 is the Legendre transform of L_0. Forced by the discontinuous setting, we need to rely on a notion of viscosity solution for discontinuous Hamiltonians. Our main results are the following: using an approximation argument, we prove in Proposition 7.4 that there exists a Lipschitz continuous viscosity solution u to the equation

$$H(x, Du) = \bar{H}, \quad x \in \mathbb{T}^N. \tag{3.2}$$

To prove uniqueness of the value λ such that the equation $H(x, Du) = \lambda$ has a viscosity solution in $\mathbb{T}^N$ we invoke a comparison argument due to Soner (see [21]).

Concerning the regularity of u, solution to (3.2), in Proposition 7.9 we show that $Du(x)$ exists for $\mu_{\mathbb{T}^N}$-a.e. $x \in \operatorname{supp}(\bar{\mu})\backslash\Sigma$, and that the identity $p = Du(x)$ holds for any $\bar{\mu}$-a.e. $(x,p) \in (\operatorname{supp}(\bar{\mu})\backslash\Sigma) \times \mathbb{R}^N$; here $\bar{\mu}$ is a minimizing measure for (3.1). Then, in Proposition 7.9, we prove an $L^2_{\mathrm{loc}}(\mathrm{d}\bar{\mu}_{\mathbb{T}^N})$ estimate for the gradient of u, that is:

$$\int_{\mathbb{T}^N} \varphi^2(x) |D_x u(x+h) - D_x u(x)|^2 \, \mathrm{d}\bar{\mu}_{\mathbb{T}^N} \leq C|h|^2,$$

for any cutoff φ compactly supported away from Σ.

Properties of minimizing measures. Properties for Mather measures are studied in Sect. 8. More precisely, in Sect. 8.2 we investigate the behavior of minimizing measure *away from the discontinuity set*. Using the L^2_{loc} estimates on Du above mentioned, we can prove that $v^\epsilon := \mathbf{v} * \eta_\epsilon$ (where η_ϵ is a standard mollification kernel) converges to $\mathbf{v}$ locally in L^2, away from Σ. This is actually enough to establish, in Theorem 8.4, invariance of a minimizing measure $\bar{\mu}$ with respect to the Euler–Lagrange flow χ outside Σ:

$$\int \chi \cdot D\phi \, \mathrm{d}\bar{\mu} = 0,$$

for any $\phi \in C^1(\mathbb{T}^N \times \mathbb{R}^N)$, compactly supported in $(\mathbb{T}^N\backslash\Sigma) \times \mathbb{R}^N$.

An example in Sect. 8.1 shows that in general we cannot expect a minimizing measure to be globally invariant under the Euler–Lagrange flow. To deduce properties of a Mather measure *on the discontinuity set*, we impose in Sect. 8.3 an extra assumption which implies that a minimizing measure $\bar{\mu}$ does not give mass to the discontinuity locus; see (8.14). Under this additional requirement we prove that there exist measures σ^+ and σ^- on $\Sigma \times \mathbb{R}^N$ such that

$$\int_{(\mathbb{T}^N\backslash\Sigma)\times\mathbb{R}^N} \chi \cdot D\phi \, \mathrm{d}\bar{\mu} = \int_{\Sigma\times\mathbb{R}^N} \phi \, \mathrm{d}\sigma^+ - \int_{\Sigma\times\mathbb{R}^N} \phi \, \mathrm{d}\sigma^+;$$

see Theorem 8.8. Another example in Sect. 8.1 shows that such a representation may fail to hold if $\bar{\mu}$ gives positive mass to the discontinuity set.

The statement of Theorem 8.8 also includes further properties of σ^+ and σ^- that—in strong analogy with the necessary conditions for minimizing trajectories—we interpret as conservation of total mass, conservation of tangential momentum, and conservation of energy (see Remark 8.10).

4. Optimality conditions for minimizing trajectories

In this section we discuss the main properties of minimizing trajectories of discontinuous Lagrangians. A well known result in classical calculus of variations is that for non-convex Lagrangians minimizing trajectories of the action may fail to have continuous derivatives, see for instance [5, Chapter 1, §9]. In this section we address the case of discontinuous Lagrangians, where similar discontinuities can also arise.

Let $T > 0$ and $a, b \in \mathbb{T}^N$ be fixed. Define, for any curve $\mathbf{x}(t)$ belonging to $\mathbf{X}^T_{a,b} := \{\mathbf{x} \in W^{1,\infty}([0,T];\mathbb{T}^N) : \mathbf{x}(0) = a,\, \mathbf{x}(T) = b\}$ the action

$$A(\mathbf{x}) := \int_0^T L(\mathbf{x}(t), \dot{\mathbf{x}}(t))\, \mathrm{d}t. \tag{4.1}$$

Consider the action-minimization problem, i.e. the problem of finding

$$\min A(\mathbf{x}) \text{ among all curves } \mathbf{x} \in \mathbf{X}^T_{a,b}.$$

Proposition 4.1. *There exists* $\mathbf{x} \in \mathbf{X}^T_{a,b}$ *minimizing the action.*

Proof. The result follows by the direct method of the calculus of variations. Note that here we need to use the lower semicontinuity (2.1) of L to show that the limit of a minimizing sequence is a minimizer. □

Necessary conditions for minimizing trajectories. We discuss next necessary conditions satisfied by minimizing trajectories of the action functional. We immediately observe that a minimizing trajectory $\mathbf{x}(\cdot)$ satisfies the usual Euler–Lagrange equation

$$-\frac{\mathrm{d}}{\mathrm{d}t} D_v L(\mathbf{x}(t), \dot{\mathbf{x}}(t)) + D_x L(\mathbf{x}(t), \dot{\mathbf{x}}(t)) = 0, \quad \text{for any } t \text{ such that } \mathbf{x}(t) \notin \Sigma. \tag{4.2}$$

Conditions for a minimizing trajectory $\mathbf{x}$ passing through Σ are investigated in the next propositions. To do this, we assume that such trajectories are piecewise C^2. The proof of the following propositions are postponed in the Appendix 8.3 as they use standard techniques in calculus of variations. The difficulty of having trajectories passing through the discontinuity set is handled considering special variations of minimizers and isolating the singularities.

Definition 4.2. Let $\mathbf{x} : [0,T] \to \mathbb{T}^N$ be a Lipschitz curve, $t_0 \in (0,T)$ and $x_0 \in \Sigma$. We say that $\mathbf{x}$ *crosses* Σ in x_0 at time t_0 if $\mathbf{x}(t_0) = x_0$, and there exists $\delta > 0$ such that $\mathbf{x}(t) \notin \Sigma$ for any t in the punctured interval $(t_0-\delta, t_0+\delta)\setminus\{t_0\}$.

According to the previous definition, a trajectory also crosses Σ if it bounces or reflects in Σ. The next proposition is also valid in these cases.

Proposition 4.3. *Let* $\mathbf{x} \in \mathbf{X}^T_{a,b}$ *be a piecewise* C^2 *minimizer of the action. Assume that* $\mathbf{x}$ *crosses* Σ *at* $x_0 = \mathbf{x}(t_0)$*, for some* $t_0 \in (0, T)$*. Then,*

$$\text{for any } \xi \in T_{\mathbf{x}(t_0)}\Sigma, \quad [\![\xi \cdot D_v L(\mathbf{x}(t), \dot{\mathbf{x}}(t))]\!]_{t_0} = 0. \tag{4.3}$$

Proof. See Appendix 8.3. □

We consider in the next Proposition the case in which a minimizer stays inside Σ during an interval:

Proposition 4.4. *For any* $x \in \Sigma$*, let* $\nu(x) \in \mathbb{R}^N$ *denote the unit normal to* Σ *at* x *pointing towards* Ω^-*. Let* $\mathbf{x} \in \mathbf{X}^T_{a,b}$ *be a piecewise* C^2 *minimizer of the action. Assume that there exist* $t_1, t_2, 0 < t_1 < t_2 < T$ *such that* $\mathbf{x}(t) \in \Sigma$ *for any* $t \in (t_1, t_2)$*. Then*

$$\left[D_x L(\mathbf{x}(t), \dot{\mathbf{x}}(t)) - \frac{\mathrm{d}}{\mathrm{d}t} D_v L(\mathbf{x}(t), \dot{\mathbf{x}}(t))\right] \cdot \nu(\mathbf{x}(t)) \geq 0, \quad \text{for any } t \in (t_1, t_2). \tag{4.4}$$

Moreover, for any $t \in (t_1, t_2)$ *and any* $\xi \in T_{\mathbf{x}(t)}\Sigma$*,*

$$\left[D_x L(\mathbf{x}(t), \dot{\mathbf{x}}(t)) - \frac{\mathrm{d}}{\mathrm{d}t} D_v L(\mathbf{x}(t), \dot{\mathbf{x}}(t))\right] \cdot \xi = 0. \tag{4.5}$$

Proof. See Appendix 8.3. □

Definition 4.5. Let $\mathbf{x} : [0, T] \to \mathbb{T}^N$ be a Lipschitz curve, $t_0 \in (0, T)$ and $x_0 \in \Sigma$. We say that $\mathbf{x}$ *enters* Σ at x_0 at time t_0 if $\mathbf{x}(t_0) = x_0$ and there exists $\delta > 0$ such that $\mathbf{x}(t) \notin \Sigma$ for any $t \in (t_0 - \delta, t_0)$ and $\mathbf{x}(t) \in \Sigma$ for any $t \in [t_0, t_0 + \delta)$. The definition of a curve *exiting* Σ at x_0 is symmetric.

Proposition 4.6. *Let* $\mathbf{x} \in \mathbf{X}^T_{a,b}$ *be a piecewise* C^2 *minimizer of the action. Assume that* $\mathbf{x}$ *enters* (*or exits*) Σ *at* x_0 *at time* $t_0 \in (0, T)$*. Then* (4.3) *holds.*

Proof. See Appendix 8.3. □

Conservation of energy. The energy associated to a Lipschitz curve $\mathbf{x} \in \mathbf{X}^T_{a,b}$ is the function $E_{\mathbf{x}} : (0, T) \to \mathbb{R}$ defined as

$$E_{\mathbf{x}}(t) := D_v L(\mathbf{x}(t), \dot{\mathbf{x}}(t))\dot{\mathbf{x}}(t) - L(\mathbf{x}(t), \dot{\mathbf{x}}(t)).$$

Proposition 4.7. *Let* $\mathbf{x} \in \mathbf{X}^T_{a,b}$ *be a piecewise* C^2 *minimizer of the action. Then the energy of* $\mathbf{x}$ *is conserved on* $[0, T]$*, that is* $E_{\mathbf{x}}(0) = E_{\mathbf{x}}(T)$*.*

Proof. See Appendix 8.3. □

5. The Mather problem, minimizing measures and duality

In this section we establish existence of minimizing holonomic probability measures, a duality result and a representation formula for the value of the Mather problem. We point out that the most part of this section is an adaptation of known results, see for instance [4], and references therein.

By coercivity (**H1**), there exists a function γ which satisfies

$$\lim_{|v|\to+\infty} \frac{\gamma(v)}{|v|} = +\infty, \quad \text{and} \quad \lim_{|v|\to+\infty} \frac{L(x,v)}{\gamma(v)} = +\infty. \tag{5.1}$$

We consider the following set of σ-finite, γ-weighted signed Borel measures on $\mathbb{T}^N \times \mathbb{R}^N$:

$$\mathcal{M} := \left\{ \mu \text{ signed measures on } \mathbb{T}^N \times \mathbb{R}^N \text{ such that } \int \gamma(v) \mathrm{d}|\mu| < \infty \right\}.$$

Recall that, for any Borel subset S, the total variation measure $|\mu|(S)$ is the supremum of $\sum_{i=1}^{\infty} |\mu(S_i)|$ among all Borel partitions $\{S_i\}$ of S. Thus, $|\mu|$ is σ-finite if and only if μ is so.

Remark 5.1. By applying the Riesz representation Theorem we see that $\mathcal{M}$ is the dual of the set $\mathcal{C}_0^\gamma$ of continuous functions $\phi : \mathbb{T}^N \times \mathbb{R}^N \to \mathbb{R}$ satisfying

$$\|\phi\|_\gamma := \sup_{\mathbb{T}^N\times\mathbb{R}^N} \left| \frac{\phi(x,v)}{\gamma(v)} \right| < \infty, \quad \lim_{|v|\to+\infty} \frac{\phi(x,v)}{\gamma(v)} = 0, \text{ unif. in } x \in \mathbb{T}^N. \tag{5.2}$$

□

We further define the subsets of $\mathcal{M}$ of holonomic measures and probability measures:

$$\mathcal{M}_{\mathrm{hol}} := \left\{ \mu \in \mathcal{M} \text{ such that } \int vD\varphi \, \mathrm{d}\mu = 0 \text{ for any } \varphi \in C^1(\mathbb{T}^N) \right\};$$

$$\mathcal{M}_1^+ := \left\{ \mu \in \mathcal{M}, \text{nonnegative and such that } \int \mathrm{d}\mu = 1 \right\}.$$

5.1. Existence of minimal holonomic probability measures

In the next proposition we revisit the well known result, due to Mañé [18], concerning the existence of minimizing holonomic probability measures. The main point that we make is that such existence holds even if L is only lower semicontinuous. We also show, in Proposition 5.5, that, as in the continuous case, the minimizing measure is supported on a graph.

Proposition 5.2. *There exists a solution $\bar{\mu} \in \mathcal{M}_{\mathrm{hol}} \cap \mathcal{M}_1^+$ to the problem*

$$\inf_{\mu\in\mathcal{M}_1^+\cap\mathcal{M}_{\mathrm{hol}}} \int L(x,v) \, \mathrm{d}\mu. \tag{5.3}$$

Proof. Let $\{\mu_n\}_n \subset \mathcal{M}_1^+ \cap \mathcal{M}_{\mathrm{hol}}$ be a minimizing sequence. The sequence $\int_{\mathbb{T}^N\times\mathbb{R}^N} L(x,v) \, \mathrm{d}\mu_n$ is bounded (because μ_n is minimizing). Then, μ_n converges weak-star to a measure $\bar{\mu}$, up to a subsequence. Such measure $\bar{\mu}$ belongs to $\mathcal{M}_1^+ \cap \mathcal{M}_{\mathrm{hol}}$. In fact each μ_n is a probability measure, and for any $\varphi \in C^1(\mathbb{T}^N)$,

$$\int vD\varphi \, \mathrm{d}\bar{\mu} = \lim_n \int vD\varphi \, \mathrm{d}\mu_n = 0. \tag{5.4}$$

We now show that $\bar{\mu}$ is a minimizer. Since $L(x, v)$ is lower semicontinuous in x, there exists a sequence of continuous functions $L_k(x, v)$, satisfying (5.2) and converging pointwise to $L(x, v)$ from below (see Remark 5.3 later). Furthermore, as $L \geq 1$ we can assume $L_k \geq 0$ for any k. We have for any k

$$\int L(x, v)\ \mathrm{d}\mu_n \geq \int L_k(x, v)\ \mathrm{d}\mu_n \to \int L_k(x, v)\ \mathrm{d}\bar{\mu}, \quad \text{as } n \to \infty$$

Then, by Fatou's lemma, we conclude:

$$\begin{aligned}\inf_{\mu \in \mathcal{M}_{\mathrm{hol}} \cap \mathcal{M}_1^+} \int L(x, v)\ \mathrm{d}\mu &:= \lim_n \int L(x, v) \mathrm{d}\mu_n \\ &\geq \liminf_k \int L_k(x, v)\ \mathrm{d}\bar{\mu} \geq \int L(x, v)\ \mathrm{d}\bar{\mu}.\end{aligned}$$

□

Remark 5.3. An example of function L_k approximating L as in the previous proof can be constructed as follows: first define $\tilde{L}_k(x, v) := \inf_y\{L(y, v) + k|x - y|^2\}$, the inf convolution, which is continuous. Then take $g : [0, \infty) \to \mathbb{R}$, a non-negative decreasing compactly supported function, with $g(r) = 1$ for r in a neighborhood of 0. Define

$$L_k(x, v) = \tilde{L}_k(x, v) g\left(\frac{|v|}{k}\right).$$

For any k, L_k is continuous, compactly supported (and hence in $\mathcal{C}_0^\gamma$), and, for any $(x, v), L_k(x, v) \leq L(x, v)$. Moreover L_k increases to L pointwise as k goes to $+\infty$. □

Remark 5.4. Let $\bar{\mu} \in \mathcal{M}_{\mathrm{hol}} \cap \mathcal{M}_1^+$ be a minimizing measure. Then, by disintegration (see Evans [12]) there exists a probability measure $\theta(\mathrm{d}x)$ on $\mathbb{T}^N$ and, for θ-a.e. x in $\mathbb{T}^N$ there exists a probability measure $\eta(\mathrm{d}v; x)$ on $\mathbb{R}^N$ such that $\bar{\mu}(\mathrm{d}x, \mathrm{d}v) = \theta(\mathrm{d}x)\eta(\mathrm{d}v; x)$. □

Proposition 5.5. *If $\bar{\mu} \in \mathcal{M}_{\mathrm{hol}} \cap \mathcal{M}_1^+$ minimizes (5.3) then $\bar{\mu}$ is supported on a graph. That is, there exists a measurable function $\mathbf{v} : \mathbb{T}^N \to \mathbb{R}^N$ such that*

$$\operatorname{supp}(\bar{\mu}) \subset \left\{(x, v) \in \mathbb{T}^N \times \mathbb{R}^N \text{ such that } v = \mathbf{v}(x)\right\}.$$

Proof. See Appendix 8.3. □

Remark 5.6. In Proposition 5.5 we only show that $\mathbf{v}$ is well defined μ-a.e. in suppμ. The value of $\mathbf{v}$ is not assigned, in principle, everywhere, but it is possible to extend it Lebesgue a.e. in $\mathbb{T}^N$ in a canonical way, as we will explain in Remark 7.10. □

5.2. Duality

The main purpose of this section is to prove the following duality result:

Theorem 5.7. *The following identity holds:*

$$\inf_{\mu\in\mathcal{M}_1^+}\sup_{\varphi\in C^1(\mathbb{T}^N)}\int[L(x,v)-vD_x\varphi(x)]\ \mathrm{d}\mu$$
$$=\sup_{\varphi\in C^1(\mathbb{T}^N)}\inf_{\mu\in\mathcal{M}_1^+}\int[L(x,v)-vD_x\varphi(x)]\ \mathrm{d}\mu.$$

The proof of Theorem 5.7 uses the following well known fact:

Theorem 5.8. (Legendre–Fenchel–Rockafellar duality Theorem) *Let E be a locally convex topological vector space over $\mathbb{R}$ with dual E'. Let $h : E \to (-\infty,+\infty]$ a convex function and $g : E \to [-\infty,+\infty)$ a concave function and let $\hat{h}$ and $\hat{g}$ their Legendre–Fenchel transform that is*

$$\hat{h}(y)=\sup_{x\in E}\{x\cdot y-h(x)\},$$
$$\hat{g}(y)=\inf_{x\in E}\{x\cdot y-g(x)\}.$$

Assume also that there exists $x_0\in E$ such that $g(x_0)$ and $h(x_0)$ are finite, and at least one of them is continuous. Then

$$\min_{y\in E'}[\hat{h}(y)-\hat{g}(y)]=\sup_{x\in E}[g(x)-h(x)].$$

Proof. See [24] or [4]. □

Consider the following convex subset of $\mathcal{C}_0^\gamma$ introduced in Remark 5.1:

$$\mathcal{C}:=\left\{\psi\in\mathcal{C}_0^\gamma(\mathbb{T}^N\times\mathbb{R}^N)\,\middle|\,\psi(x,v)=vD\varphi(x),\ \varphi\in C^1(\mathbb{T}^N)\right\},\tag{5.5}$$

and define $g,h:\mathcal{C}_0^\gamma(\mathbb{T}^N\times\mathbb{R}^N)\to\mathbb{R}$ as follows:

$$g(\psi):=\min_{\substack{x\in\mathbb{T}^N\\ v\in\mathbb{R}^N}}\{L(x,v)+\psi(x,v)\};\quad h(\psi):=\begin{cases}0, & \text{if }\psi\in\mathcal{C}\\ +\infty, & \text{if }\psi\notin\mathcal{C}.\end{cases}\tag{5.6}$$

Observe that $\mathcal{M}_{\mathrm{hol}}=\{\mu\in\mathcal{M}$ such that $\int\psi\ \mathrm{d}\mu=0$ for any $\psi\in\mathcal{C}\}$. We immediately note that h is finite on $\mathcal{C}$ and convex, because $\mathcal{C}$ is so. The function g is concave, because is the infimum of an affine function and finite on $\mathcal{C}$, because L is superlinear. Consequently, the Legendre–Fenchel transform of h and g can be defined. In the next Lemmas we establish further properties of h and g we will use in the proof of Theorem 5.7.

Lemma 5.9. *The function g defined in (5.6) is continuous in $\mathcal{C}_0^\gamma(\mathbb{T}^N\times\mathbb{R}^N)$.*

Proof. Let $\psi_n\to\psi$ in $\mathcal{C}_0^\gamma$. Then $\|\psi_n\|_\gamma$, and $\|\psi\|_\gamma$ are uniformly bounded. Thus, by (**H1**), there exists $R>0$ such that the minima of $L+\psi$ and $L+\psi_n$ in $\mathbb{T}^N\times\mathbb{R}^N$ are achieved on $\mathbb{T}^N\times B(0,R)$. As in $\mathbb{T}^N\times B(0,R)$ the sequence ψ_n converges to ψ uniformly, we get the conclusion:

$$\min_{\mathbb{T}^N\times\mathbb{R}^N}L+\psi_n=\min_{\mathbb{T}^N\times B(0,R)}L+\psi_n\to\min_{\mathbb{T}^N\times B(0,R)}L+\psi=\min_{\mathbb{T}^N\times\mathbb{R}^N}L+\psi.$$

□

Lemma 5.10. *The Legendre–Fenchel transform of h and g defined in* (5.6) *are respectively*

$$\hat{h}(\mu) := \begin{cases} 0, & \text{if } \mu \in \mathcal{M}_{\text{hol}} \\ +\infty, & \text{otherwise} \end{cases}$$

$$\hat{g}(\mu) := \begin{cases} -\int L(x,v)\,\mathrm{d}\mu & \text{if } \mu \in \mathcal{M}_1^+ \\ -\infty, & \text{otherwise.} \end{cases}$$

Proof. See Appendix 8.3. □

Proof of Theorem 5.7*.* Observe that, for any $\varphi \in C^1(\mathbb{T}^N)$,

$$\inf_{\mu\in\mathcal{M}_1^+} \int [L(x,v) - vD\varphi(x)]\,\mathrm{d}\mu = \min_{x,v}\{L(x,v) - vD\varphi(x)\},$$

then

$$\sup_{\varphi\in C^1(\mathbb{T}^N)} \inf_{\mu\in\mathcal{M}_1^+} \int [L(x,v) - vD\varphi(x)]\,\mathrm{d}\mu = \sup_{\psi\in\mathcal{C}_0^\gamma} [g(\psi) - h(\psi)]$$

where $g, h : \mathcal{C}_0^\gamma \to \mathbb{R}$ are defined in (5.6). Note also that g and h are finite on $\mathcal{C}$, and g is continuous by Lemma 5.9. Then, by applying Theorem 5.8 with $E = \mathcal{C}_0^\gamma$ (thus $E' = \mathcal{M}$ by Remark 5.1), and taking into account Lemma 5.10, we conclude the proof:

$$\begin{aligned}\sup_{\varphi\in C^1(\mathbb{T}^N)} \inf_{\mu\in\mathcal{M}_1^+} \int [L(x,v) - vD\varphi(x)]\,\mathrm{d}\mu &= \sup_{\psi\in\mathcal{C}_0^\gamma} [g(\psi) - h(\psi)] \\ = \inf_{\mu\in\mathcal{M}} [\hat{h}(\mu) - \hat{g}(\mu)] &= \inf_{\mu\in\mathcal{M}_1^+\cap\mathcal{M}_{\text{hol}}} \int L(x,v)\,\mathrm{d}\mu \\ = \inf_{\mu\in\mathcal{M}_1^+} \sup_{\varphi\in C^1(\mathbb{T}^N)} &\int [L(x,v) - vD\varphi(x)]\,\mathrm{d}\mu.\end{aligned}$$

□

As a consequence of Theorem 5.7, we have the following

Corollary 5.11. *Let $H(x,p)$ be the Legendre transform of $L(x,v)$. The quantity*

$$\bar{H} := \inf_{\varphi\in C^1(\mathbb{T}^N)} \sup_{x\in\mathbb{T}^N} H(x, D\varphi(x)) \tag{5.7}$$

is well defined, and coincides with

$$-\inf_{\mu\in\mathcal{M}_1^+\cap\mathcal{M}_{\text{hol}}} \int L(x,v)\,\mathrm{d}\mu.$$

Proof. By applying the Legendre transform, we first observe that

$$\begin{aligned}\sup_{\mu\in\mathcal{M}_1^+} \int [vD\varphi(x) - L(x,v)]\,\mathrm{d}\mu &= \sup_x \sup_v (vD\varphi(x) - L(x,v)) \\ &= \sup_x H(x, D\varphi(x)).\end{aligned} \tag{5.8}$$

Moreover, since H is upper semicontinuous, the supremum in the right hand side of (5.8) is well defined. Using Theorem 5.7, and taking into account (5.8), we compute:

$$\begin{aligned}
\inf_{\mu\in\mathcal{M}_1^+\cap\mathcal{M}_{\text{hol}}}\int L(x,v)\,\mathrm{d}\mu &= \inf_{\mu\in\mathcal{M}_1^+}\sup_{\varphi\in C^1(\mathbb{T}^N)}\int[L(x,v)-vD\varphi(x)]\,\mathrm{d}\mu\\
&= \sup_{\varphi\in C^1(\mathbb{T}^N)}\inf_{\mu\in\mathcal{M}_1^+}\int[L(x,v)-vD\varphi(x)]\,\mathrm{d}\mu\\
&= -\inf_{\varphi\in C^1(\mathbb{T}^N)}\left[-\inf_{\mu\in\mathcal{M}_1^+}\int[L(x,v)-vD\varphi(x)]\,\mathrm{d}\mu\right]\\
&= -\inf_{\varphi\in C^1(\mathbb{T}^N)}\left[\sup_{\mu\in\mathcal{M}_1^+}\int[vD\varphi(x)-L(x,v)]\,\mathrm{d}\mu\right]\\
&= -\inf_{\varphi\in C^1(\mathbb{T}^N)}\left[\sup_x H(x,D\varphi(x))\right].
\end{aligned}$$

□

5.3. Holonomy preserving variations

We next describe a class of variations, introduced in [4], that preserves the holonomy constraint. This type of variation will be used in Sect. 8 to establish the properties of a minimal holonomic measure.

We consider next a C^1 vector field $\xi:\mathbb{T}^N\to\mathbb{R}^N$ which is either compactly supported on $\mathbb{T}^N\backslash\Sigma$ or tangent to Σ and compactly supported in a neighborhood of it. Due to the lack of continuity of L we cannot make variations in the normal direction to Σ, and this explains the previous conditions on ξ.

Proposition 5.12. *Let $\bar\mu$ be a minimizing holonomic measure and $\xi:\mathbb{T}^N\to\mathbb{R}^N$, a C^1 vector field. If*

- *either ξ is compactly supported in $\mathbb{T}^N\backslash\Sigma$,*
- *or ξ is compactly supported in a neighborhood of Σ and tangent to it,*

then,

$$\int_{\mathbb{T}^N\times\mathbb{R}^N}\left[\xi_s(x)\frac{\partial L}{\partial x_s}(x,v)+v_k\frac{\partial L}{\partial v_s}(x,v)\frac{\partial\xi_s}{\partial x_k}(x)\right]\,\mathrm{d}\bar\mu=0.\quad (s=1,\dots N). \tag{5.9}$$

Proof. The proof goes along the same lines of that of [4, Theorem 41]. Let $\Psi(t,x)$ the flow generated by ξ, that is, $\Psi(0,x)=x$ for any $x\in\mathbb{T}^N$ and

$$\frac{\mathrm{d}}{\mathrm{d}t}\Psi(t,x)=\xi(\Psi(t,x))$$

for any $t>0$ and $x\in\mathbb{T}^N$. The flow of ξ can be extended to $\mathbb{T}^N\times\mathbb{R}^N$ by considering

$$\begin{cases}\dot x_s(x,v)=\xi_s(x),\\ \dot v_s(x,v)=v_k\frac{\partial}{\partial x_k}\xi_s(x),\end{cases}\qquad (s=1,\dots,N). \tag{5.10}$$

A direct computation shows that flow $(X(t,x,v),V(t,x,v))$ associated to (5.10) is

$$X_s(t,x,v)=\Psi_s(t,x),\quad V_s(t,x,v)=v_k\frac{\partial}{\partial x_k}\Psi_s(t,x). \tag{5.11}$$

The flow (5.11) induces a flow on functions $\psi:\mathbb{T}^N\times\mathbb{R}^N\to\mathbb{R}$ given, for any $t>0$, by

$$\psi_t(x,v):=\psi(X(t,x,v),V(t,x,v)). \tag{5.12}$$

We next use the flow (5.11) and formula (5.12) to define a flow on measures. For any $t>0$ and any measure μ over $\mathbb{T}^N\times\mathbb{R}^N$ we set μ_t to be such that

$$\int_{\mathbb{T}^N\times\mathbb{R}^N}\psi(x,v)\ \mathrm{d}\mu_t(x,v)=\int_{\mathbb{T}^N\times\mathbb{R}^N}\psi_t(x,v)\ \mathrm{d}\mu(x,v),\quad \text{for any } \psi. \tag{5.13}$$

An easy computation shows that the flow (5.11) preserves the set $\mathcal{C}$ defined in (5.5). Thus the flow on measures $(t,\mu)\mapsto\mu_t$ given by (5.13) preserves the holonomy constraint; in fact, for any $\psi\in\mathcal{C}$,

$$\int\psi\,\mathrm{d}\mu_t=\int\psi_t\,\mathrm{d}\mu=0$$

because $\psi_t\in\mathcal{C}$. Then, since $\bar{\mu}$ is a minimizing holonomic measure

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0}\left(\int L\,\mathrm{d}\bar{\mu}_t\right)=0,$$

which gives (5.9), taking into account (5.10). □

5.4. An example in one dimension

We consider, for $(x,v)\in[0,1]\times\mathbb{R}$, the Lagrangian

$$L(x,v):=\frac{|v|^2}{2}+V(x),$$

with $V(x)=0$ for any $x\in[0,\frac{1}{3}]\cup[\frac{2}{3},1]$, and $V(x)=1$ for any $x\in(\frac{1}{3},\frac{2}{3})$. The discontinuity locus, for such L, is then the set $\Sigma=\{\frac{1}{3},\frac{2}{3}\}$. We exhibit next a solution to the problem

$$\min\left\{\int L(x,v)\ \mathrm{d}\mu\,\middle|\,\mu \text{ holonomic probability measures over } [0,1]\times\mathbb{R}\right\}. \tag{5.14}$$

Consider for any $\lambda\in\mathbb{R}$ the set

$$\tilde{\mathcal{M}}(\lambda):=\Big\{\mu:\ \mathrm{d}\mu(x,v)=\frac{\lambda}{v_1}\delta_{v_1}(v)\ \mathrm{d}x\Big|_{[0,\frac{1}{3}]\cup[\frac{2}{3},1)}+\frac{\lambda}{v_2}\delta_{v_2}(v)\ \mathrm{d}x\Big|_{(\frac{1}{3},\frac{2}{3})},$$
$$\text{such that } v_1,v_2>0 \text{ and } \lambda=\frac{3v_1v_2}{v_1+2v_2}\Big\}.$$

The quantity λ is the so called rotation number associated to measures in $\tilde{\mathcal{M}}(\lambda)$. In fact it is easy to verify that for any $\mu\in\tilde{\mathcal{M}}(\lambda)$,

$$\int v\,\mathrm{d}\mu=\lambda. \tag{5.15}$$

Lemma 5.13. *Let $\lambda > 0$ be fixed. The minimum in (5.14), under the constraint (5.15) is attained on the set $\tilde{\mathcal{M}}(\lambda)$.*

Proof. Observe that any measure $\mu \in \tilde{\mathcal{M}}(\lambda)$ is an holonomic probability measure. In fact the condition

$$\lambda = \frac{3v_1 v_2}{v_1 + 2v_2}$$

readily implies that the total mass of μ is 1. Moreover, for any $\varphi \in C^1([0,1])$, such that $\varphi(0) = \varphi(1)$,

$$\int_{[0,1]\times\mathbb{R}} v\varphi'(x)\ \mathrm{d}\mu(x,v) = \int_{\mathbb{R}} \left[\lambda \int_{[0,1]} \varphi'(x)\ \mathrm{d}x\right]\ \mathrm{d}v = 0.$$

Let $\bar{\mu}$ be the solution of (5.14). By Remark 5.4, $\bar{\mu}(x,v) = \theta(x)\eta(v;x)$. Furthermore, by Proposition 5.5 there exists $\mathbf{v} : [0,1] \to \mathbb{R}$ measurable whose graph supports $\bar{\mu}$; thus, taking into account (5.15) we can write

$$\theta(x) = \frac{\lambda}{|\mathbf{v}(x)|}, \quad \eta(v;x) = \delta_{\mathbf{v}(x)}(v).$$

The function $\mathbf{v}(x)$ never changes its sign, otherwise it is easy to obtain a contradiction with μ being holonomic. In addition, since we assumed $\lambda > 0, \mathbf{v}(x)$ must be positive. Finally, since L is convex in v, we have that $\mathbf{v}(x)$ is constant in each of the three subintervals of $[0,1]$. In fact if, for instance, $\mathbf{v}(x)$ is non constant in $(\frac{1}{3}, \frac{2}{3})$, then we can define the averaged velocity in this interval by

$$\tilde{v} := \left(\int_{\frac{1}{3}}^{\frac{2}{3}} \mathbf{v}(x)\theta(x)\ \mathrm{d}x\right)\left(\int_{\frac{1}{3}}^{\frac{2}{3}} \theta(x)\ \mathrm{d}x\right)^{-1}.$$

Then by convexity the measure

$$\tilde{\mu}(x,v) := \begin{cases} \frac{\lambda}{\tilde{v}}\,\delta_{\tilde{v}}(v), & \text{if } x \in \left(\frac{1}{3}, \frac{2}{3}\right) \\ \bar{\mu}(x,v), & \text{otherwise} \end{cases}$$

would satisfy $\int L\ \mathrm{d}\tilde{\mu} \le \int L\ \mathrm{d}\bar{\mu}$. That is $\bar{\mu}$ would not be a minimizer. □

We assume, for simplicity, $\lambda = 1$, and consider the problem

$$\min_{\mu \in \tilde{\mathcal{M}}(1)} \int L(x,v)\ \mathrm{d}\mu(x,v). \tag{5.16}$$

By condition

$$1 = \frac{3v_1 v_2}{v_1 + 2v_2} \tag{5.17}$$

the admissible values for the pair (v_1, v_2) must satisfy

$$v_1 = \frac{2v_2}{3v_2 - 1}.$$

In particular, any v_2 in the interval $(0, \frac{1}{3})$ is not admissible, as the corresponding value for v_1 would be negative. A direct computation shows that the minimum in (5.16) is attained for a unique pair $(\bar{v}_1, \bar{v}_2)$, satisfying (5.17), with $\bar{v}_2 > 1/3$. By Lemma 5.13 the measure

$$\mathrm{d}\bar{\mu}(x,v) = \frac{1}{\bar{v}_1}\delta_{\bar{v}_1}(v)\ \mathrm{d}x\Big|_{[0,\frac{1}{3}]\cup[\frac{2}{3},1)} + \frac{1}{\bar{v}_2}\delta_{\bar{v}_2}(v)\ \mathrm{d}x\Big|_{(\frac{1}{3},\frac{2}{3})} \tag{5.18}$$

solves (5.14) under the constrain $\int v\,\mathrm{d}\mu = 1$.

6. Approximation of minimizing measures

In Proposition 5.2 we proved the existence of a minimizing measure for the Mather problem for a lower semicontinuous Lagrangian. In this Section we show that it is possible to obtain such a measure as a weak-star limit of a sequence of minimizing measures for continuous approximating Lagrangians. Suppose L satisfies assumption (**H3**) of Sect. 2 that we rewrite for convenience,

$$L(x,v) = L_0(x,v) + V(x), \tag{6.1}$$

and consider the Yosida inf-convolution of V, which is defined as

$$V^{\epsilon}(x) := \inf_{y\in\mathbb{T}^N}\left\{V(y) + \frac{|x-y|^2}{\epsilon}\right\}, \quad \epsilon > 0. \tag{6.2}$$

Then set

$$L^{\epsilon}(x,v) := L_0(x,v) + V^{\epsilon}(x). \tag{6.3}$$

Inf-convolution is a natural way to construct a continuous approximation to V from below. The aim of this Section is to prove that any weak-star limit of a sequence of minimizers for the approximating problems is a solution to the Mather problem, that is:

Theorem 6.1. *Assume that L has the form (6.1). For any $\epsilon > 0$, let μ^{ϵ} be the solution of*

$$\min_{\mu\in\mathcal{M}_1^+\cap\mathcal{M}_{\mathrm{hol}}} \int L^{\epsilon}(x,v)\ \mathrm{d}\mu, \tag{6.4}$$

where L^{ϵ} is defined in (6.3). Assume that μ^{ϵ} converges $\bar{\mu}$ weak–star, as $\epsilon \to 0$. Then $\bar{\mu}$ is a solution to

$$\min_{\mu\in\mathcal{M}_1^+\cap\mathcal{M}_{\mathrm{hol}}} \int L(x,v)\ \mathrm{d}\mu.$$

We postpone the proof of theorem in the end of this Section, and establish first some preliminary result. We denote by $\mathcal{P}(\mathbb{T}^N)$ the space of Radon probability measures on $\mathbb{T}^N$; $\mathcal{P}(\mathbb{T}^N)$ is compact with respect to the weak-star topology.

Lemma 6.2. *The function $G : \mathcal{P}(\mathbb{T}^N) \to \mathbb{R}$ defined by*

$$G(\nu) := \int_{\mathbb{T}^N} V(x)\ \mathrm{d}\nu(x) \tag{6.5}$$

is lower semicontinuous with respect to the weak–star convergence.

Proof. Take a sequence $\{\nu^\epsilon\}_\epsilon \subset \mathcal{P}(\mathbb{T}^N)$ converging weak-star to some $\bar{\nu}$ as $\epsilon \to 0$. Since V is lower semicontinuous, there exists a sequence of continuous functions V_n converging to V from below. Then, for any n, by weak-star convergence,

$$\liminf_{\epsilon\to 0} \int_{\mathbb{T}^N} V(x)\ \mathrm{d}\nu^\epsilon(x) \geq \liminf_{\epsilon\to 0} \int_{\mathbb{T}^N} V_n(x)\ \mathrm{d}\nu^\epsilon(x) = \int_{\mathbb{T}^N} V_n(x)\ \mathrm{d}\bar{\nu}(x).$$

Thus, by the monotone convergence theorem,

$$\liminf_{\epsilon\to 0} \int_{\mathbb{T}^N} V(x)\ \mathrm{d}\nu^\epsilon(x) \geq \lim_{n\to +\infty} \int_{\mathbb{T}^N} V_n(x)\ \mathrm{d}\bar{\nu}(x) = \int_{\mathbb{T}^N} V(x)\ \mathrm{d}\bar{\nu}(x).$$

□

The space $\mathcal{P}(\mathbb{T}^N)$ can be metrized with the Wasserstein distance, defined, for any couple of probability measures ν_1 and ν_2 in $\mathcal{P}(\mathbb{T}^N)$ through the formula

$$[W(\nu,\nu')]^2 = \inf_{\pi\in\Pi(\nu,\nu')} \left\{ \int_{\mathbb{T}^N\times\mathbb{T}^N} |x-y|^2\ \mathrm{d}\pi(x,y) \right\},$$

where $\Pi(\nu,\nu')$ is the set of Radon probability measures on $\mathbb{T}^N \times \mathbb{T}^N$, whose marginal on the first coordinate is ν and the marginal on the second coordinate is ν'. We can use W to define G^ϵ, the inf-convolution of the function G given in (6.5), that is:

$$G^\epsilon(\nu) = \inf_{\nu'\in\mathcal{P}(\mathbb{T}^N)} \left\{ G(\nu') + \frac{W(\nu,\nu')^2}{\epsilon} \right\}.$$

Lemma 6.3. *For any $\bar{\nu} \in \mathcal{P}(\mathbb{T}^N)$ and any sequence $\{\nu^\epsilon\}_\epsilon \subset \mathcal{P}(\mathbb{T}^N)$ converging weak-star to $\bar{\nu}$ as $\epsilon \to 0$,*

$$G(\bar{\nu}) \leq \liminf_{\epsilon\to 0} G^\epsilon(\nu^\epsilon). \tag{6.6}$$

Proof. The statement is valid in general for lower semicontinuous functions defined in compact metric spaces, and we do not use the explicit expression of the Wasserstein distance to prove the lemma.

Suppose by contradiction that (6.6) does not hold for some $\bar{\nu} \in \mathcal{P}(\mathbb{T}^N)$ and some sequence $\nu^\epsilon \to \bar{\nu}$. Then there exists $\beta > 0$ such that $G(\bar{\nu}) - \beta > G^\epsilon(\nu^\epsilon)$, if ϵ is sufficiently small. By the definition of G^ϵ and by compactness we have, for some $\tilde{\nu}^\epsilon \in \mathcal{P}(\mathbb{T}^N)$,

$$G(\tilde{\nu}^\epsilon) + \frac{W(\nu^\epsilon,\tilde{\nu}^\epsilon)^2}{\epsilon} = G^\epsilon(\nu^\epsilon) < G(\bar{\nu}) - \beta.$$

Using again the compactness of $\mathcal{P}(\mathbb{T}^N)$, $\tilde{\nu}^\epsilon$ converges weak-star, up to a subsequence, to some $\nu_0 \in \mathcal{P}(\mathbb{T}^N)$ as $\epsilon \to 0$. Passing to the lim inf as $\epsilon \to 0$ and taking into account Lemma 6.2 we get:

$$G(\bar{\nu}) - \beta > \liminf_{\epsilon\to 0} \left\{ G(\tilde{\nu}^\epsilon) + \frac{W(\nu^\epsilon,\tilde{\nu}^\epsilon)^2}{\epsilon} \right\} \geq G(\nu_0) + \liminf_{\epsilon\to 0} \left\{ \frac{W(\nu^\epsilon,\tilde{\nu}^\epsilon)^2}{\epsilon} \right\}.$$

Now, if $\nu_0 \neq \bar{\nu}$, since $G(\nu_0)$ is bounded, we reach a contradiction as the right hand side of the previous inequality is $+\infty$. Thus $\nu_0 = \bar{\nu}$ and we have

$$G(\bar{\nu}) - \beta \geq G(\bar{\nu}) + \liminf_{\epsilon \to 0} \left\{ \frac{W(\nu^\epsilon, \tilde{\nu}^\epsilon)^2}{\epsilon} \right\} \geq G(\bar{\nu}),$$

which is impossible. □

Lemma 6.4. *For any $\nu \in \mathcal{P}(\mathbb{T}^N)$ and any $\epsilon > 0$*

$$G^\epsilon(\nu) = \int_{\mathbb{T}^N} V^\epsilon(x) \mathrm{d}\nu(x).$$

Proof. Let $\nu \in \mathcal{P}(\mathbb{T}^N)$ and $\epsilon > 0$ be fixed. We have by definition

$$\begin{aligned} G^\epsilon(\nu) &= \inf_{\nu' \in \mathcal{P}(\mathbb{T}^N)} \left\{ \int_{\mathbb{T}^N} V(y) \, \mathrm{d}\nu'(y) + \inf_{\pi \in \Pi(\nu, \nu')} \left[\int_{\mathbb{T}^N \times \mathbb{T}^N} \frac{|x-y|^2}{\epsilon} \, \mathrm{d}\pi(x,y) \right] \right\} \\ &= \inf_{\nu' \in \mathcal{P}(\mathbb{T}^N)} \inf_{\pi \in \Pi(\nu, \nu')} \left\{ \int_{\mathbb{T}^N \times \mathbb{T}^N} \left[V(y) + \frac{|x-y|^2}{\epsilon} \right] \, \mathrm{d}\pi(x,y) \right\}. \end{aligned}$$

Now observe that for each x, the minimizing π should be supported on the minima of the function $V(y) + \frac{|x-y|^2}{\epsilon}$, and so

$$G^\epsilon(\nu) = \int_{\mathbb{T}^N} \inf_{y \in \mathbb{T}^N} \left[V(y) + \frac{|x-y|^2}{\epsilon} \right] \, \mathrm{d}\nu(x) = \int_{\mathbb{T}^N} V^\epsilon(x) \, \mathrm{d}\nu(x).$$

□

The proof of Theorem 6.1 is based on the following lower semicontinuity result:

Proposition 6.5. *Assume that L has the form (6.1) and let L^ϵ be the function defined in (6.3). Let μ^ϵ be a sequence of probability measures on $\mathbb{T}^N \times \mathbb{R}^N$ and assume that μ^ϵ converges weak-star to a probability measure $\bar{\mu}$ as $\epsilon \to 0$. Then*

$$\liminf_{\epsilon \to 0^+} \int L^\epsilon(x,v) \, \mathrm{d}\mu^\epsilon(x,v) \geq \int L(x,v) \, \mathrm{d}\bar{\mu}(x,v).$$

Proof. By using Lemma 6.4 and Lemma 6.3 we have

$$\begin{aligned} \liminf_{\epsilon \to 0} \int_{\mathbb{T}^N} V^\epsilon(x) \, \mathrm{d}\mu^\epsilon_{\mathbb{T}^N}(x) &= \liminf_{\epsilon \to 0} G^\epsilon(\mu^\epsilon_{\mathbb{T}^N}) \\ &\geq G(\bar{\mu}_{\mathbb{T}^N}) = \int_{\mathbb{T}^N} V(x) \, \mathrm{d}\bar{\mu}_{\mathbb{T}^N}(x). \end{aligned}$$

Then taking into account that L_0 is continuous, by weak–star convergence we conclude

$$\begin{aligned} &\liminf_{\epsilon \to 0} \int_{\mathbb{T}^N \times \mathbb{R}^N} L^\epsilon(x,v) \, \mathrm{d}\mu^\epsilon(x,v) \\ &\geq \liminf_{\epsilon \to 0} \int_{\mathbb{T}^N \times \mathbb{R}^N} L_0(x,v) \mathrm{d}\mu^\epsilon(x,v) + \liminf_{\epsilon \to 0} \int_{\mathbb{T}^N} V^\epsilon(x) \, \mathrm{d}\mu^\epsilon_{\mathbb{T}^N}(x) \\ &\geq \int_{\mathbb{T}^N \times \mathbb{R}^N} L_0(x,v) \, \mathrm{d}\bar{\mu}(x,v) + \int_{\mathbb{T}^N} V(x) \, \mathrm{d}\bar{\mu}_{\mathbb{T}^N}(x) = \int_{\mathbb{T}^N \times \mathbb{R}^N} L(x,v) \, \mathrm{d}\bar{\mu}(x,v). \end{aligned}$$

□

We conclude this section with the proof of Theorem 6.1.

Proof of Theorem 6.1. Since $V^\epsilon \leq V$ for any ϵ,

$$\int L^\epsilon(x,v)\,\mathrm{d}\mu \leq \int L(x,v)\,\mathrm{d}\mu, \quad \text{for any } \mu.$$

Moreover, because μ^ϵ is a minimizer,

$$\int L^\epsilon(x,v)\,\mathrm{d}\mu^\epsilon \leq \int L^\epsilon(x,v)\,\mathrm{d}\mu, \quad \text{for any } \epsilon \text{ and any } \mu.$$

Thus, by Proposition 6.5 we get, for any μ,

$$\int L(x,v)\,\mathrm{d}\bar{\mu} \leq \liminf_{\epsilon\to 0}\int L^\epsilon(x,v)\,\mathrm{d}\mu^\epsilon \leq \liminf_{\epsilon\to 0}\int L^\epsilon(x,v)\,\mathrm{d}\mu \leq \int L(x,v)\,\mathrm{d}\mu.$$

□

7. Viscosity solutions

Let $H(x,p)$ be the Legendre transform of $L(x,v)$, as in Sect. 2. In this section we study viscosity solutions to the Hamilton–Jacobi equation with H as Hamiltonian.

Observe that, since H is uniformly convex in p, we have, for some $\kappa > 0$,

$$H(x,p) \geq H(x,q) + D_pH(x,q)(p-q) + \frac{\kappa}{2}|p-q|^2, \quad \text{for all } x,p,q \in \mathbb{R}^N. \tag{7.1}$$

We are particularly interested in the Hamiltonian associated to the Lagrangian L defined in (6.1), that is:

$$H(x,p) = H_0(x,p) - V(x). \tag{7.2}$$

where H_0 is the Legendre transform of L_0.

Since $H(x,p)$ is not continuous in x, we rely on a special notion of viscosity solution for discontinuous equations; see [1], [2]. Recall first the definition of lower and upper envelopes of a locally bounded function $\psi(z)$:

$$\psi_*(z) := \liminf_{z_n\to z}\psi(z_n), \quad \psi^*(z) := \limsup_{z_n\to z}\psi(z_n). \tag{7.3}$$

Whenever ψ is a function of several variables, the envelopes have to be understood with respect to all variables.

Definition 7.1. We say that an upper semicontinuous function u, locally bounded in $\mathbb{T}^N$ is a viscosity *subsolution* of the equation

$$H(x,Du(x)) = 0, \quad x \in \mathbb{T}^N \tag{7.4}$$

if $H_*(x,D\varphi(x)) \leq 0$ for every $\varphi \in C^1(\mathbb{T}^N)$ and any local maximum point x of $u-\varphi$.

Symmetrically, a lower semicontinuous function u, locally bounded in $\mathbb{T}^N$ is a viscosity *supersolution* of (7.4) if $H^*(x,D\varphi(x)) \geq 0$ for every $\varphi \in C^1(\mathbb{T}^N)$ and any local minimum point x of $u-\varphi$.

A continuous function that satisfies both the conditions above is called a viscosity *solution* of (7.4).

Observe that any viscosity solution of (7.4) is a viscosity solution (in the usual sense) in $\mathbb{T}^N \backslash \Sigma$. In fact the lower and upper envelopes H_* and H^* coincide with H in $(\mathbb{T}^N \backslash \Sigma) \times \mathbb{R}^N$.

Remark 7.2. For viscosity solutions of discontinuous Hamiltonians introduced in Definition 7.1, the following stability result holds; see [2]. Let u^α be an upper semicontinuous subsolution (resp. lower semicontinuous supersolution) to the equation

$$H^\alpha(x, Du^\alpha(x)) = 0, \quad x \in \mathbb{T}^N$$

with H^α locally uniformly bounded in $\mathbb{T}^N \times \mathbb{R}^N$. Consider the lower and upper semilimits of the family H^α defined respectively as:

$$H_\sharp(x,p) := \liminf{}_\sharp H^\alpha(x,p) = \liminf_{\substack{(x_n,p_n)\to(x,p)\\ \alpha_n\to 0}} H^{\alpha_n}(x_n,p_n),$$

$$H^\sharp(x,p) := \limsup{}^\sharp H^\alpha(x,p) = \limsup_{\substack{(x_n,p_n)\to(x,p)\\ \alpha_n\to 0}} H^{\alpha_n}(x_n,p_n).$$

If u^α are locally uniformly bounded in $\mathbb{T}^N$ then

$$u^\sharp(x) := \limsup{}^\sharp u^\alpha(x) := \limsup_{\substack{y\to x\\ \alpha\to 0}} u^\alpha(y) \quad (\text{resp. } u_\sharp := \liminf_\sharp u^\alpha)$$

is a subsolution (resp. a supersolution) to the equation

$$H_\sharp(x, Du) = 0 \quad (\text{resp. } H^\sharp(x, Du) = 0) \quad \text{in } \mathbb{T}^N.$$

□

We will need to consider the following approximating problems. For any fixed ϵ, let V^ϵ be the inf-convolution of V, already considered in Sect. 6 and set

$$H^\epsilon(x,p) = H_0(x,p) - V^\epsilon(x). \tag{7.5}$$

For these functions, the lower and upper semilimits introduced in the previous Remark 7.2 are:

$$H_\sharp(x,p) := \liminf{}_\sharp H^\epsilon(x,p) = H_0(x,p) - \limsup_{\substack{x_n\to x\\ \epsilon_n\to 0}} V^{\epsilon_n}(x_n),$$

$$H^\sharp(x,p) := \limsup{}^\sharp H^\epsilon(x,p) = H_0(x,p) - \liminf_{\substack{x_n\to x\\ \epsilon_n\to 0}} V^{\epsilon_n}(x_n).$$

In the next lemma we compare the upper and lower semilimits of $H^\epsilon(x,p)$ with the envelopes of $H(x,p)$ defined through the formula (7.3).

Lemma 7.3. *Assume H has the form (7.2). For any x and p, $H^*(x,p) \geq H^\sharp(x,p)$ and $H_*(x,p) \leq H_\sharp(x,p)$.*

Proof. Because of the structure of $H(x,p)$ displayed in (7.2) we concentrate just on the discontinuous function V. Moreover, since $V(x)$ is continuous outside Σ, we need to prove the statement only in the case in which $x \in \Sigma$. We first compare H^* with $H^\sharp$.

Fix $\rho > 0$ and take sequences $z_n \to x$ and $\bar{\epsilon}_n \to 0$ such that

$$\liminf_n V^{\bar{\epsilon}_n}(z_n) \le \rho + \inf_{\substack{x_n \to x \\ \epsilon_n \to 0}} \liminf_n V^{\epsilon_n}(x_n).$$

Since $z_n \to x$, then there exists $\beta_n \to 0$ with $|\beta_n| \ge 2\bar{\epsilon}_n$ such that for any n,

$$B\left(z_n - \beta_n, \frac{\beta_n}{2}\right) \subset \Omega^-.$$

This ensures that the inf-convolution $V^{\bar{\epsilon}_n}$ at the point $z_n - \beta_n$ do not depend on V in points in Ω^+. Hence

$$|V(z_n - \beta_n) - V^{\bar{\epsilon}_n}(z_n - \beta_n)| \to 0,$$

as $n \to \infty$. Moreover, because $z_n - \beta_n \in \Omega^-$, $V^{\bar{\epsilon}_n}(z_n - \beta_n) \le V^{\bar{\epsilon}_n}(z_n) + O(|\beta_n|)$. Consequently we have

$$\begin{aligned}\inf_{x_n \to x} \liminf_n V(x_n) &\le \liminf_n V(z_n - \beta_n) = \liminf_n V^{\bar{\epsilon}_n}(z_n - \beta_n) \\ &\le \liminf_n V^{\bar{\epsilon}_n}(z_n) + O(|\beta_n|).\end{aligned}$$

Thus,

$$\inf_{x_n \to x} \liminf_n V(x_n) \le \rho + \inf_{\substack{x_n \to x \\ \epsilon_n \to 0}} \liminf_n V^{\epsilon_n}(x_n),$$

which implies $H^*(x,p) \ge H^\sharp(x,p)$, as ρ is arbitrarily small. The other inequality in the statement can be established with a symmetric argument. □

7.1. Existence of viscosity solutions

Proposition 7.4. *Assume H has the form (7.2) and let $\bar{H}$ be the number defined in (5.7). Then there exists a Lipschitz continuous viscosity solution to the equation*

$$H(x, Du(x)) = \bar{H} \quad \text{in } \mathbb{T}^N. \tag{7.6}$$

Proof. The proof is divided in two parts. We first prove that there exists $\hat{H} \in \mathbb{R}$ such that the equation

$$H(x, Du(x)) = \hat{H} \quad \text{in } \mathbb{T}^N. \tag{7.7}$$

has Lipschitz continuous viscosity solutions. Then we prove that $\hat{H} = \bar{H}$.

Consider $H^\epsilon(x,p)$ defined in (7.5). By standard theory, for any ϵ there exists $\bar{H}^\epsilon$ such that the equation

$$H^\epsilon(x, Du) = \bar{H}^\epsilon \tag{7.8}$$

admits a viscosity solution, denoted by u^ϵ. The sequence $\{u^\epsilon\}$ is equi-bounded up to addition of a constant and equi-Lipschitz. We denote by u a uniform limit of u^ϵ as $\epsilon \to 0$.

Notice that $\bar{H}^\epsilon \ge -C$ for some constant C for any ϵ; moreover a direct computation shows that the sequence $\{\bar{H}^\epsilon\}$ is monotone non increasing. Then there exists

$$\lim_{\epsilon \to 0} \bar{H}^\epsilon =: \hat{H}. \tag{7.9}$$

To conclude the proof we show that u is a viscosity solution of (7.7). Because the limit equation is discontinuous we need to use the notion of viscosity solutions introduced in Definition 7.1, but of course from the proof it follows that u is Lipschitz.

Notice first that $H^\epsilon(x,p)$ converges uniformly to $H(x,p)$ as $\epsilon \to 0$, in compact subsets of $(\mathbb{T}^N \backslash \Sigma) \times \mathbb{R}^N$. Then, by stability, u is a solution to (7.7) in $\mathbb{T}^N \backslash \Sigma$. Furthermore u^ϵ is, in particular, a subsolution to (7.8) in $\mathbb{T}^N$. Then, combining Remark 7.2 and Lemma 7.3 we discover that the function

$$u^\sharp(x) := \limsup{}^\sharp u^\epsilon(x) := \limsup_{\substack{y \to x \\ \epsilon \to 0}} u^\epsilon(y)$$

satisfies

$$H_*(x, Du^\sharp) \le H_\sharp(x, Du^\sharp) \le \bar{H}, \quad \text{in } \mathbb{T}^N.$$

Analogously, $u_\sharp := \liminf{}_\sharp u^\epsilon$ satisfies

$$H^*(x, Du_\sharp) \ge H^\sharp(x, Du_\sharp) \ge \bar{H}, \quad \text{in } \mathbb{T}^N.$$

But since u is continuous, $u^\sharp, u_\sharp$ and u agree in $\mathbb{T}^N$. Thus u is a viscosity solution of (7.7) in the sense of Definition 7.1. The first part of the proof is then completed.

It remains to prove that $\hat{H}$ coincides with

$$\bar{H} = -\inf_{\mu \in \mathcal{M}_1^+ \cap \mathcal{M}_{\text{hol}}} \int L(x,v) \, \mathrm{d}\mu.$$

Let $L^\epsilon(x,v)$ be the approximation of $L(x,v)$ defined in (6.3). For any ϵ we have

$$-\bar{H}^\epsilon = \inf_{\mu \in \mathcal{M}_1^+ \cap \mathcal{M}_{\text{hol}}} \int L^\epsilon(x,v) \, \mathrm{d}\mu(x,v) = \int L^\epsilon(x,v) \, \mathrm{d}\mu^\epsilon(x,v),$$

for some measure $\mu^\epsilon \in \mathcal{M}_1^+ \cap \mathcal{M}_{\text{hol}}$. Furthermore, since H^ϵ is uniformly coercive, the supports of μ^ϵ are uniformly bounded. Thus μ^ϵ converges weak-star to some $\bar{\mu}$ up to subsequence. By Theorem 6.1 $\bar{\mu}$ minimizes $\int L \, \mathrm{d}\mu$. Then recalling that $L^\epsilon \le L$ for any ϵ we obtain

$$-\bar{H}^\epsilon = \int L^\epsilon \, \mathrm{d}\mu^\epsilon \le \int L^\epsilon \, \mathrm{d}\bar{\mu} \le \int L \, \mathrm{d}\bar{\mu} = -\bar{H} \quad \text{for any } \epsilon.$$

Passing to the limit as $\epsilon \to 0$ this gives $\hat{H} \ge \bar{H}$. Using Proposition 6.5 we get the opposite inequality:

$$-\bar{H} = \int L \, \mathrm{d}\bar{\mu} \le \liminf_{\epsilon \to 0} \int L^\epsilon \, \mathrm{d}\mu^\epsilon = \liminf_{\epsilon \to 0} -\bar{H}^\epsilon = -\hat{H}.$$

This completes the proof. □

Remark 7.5. A function is locally semiconcave on an open set if it is semiconcave in every compact subset. The solution u of (7.6) is locally semiconcave in $\mathbb{T}^N \backslash \Sigma$, because it is the uniform limit of the sequence of locally semiconcave functions u^ϵ satisfying (7.8), and for each compact subset of $\mathbb{T}^N \backslash \Sigma$ we can obtain an estimate on the semiconcavity constant uniform in ϵ. The semiconcavity constant is, of course, not globally bounded. □

Proposition 7.6. *There exists at most one* $\lambda \in \mathbb{R}$ *such that the equation*

$$H(x, Du) = \lambda \tag{7.10}$$

admits a Lipschitz viscosity solution u *in* $\mathbb{T}^N$.

Proof. The statement is a direct consequence of the comparison between sub- and super-solutions of state-constraint problems. This property has been shown by Soner in [21]; we repeat here the argument for completeness. Assume that u and v are two functions defined on $\mathbb{T}^N$ satisfying respectively

$$H(x, Du) \leq a, \quad H(x, Dv) \geq b, \qquad x \in \mathbb{T}^N.$$

Assume further that v is lower semicontinuous on $\mathbb{T}^N$. We claim that $a \geq b$. Let $\nu(x)$ be a vector field defined on Σ, pointing toward Ω^- and $\eta : \mathbb{T}^N \to \mathbb{R}^N$ such that $\eta(x) = \eta(x)$ for every $x \in \Sigma$. For every small $\epsilon > 0$ define

$$\Phi(x, y) = u(x) - v(y) - \left|\frac{x-y}{\epsilon} - \eta(y)\right|^2, \quad (x, y) \in \mathbb{T}^{2N}.$$

By coercivity of H, u is Lipschitz continuous on $\mathbb{T}^N$, with Lipschitz constant $M > 0$. Then, since v is assumed to be lower semicontinuous, the function Φ achieves a maximum over $\mathbb{T}^{2N}$. Let us denote by (x_ϵ, y_ϵ) the point of maximum of Φ, then $\Phi(x_\epsilon, y_\epsilon) \geq \Phi(y_\epsilon + \epsilon\eta(y_\epsilon), y_\epsilon)$, which yields

$$\left|\frac{x_\epsilon - y_\epsilon}{\epsilon} - \eta(y_\epsilon)\right|^2 \leq u(x_\epsilon) - u(y_\epsilon + \epsilon\eta(y_\epsilon)) \leq \epsilon M \left|\frac{x_\epsilon - y_\epsilon}{\epsilon} - \eta(y_\epsilon)\right|.$$

Hence,

$$\left|\frac{x_\epsilon - y_\epsilon}{\epsilon} - \eta(y_\epsilon)\right| \leq \epsilon M. \tag{7.11}$$

This implies that $x_\epsilon = y_\epsilon + \epsilon\eta(y_\epsilon) + O(\epsilon^2)$, as $\epsilon \to 0$. Then, for ϵ small enough we get the following property:

$$y_\epsilon \in \bar{\Omega}^- \Longrightarrow x_\epsilon \in \Omega^-. \tag{7.12}$$

Set further

$$\phi(x, y) = \left|\frac{x-y}{\epsilon} - \eta(y)\right|^2, \quad (x, y) \in \mathbb{T}^{2N}$$

and observe that

$$\begin{aligned} D_x\phi(x, y) &= \frac{2}{\epsilon}\left(\frac{x-y}{\epsilon} - \eta(y)\right) \\ -D_y\phi(x, y) &= 2\left(\frac{I}{\epsilon} + D\eta(y)^T\right)\left(\frac{x-y}{\epsilon} - \eta(y)\right) \\ &= D_x\phi(x, y) + 2D\eta(y)^T\left(\frac{x-y}{\epsilon} - \eta(y)\right), \end{aligned}$$

where $D\eta(y)^T$ denotes the transposed matrix of $D\eta(y)$. Notice also that, by (7.11), $D_x\phi(x_\epsilon, y_\epsilon) \leq 2M$ and

$$-D_y\phi(x_\epsilon, y_\epsilon) = D_x\phi(x_\epsilon, y_\epsilon) + O(\epsilon), \quad \text{as } \epsilon \to 0.$$

By extracting a subsequence if necessary, we can assume that either $y_\epsilon \in \bar{\Omega}^-$ as $\epsilon \to 0$, or $y_\epsilon \in \Omega^+$. If $y_\epsilon \in \bar{\Omega}^-$, then $x_\epsilon \in \bar{\Omega}^-$. By definition of viscosity solution, in view of (7.12) we get,

$$H^-(y_\epsilon, -D_y\phi(x_\epsilon, y_\epsilon)) \geq b \quad \text{and} \quad H^-(x_\epsilon, D_x\phi(x_\epsilon, y_\epsilon)) \leq a \tag{7.13}$$

Since $y_\epsilon - x_\epsilon \to 0$, and $D_y\phi(x_\epsilon, y_\epsilon) - D_x\phi(x_\epsilon, y_\epsilon) \to 0$ we conclude $a \geq b$. If instead $y_\epsilon \in \Omega^+$ we get

$$H^+(y_\epsilon, -D_y\phi(x_\epsilon, y_\epsilon)) \geq b \tag{7.14}$$

If $x_\epsilon \in \Omega^+$ infinitely often, by extracting a further subsequence we can assume that

$$H^+(x_\epsilon, D_x\phi(x_\epsilon, y_\epsilon)) \leq a,$$

from which we conclude $a \geq b$. Alternatively, $x_\epsilon \in \bar{\Omega}^-$ infinitely often. Then, since $x_\epsilon - y_\epsilon \to 0$, taking into account the discontinuity at Σ we have

$$a \geq H^-(x_\epsilon, D_x\phi(x_\epsilon, y_\epsilon)) \geq H^+(y_\epsilon, D_x\phi(x_\epsilon, y_\epsilon)).$$

Then, using $D_y\phi(x_\epsilon, y_\epsilon) - D_x\phi(x_\epsilon, y_\epsilon) \to 0$ we conclude $a \geq b$. □

Remark 7.7. Putting together Proposition 7.4 and Proposition 7.6 we discover that $\bar{H}$ is the unique value λ for which Eq. (7.10) admits a Lipschitz continuous viscosity solution. □

7.2. Regularity

We know, by Proposition 7.4 that there are Lipschitz solutions of (7.6). In this subsection we are interested in local $L^2(\mathrm{d}\mu_{\mathbb{T}^N\setminus\Sigma})$ estimates for the gradient quotient of solutions of such equation, where μ is a minimizing holonomic measure. The estimates presented in this section are actually a localized version of those in [13].

Given μ probability measure on $\mathbb{T}^N \times \mathbb{R}^N$, we denote as in the preceding section with $\mu_{\mathbb{T}^N}$ its projection onto the state space $\mathbb{T}^N$. We introduce next the push-forward of μ. Since L is strictly convex and superlinear, there exists a one to one correspondence between the Hamiltonian space and the Lagrange space of coordinates, (x, p) and (x, v), respectively. Such correspondence is expressed through the map $\Phi : \mathbb{T}^N \times \mathbb{R}^N \to \mathbb{T}^N \times \mathbb{R}^N$

$$\Phi(x, v) := (x, -D_vL(x, v)).$$

Even though L is discontinuous, since we are assuming hypothesis (**H3**) holds, $D_vL = D_vL_0$ is a smooth diffeomorphism.

The push-forward of any measure μ with respect to Φ is the measure $\mu_\# := \Phi_\#\mu$ defined by:

$$\int \psi(x, p)\ \mathrm{d}\mu_\#(x, p) := \int \psi(x, -D_vL(x, v))\ \mathrm{d}\mu(x, v), \tag{7.15}$$

for every $\psi \in C(\mathbb{T}^N \times \mathbb{R}^N)$ compactly supported.

Next Lemma 7.8 is a version for this problem of the well known fact that, for continuous Hamiltonians, there exist smooth strict subsolutions of the equation (7.6) (Cfr. [15]).

Lemma 7.8. *Let $\bar{H}$ be the number defined in (5.7). Then for any $a > \bar{H}$ there exists u_a smooth such that*

$$H(x, Du_a(x)) \le a \quad in\ \mathbb{T}^N.$$

Proof. For every ϵ, let $\bar{H}^\epsilon$ be the sequence of values such that Eq. (7.8) admits a viscosity solution u_ϵ. Since $\bar{H}^\epsilon$ converges to $\bar{H}$ from above and is monotonic non increasing, $a > \bar{H}^\epsilon$ for every ϵ sufficiently small. Fix such an ϵ; by the usual theory for continuous Hamiltonians, we know that a smooth strict subsolution u_a of $H^\epsilon = a$ there exists. We conclude:

$$a > H^\epsilon(x, Du_a) \ge H(x, Du_a) \quad \text{in } \mathbb{T}^N.$$

□

Proposition 7.9. *Let u be a solution of (7.6) and $\bar{\mu}$ a minimizing holonomic probability measure. Then, for $\bar{\mu}_{\mathbb{T}^N}$-a.e. $x \in \operatorname{supp}(\bar{\mu}_{\mathbb{T}^N})\backslash\Sigma$, $Du(x)$ exists. Moreover*

$$p = Du(x), \qquad for\ \bar{\mu}_{\#}\text{-}a.e.\ (x,p) \in (\operatorname{supp}(\bar{\mu}_{\mathbb{T}^N})\backslash\Sigma) \times \mathbb{R}^N.$$

Proof. Consider the following two sequence of functions. Let $\{u_n\}$ be a sequence of of smooth approximate subsolutions of

$$H(x, Du_n(x)) \le \bar{H} + \frac{1}{n}$$

uniformly converging to u as $n \to +\infty$. Such a sequence exists in force of Lemma 7.8. Consider further the sequence $\{v_n\}$ of functions obtained by convolving the solution u for some mollification kernel η_n of size $1/n$; of course v_n fail to be an approximate subsolutions around the singular set. Fix a radius k_0 and consider a tube Σ_{k_0} around Σ. Define also a function λ on $\mathbb{T}^N$ with $0 \le \lambda \le 1$ in such a way that λ is supported in Σ_{k_0} and $1-\lambda$ is supported in $\mathbb{T}^N\backslash\Sigma_{k_0}$. Finally set

$$w_n(x) := \lambda(x)u_n(x) + (1-\lambda(x))v_n(x).$$

By convexity of H, w_n is a smooth approximate subsolution, that is

$$H(x, Dw_n(x)) \le \bar{H} + \frac{1}{n};$$

moreover it coincides with v_n outside Σ_{k_0}. By (7.1), with $p = Dv_n(y)$ and $q = Dw_n(x)$,

$$\begin{aligned} H(x, Dv_n(y)) \ge{} & H(x, Dw_n(x)) + D_pH(x, Dw_n(x))(Dv_n(y) - Dw_n(x)) \\ & + \frac{\kappa}{2}|Dv_n(y) - Dw_n(x)|^2. \end{aligned}$$

Multiplying by $\eta_n(x-y)$, integrating with respect to y we get

$$\begin{aligned} & H(x, Dw_n(x)) + \frac{\kappa}{2}\int_{\mathbb{T}^N\backslash\Sigma_{k_0}} \eta_n(x-y)|Dv_n(y) - Dw_n(x)|^2 \,\mathrm{d}y \\ & \quad \le \int_{\mathbb{T}^N\backslash\Sigma_{k_0}} \eta_n(x-y)H(x, Dv_n(y))\,\mathrm{d}y \le \bar{H} + \frac{1}{n}. \end{aligned} \tag{7.16}$$

Using again (7.1) we have

$$\begin{aligned}
&\frac{\kappa}{2}\int |Dw_n(x)-p|^2 \, \mathrm{d}\bar\mu_\#(x,p)\\
&\quad\le \int [H(x,Dw_n(x))-H(x,p)-D_pH(x,p)(Dw_n(x)-p)] \, \mathrm{d}\bar\mu_\#(x,p)\\
&\quad= -\bar H + \int H(x,Dw_n(x)) \, \mathrm{d}\bar\mu_\#(x,p).
\end{aligned}\tag{7.17}$$

In the last identity we have used the fact that, by Corollary 5.11,

$$\int [pD_pH(x,p)-H(x,p)] \, \mathrm{d}\bar\mu_\#(x,p) = \int L(x,v) \, \mathrm{d}\bar\mu(x,v) = -\bar H$$

and that, by holonomy,

$$\int D_pH(x,p)Dw_n(x) \, \mathrm{d}\bar\mu_\#(x,p) = 0.$$

Putting together (7.16) and (7.17) we obtain

$$\frac{\kappa}{2}\int |Dw_n(x)-p|^2 \, \mathrm{d}\bar\mu_\#(x,p) + \int \beta_n(x) \, \mathrm{d}\bar\mu_{\mathbb{T}^N}(x) \le \frac{1}{n},\tag{7.18}$$

where we denoted

$$\beta_n(x) := \frac{\kappa}{2}\int_{\mathbb{T}^N\setminus\Sigma_{k_0}} \eta_n(x-y)|Dv_n(y)-Dw_n(x)|^2 \, \mathrm{d}y.$$

Then for $\bar\mu_{\mathbb{T}^N}$-a.e. $x, \beta_n(x)\to 0$ as $n\to+\infty$. Now, notice that, by (7.18), $Dw_n(x)$ converges to p in $L^2(d\bar\mu_\#)$. Moreover, $\bar\mu_{\mathbb{T}^N}$-a.e. x in $\mathbb{T}^N\setminus\Sigma_{k_0}$ is a point of approximating continuity for Du. Then, since w_n and v_n agree outside Σ_{k_0}, $Dw_n(x)$ converges to $Du(x)$ for $\bar\mu_{\mathbb{T}^N}$-a.e. x in $\mathbb{T}^N\setminus\Sigma_{k_0}$. This in turn entails $p = Du(x)$. The statement is then established, as k_0 can be chosen arbitrarily small. □

Remark 7.10. By arguing as in [4, Corollary 33], we can rephrase the statement of Proposition 7.9 by saying that $D_xu(x)$ exists for $\bar\mu_{\mathbb{T}^N}$-a.e. $x\in\mathbb{T}^N\setminus\Sigma$ and satisfies

$$D_vL(x,v) = D_xu(x), \quad \bar\mu\text{-a.e. in } (\mathbb{T}^N\setminus\Sigma)\times\mathbb{R}^N.\tag{7.19}$$

Moreover, we will prove, under an additional assumption, that Σ is negligible with respect to $\bar\mu_{\mathbb{T}_N}$; see Remark 8.9 and Lemma 8.11 in the next Section. Thus we can assume, for our purposes, that (7.19) holds $\bar\mu$-a.e. in $\mathbb{T}^N\times\mathbb{R}^N$.

Another important consequence of the previous Proposition is the following. In Proposition 5.5 we proved that a minimizing holonomic measure is supported on the graph of a function $\mathbf{v}$ which is defined, in principle, on $\mathrm{supp}(\bar\mu)$; see Remark 5.6. But then, thanks to Proposition 7.9, there is a canonical extension Lebesgue a.e. in $\mathbb{T}^N$ given by

$$\mathbf{v}(x) = -D_pH(x,D_xu(x)).$$

In particular this implies $|\mathbf{v}|$ bounded. □

Remark 7.11. By (7.6) and Proposition 7.9 immediately follows that

$$H(x,p) - \bar{H} = 0 \quad \bar{\mu}_{\#}(x,p) \text{ a.e. in } (\mathbb{T}^N \backslash \Sigma) \times \mathbb{R}^N.$$

□

In the next proposition we provide an L^2_{loc} estimate of the gradient quotient of a solution of (7.6).

Proposition 7.12. *Assume H has the form (7.2). Let u be a viscosity solution of (7.6), $\bar{\mu}$ a minimizing holonomic measure and $\bar{\mu}_{\mathbb{T}^N}$ its projection onto $\mathbb{T}^N$. Then there exists $C > 0$ such that, for any $\delta > 0$ and any C^∞ cutoff function φ supported in $\mathbb{T}^N \backslash \Sigma_\delta$,*

$$\int_{\mathbb{T}^N} \varphi^2(x) |D_x u(x+h) - D_x u(x)|^2 \, \mathrm{d}\bar{\mu}_{\mathbb{T}^N} \leq C|h|^2,$$

for any $h \in \mathbb{R}^N$ with $|h| < \delta/2$.

The proof of Proposition 7.12 makes use of the following:

Lemma 7.13. *Let u be a solution to (7.6) and u^ϵ the convolution of u with a standard mollifier η_ϵ supported in $B(0,\epsilon), \epsilon > 0$. Then there exists $C > 0$ such that, for any $\delta > 0$,*

$$H(x, Du^\epsilon(x)) \leq \bar{H} + C\epsilon \quad \text{in } \mathbb{T}^N \backslash \Sigma_\delta. \tag{7.20}$$

if ϵ is sufficiently small.

Proof. Recall first that, by Remark 7.10 $|Du(x)| \leq M$ for μ-a.e. $x \in \mathbb{T}^N \setminus \Sigma$, for a minimizing holonomic measure μ. Then we set

$$C := \sup\{|D_x H(x,p)| \,:\, x \in \mathbb{T}^N \backslash \Sigma,\ |p| \leq M\} < \infty.$$

For any $x \in \mathbb{T}^N \backslash \Sigma_\delta$ and $y \in B(x,\epsilon)$ with $\epsilon < \delta$, we have

$$|H(x,p) - H(x-y,p)| \leq C|y|, \quad \text{for any } |p| \leq M.$$

Thus, by Jensen's inequality

$$\begin{aligned} \bar{H} &\geq \int \eta_\epsilon(y) H(x, Du(x-y)) \, \mathrm{d}y - C \int |y| \eta_\epsilon(y) \, \mathrm{d}y \\ &\geq H\left(x, \int \eta_\epsilon(y) Du(x-y) \, \mathrm{d}y\right) - C \int |y| \eta_\epsilon(y) \, \mathrm{d}y \\ &= H(x, Du^\epsilon(x)) - C \int |y| \eta_\epsilon(y) \, \mathrm{d}y \geq H(x, Du^\epsilon(x)) - C, \end{aligned}$$

that is (7.20) holds. □

Proof of Proposition 7.12. Let u^ϵ be the convolution of u with a standard smooth mollifier supported in $B(0,\epsilon)$. Let $x \in \mathbb{T}^N \backslash \Sigma_\delta$ and $h \in \mathbb{R}^N$ with $|h| < \delta/2$. Then $x \pm h \notin \Sigma_{\delta/2}$ and, by Lemma 7.13, for ϵ sufficiently small,

$$H(x \pm h, Du^\epsilon(x \pm h)) \leq C\epsilon + \bar{H}.$$

Thus,

$$\begin{aligned} &H(x, Du^\epsilon(x \pm h)) - H(x, Du(x)) \\ &\quad \leq C\epsilon + H(x, Du^\epsilon(x \pm h)) - H(x \pm h, Du^\epsilon(x \pm h)) \end{aligned}$$

and consequently, by convexity (7.1),

$$\begin{aligned}
&\frac{\kappa}{2}|Du^\epsilon(x\pm h)-Du(x)|^2+D_pH_0(x,Du(x))(Du^\epsilon(x\pm h)-Du(x))\\
&\quad\le H(x,Du^\epsilon(x\pm h))-H(x,Du(x))\\
&\quad\le C\epsilon+H_0(x,Du^\epsilon(x\pm h))-H_0(x\pm h,Du^\epsilon(x\pm h))-V(x)+V(x\pm h)\\
&\quad\le C(\epsilon+|h|^2)\pm h\cdot DV(x)\pm h\cdot D_xH_0(x,Du^\epsilon(x\pm h)).
\end{aligned}$$

Now notice that, for some $\xi\in[Du^\epsilon(x\pm h),Du(x)]$,

$$\begin{aligned}
&\pm h\cdot D_xH_0(x,Du^\epsilon(x\pm h))\\
&\quad=\pm h\cdot D_xH_0(x,Du(x))\pm h\cdot D^2_{x,p}H_0(x,\xi)\cdot(Du^\epsilon(x\pm h)-Du(x));
\end{aligned}$$

then, observing that $|D^2_{xp}H_0(x,\xi)|\le C$ for some constant $C>0$, for any $\xi\in[Du^\epsilon(x\pm h),Du(x)]$ and any ϵ, and using a Cauchy inequality weighted with $\beta>0$, we can write

$$\begin{aligned}
&\pm h\cdot D_xH_0(x,Du^\epsilon(x\pm h))\\
&\quad\le\pm h\cdot D_xH_0(x,Du(x))+\frac{1}{\beta}|h|^2+\beta|Du^\epsilon(x\pm h)-Du(x)|^2.
\end{aligned}$$

Then, by choosing β appropriately small,

$$\begin{aligned}
&C|Du^\epsilon(x\pm h)-Du(x)|^2\\
&\quad\le C(\epsilon+|h|^2)\pm h\cdot DV(x)\pm h\cdot D_xH_0(x,Du(x))\\
&\qquad-D_pH_0(x,Du(x))(Du^\epsilon(x\pm h)-Du(x))
\end{aligned}$$

and, after some cancellations, we get

$$\begin{aligned}
&|Du^\epsilon(x+h)-Du(x)|^2+|Du^\epsilon(x-h)-Du(x)|^2\\
&\quad\le C(\epsilon+|h|^2)-CD_pH_0(x,Du(x))(Du^\epsilon(x+h)-2Du(x)+Du^\epsilon(x-h)).
\end{aligned}$$

Multiplying by φ^2 and integrating with respect to $\bar\mu_{\mathbb{T}^N}(x)$ we have

$$\begin{aligned}
&\int_{\mathbb{T}^N}\varphi^2|Du^\epsilon(x+h)-Du(x)|^2\,\mathrm{d}\bar\mu_{\mathbb{T}^N}\\
&\quad\le C(\epsilon+|h|^2)\\
&\qquad-C\int_{\mathbb{T}^N}\varphi^2D_pH_0(x,Du(x))\cdot(Du^\epsilon(x+h)-2Du(x)+Du^\epsilon(x-h))\,\mathrm{d}\bar\mu_{\mathbb{T}^N}.
\end{aligned}\tag{7.21}$$

Now, integrating by parts we can rewrite the integral in the right hand side of the previous inequality as

$$\begin{aligned}
&-\int_{\mathbb{T}^N}\varphi^2D_pH_0(x,Du(x))\cdot(Du^\epsilon(x+h)-2Du(x)+Du^\epsilon(x-h))\,\mathrm{d}\bar\mu_{\mathbb{T}^N}(x)\\
&\quad=\int_{\mathbb{T}^N}2\varphi D\varphi D_pH_0(x,Du(x))\cdot[u^\epsilon(x+h)-2u(x)+u^\epsilon(x-h)]\,\mathrm{d}\bar\mu_{\mathbb{T}^N}(x).
\end{aligned}$$

Moreover, since u^ϵ is semiconcave, being the convolution of a smooth mollifier with u which is semiconcave (see Remark 7.5), we have

$$u^\epsilon(x+h)-2u(x)+u^\epsilon(x-h)\le C|h|^2+2|u^\epsilon-u|_\infty.$$

From this, taking into account that $|D_pH_0(x,Du(x))|$ is bounded because u is Lipschitz, that $|\varphi|$ and $|D\varphi|$ are also bounded in $\mathbb{T}^N$, and that u^ϵ converges uniformly to u, we conclude after (7.21) that

$$\int_{\mathbb{T}^N} \varphi^2 |Du^\epsilon(x+h) - Du(x)|^2 \, \mathrm{d}\bar{\mu}_{\mathbb{T}^N}(x) \le C(\epsilon + |h|^2).$$

We send $\epsilon \to 0$. Up to subsequences, $Du^\epsilon(x+h) \to \mathbf{p}(x)$ weakly in $L^2_{\mathrm{loc}}(\mathrm{d}\bar{\mu}_{\mathbb{T}^N})$ and

$$\int_{\mathbb{T}^N} \varphi^2 |\mathbf{p} - Du|^2 \, \mathrm{d}\bar{\mu}_{\mathbb{T}^N} \le C|h|^2.$$

To complete the proof we must show that $\mathbf{p} \in Du(x+h)$ $\bar{\mu}_{\mathbb{T}^N}$-a.e. namely, that for $\bar{\mu}_{\mathbb{T}^N}$-a.e. x there exists $C > 0$ such that

$$u(y+h) \le u(x+h) + \mathbf{p}\cdot(y-x) + C|y-x|^2 \quad \text{for any } y. \tag{7.22}$$

First note that $u^\epsilon(\cdot + h)$ is semiconcave, that is

$$u^\epsilon(y+h) \le u^\epsilon(x+h) + Du^\epsilon(x)\cdot(y-x) + C|y-x|^2, \quad \text{for any } x, y.$$

Then, for any nonnegative function $g \in L^2(\mathrm{d}\bar{\mu}_{\mathbb{T}^N})$,

$$0 \le \int [-u^\epsilon(y+h) + u^\epsilon(x+h) + Du^\epsilon(x)\cdot(y-x) + C|y-x|^2] g(x) \, \mathrm{d}\bar{\mu}_{\mathbb{T}^N}.$$

By sending $\epsilon \to 0$ and taking into account that $u^\epsilon \to u$ uniformly,

$$0 \le \int [-u(y+h) + u(x+h) + \mathbf{p}\cdot(y-x) + C|y-x|^2] g(x) \, \mathrm{d}\bar{\mu}_{\mathbb{T}^N}.$$

Since g is arbitrary, (7.22) is then established, and the proof is completed. □

8. Properties of minimizing measures

We study now some examples and additional properties of minimizing measures. As in the previous Sections, we assume hypothesis (**H3**) concerning the structure of L.

Definition 8.1. Let Ω be an open subset of $\mathbb{T}^N \times \mathbb{R}^N$. We say that a measure μ is *invariant in* Ω under the flow generated by the Euler–Lagrange equation if for any function $\phi \in C^1(\mathbb{T}^N \times \mathbb{R}^N)$, compactly supported in Ω,

$$\int_{\mathbb{T}^N \times \mathbb{R}^N} \chi(x,v) D\phi(x,v) \, \mathrm{d}\mu(x,v) = 0,$$

where χ is the vector filed corresponding to the Euler–Lagrange equation, or more explicitly, if

$$\int v_k \frac{\partial \phi}{\partial x_k} + \frac{\partial \phi}{\partial v_j} \left(\frac{\partial^2 L}{\partial v}\right)^{-1}_{js} \left[\frac{\partial L}{\partial x_s} - v_k \frac{\partial^2 L}{\partial x_k \partial v_s}\right] \, \mathrm{d}\mu(x,v) = 0. \tag{8.1}$$

Remark 8.2. It is sometimes convenient to use the following equivalent definition of invariance (see [4] and [17]), that is, μ is invariant in Ω under the flow generated by the Euler–Lagrange equation if and only if for any function $\phi \in C^1(\mathbb{T}^N \times \mathbb{R}^N)$, compactly supported in Ω,

$$\int_{\mathbb{T}^N \times \mathbb{R}^N} \{\phi, H\}\ \mathrm{d}\mu_{\#}(x,p) = 0, \tag{8.2}$$

where $\mu_{\#}$ is the push-forward of μ, defined in (7.15), and the notation $\{\cdot,\cdot\}$ stands for the Poisson bracket, which is defined, for any couple of functions $A(x,p), B(x,p)$ in $C^1(\mathbb{T}^N \times \mathbb{R}^N)$ as

$$\{A, B\} := D_pA \cdot D_xB - D_xA \cdot D_pB.$$

□

Remark 8.3. Note that if $\mu_{\#}$ does not give mass to the set where H is non-differentiable, the integral in (8.2) is well defined if we use the convention that we identify $\{\phi, H\}$ with a function defined almost everywhere in the support of $\mu_{\#}$ and not as a distribution. □

8.1. Examples

A minimizing measure not giving mass to the singular set. We continue the study of the example of Sect. 5.4. With $\bar{\mu}$ we denote the measure obtained in (5.18), and with $\bar{\mu}_{\#}$ its push-forward. Observe that $\bar{\mu}(\Sigma \times \mathbb{R}^N) = 0$. We have $H(x,p) = \frac{|p|^2}{2} - V(x)$. It is easy to check that $\bar{\mu}$ is invariant outside Σ. In fact, if $\phi \in C^1([0,1] \times \mathbb{R})$, compactly supported in $([0,1]\backslash\Sigma) \times \mathbb{R}$,

$$\int_{[0,1]\times\mathbb{R}} \{H, \phi\}\ \mathrm{d}\bar{\mu}_{\#}(x,p) = 0,$$

because $H(x,p)$ is differentiable in $([0,1]\backslash\Sigma) \times \mathbb{R}$; see Remark 8.3.

In the general case of a ϕ differentiable and compactly supported in $[0,1]\times\mathbb{R}$, taking into account (7.15) and the representation (5.18) for $\bar{\mu}$ we have,

$$\begin{aligned}
&\int_{([0,1]\backslash\Sigma)\times\mathbb{R}} D_pH(x,p) \cdot D_x\phi(x,p)\ \mathrm{d}\bar{\mu}_{\#}(x,p) \\
&\quad = \int_{([0,1]\backslash\Sigma)\times\mathbb{R}} v \cdot D_x\phi(x,v)\ \mathrm{d}\bar{\mu}(x,v) \\
&\quad = -\int_{[0,\frac{1}{3})} D_x\phi(x,\bar{v}_1)\ \mathrm{d}x - \int_{(\frac{1}{3},\frac{2}{3})} D_x\phi(x,\bar{v}_2)\ \mathrm{d}x - \int_{(\frac{2}{3},1)} D_x\phi(x,\bar{v}_1)\ \mathrm{d}x \\
&\quad = \phi\left(\frac{1}{3},\bar{v}_2\right) + \phi\left(\frac{2}{3},\bar{v}_1\right) - \phi\left(\frac{1}{3},\bar{v}_1\right) - \phi\left(\frac{2}{3},\bar{v}_2\right)
\end{aligned}$$

where $\bar{v}_1$ and $\bar{v}_2$ are the two values in (5.18). Furthermore,

$$\int_{([0,1]\backslash\Sigma)\times\mathbb{R}} D_xH(x,p)D_p\phi(x,p)\ \mathrm{d}\mu_{\#} = 0,$$

because $D_xH = 0$ on $([0,1]\backslash\Sigma) \times \mathbb{R}$. Thus, we can write

$$\int_{([0,1]\backslash\Sigma)\times\mathbb{R}} \{H, \phi\}\ \mathrm{d}\bar{\mu}_{\#} = \int_{\Sigma\times\mathbb{R}} \phi\, \mathrm{d}\sigma^+ - \int_{\Sigma\times\mathbb{R}} \phi\, \mathrm{d}\sigma^-, \tag{8.3}$$

where

$$\sigma^+ := \delta_{(\frac{1}{3},\bar{v}_2)} + \delta_{(\frac{1}{3},\bar{v}_1)}, \quad \sigma^- := \delta_{(\frac{1}{3},\bar{v}_1)} + \delta_{(\frac{2}{3},\bar{v}_2)}.$$

Observe that σ^+ and σ^- are not probability measures.

A minimizing measure giving mass to the singular set. In the previous example the minimizing measure is supported away from the singular set. We consider now an example in which the Mather measure is concentrated on the singularity.

We consider again the Lagrangian $L(x, v) = \frac{|v|^2}{2} + V(x)$, $(x, v) \in [0, 1]\times\mathbb{R}$; we are assuming now $V(x)$ to be lower semicontinuous, with only one jump at the point of minimum $x_0 \in (0, 1)$, and smooth elsewhere. We further assume that the minimum is strict. It is easy to check that in this case the minimizing measure is $\bar{\mu}(x, v) = \delta_0(v)\delta_{x_0}(x)$. In fact,

$$\int_{[0,1]\times\mathbb{R}} L(x, v)\, \mathrm{d}\bar{\mu}(x, v) = V(x_0) = \min_{x\in[0,1]} V(x).$$

As in the previous example, if ϕ is a C^1 function compactly supported away from x_0, we have

$$\int_{[0,1]\times\mathbb{R}} \{H, \phi\}\, \mathrm{d}\bar{\mu}_\#(x, p) = \int_{[0,1]\times\mathbb{R}} V'(x)D_p\phi(x, p)\, \mathrm{d}\bar{\mu}_\#(x, p) = 0.$$

Thus $\bar{\mu}$ is invariant on $[0, 1]\backslash\{x_0\}$, under the Euler–Lagrange flow, which is obvious since $\bar{\mu}$ gives no mass to that set. However because $\bar{\mu}_\#$ gives mass 1 to the set where H is not differentiable, the integral $\int\{\phi, H\}\, \mathrm{d}\bar{\mu}_\#$ is not defined if $x_0 \in \mathrm{supp}\bar{\mu}_\#$. In this case is unclear whether an analog to (8.3) can be established.

8.2. Invariance of minimizing measures outside the singular set

The goal of this Section is to prove the following

Theorem 8.4. *Assume* (**H3**) *holds. Let* $\bar{\mu}$ *be a minimizing holonomic measure. Then* $\bar{\mu}$ *is invariant under the Euler–Lagrange flow in* $(\mathbb{T}^N\backslash\Sigma) \times \mathbb{R}^N$.

The result is a direct application of the following

Proposition 8.5. *Assume* (**H3**) *holds. Let* $\bar{\mu}$ *be a minimizing holonomic measure and* $\mathbf{v}(x)$ *be a function as in Proposition* 5.5 *whose graph supports* $\bar{\mu}$. *Let* v^ϵ *be the convolution of* $\mathbf{v}$ *with a smooth mollifier. Then* v^ϵ *converges to* $\mathbf{v}$ *in* $L^2_{\mathrm{loc}}(\mathrm{d}\bar{\mu}_{\mathbb{T}^N\backslash\Sigma})$ *as* $\epsilon \to 0$.

Proof. First of all observe that, by Remark 7.10,

$$\mathbf{v}(x) = -D_pH(x, Du(x)), \tag{8.4}$$

where u is a solution to (7.6). The previous identity holds for $\bar{\mu}_{\mathbb{T}^N}$-a.e. x, and also for Lebesgue a.e x. Let η_ϵ be a smooth mollifier supported in $B(0, \epsilon/2)$. By (8.4) we can then write

$$v^\epsilon(x) - \mathbf{v}(x) = \int_{\mathbb{T}^N} [D_pH(x + y, Du(x + y)) - D_pH(x, Du(x))]\, \eta_\epsilon(y)\, \mathrm{d}y.$$

Let $\varphi(x)$ be a cutoff function compactly supported in $\mathbb{T}^N\setminus\Sigma_\epsilon$. Adding and subtracting the quantity $D_pH(Du(x+y),x)$ in the right hand side of the previous line, multiplying by φ, taking the square, integrating with respect to $\bar{\mu}_{\mathbb{T}^N}(x)$ and using Jensen's inequality we get

$$\begin{aligned}
&\int_{\mathbb{T}^N} \varphi^2(x)|v^\epsilon(x)-\mathbf{v}(x)|^2 \ \mathrm{d}\bar{\mu}_{\mathbb{T}^N}\\
&\quad\le C\int_{\mathbb{T}^N\times\mathbb{T}^N} \varphi^2(x)|D_pH(x,Du(x+y))-D_pH(x,Du(x))|^2\eta_\epsilon(y)\,\mathrm{d}y\,\mathrm{d}\bar{\mu}_{\mathbb{T}^N}\\
&\qquad+C\int_{\mathbb{T}^N\times\mathbb{T}^N} \varphi^2(x)|D_pH(x+y,Du(x+y))\\
&\qquad-D_pH(x,Du(x+y))|^2\eta_\epsilon(y)\ \mathrm{d}y\ \mathrm{d}\bar{\mu}_{\mathbb{T}^N},
\end{aligned}$$

for some constant $C>0$. Let us denote by I_1 and I_2 respectively the first and the second integral in the right hand side of the previous inequality.

By the mean value theorem we have

$$\begin{aligned}
&|D_pH(x,Du(x+y))-D_pH(x,Du(x))|\\
&\quad\le \|D^2_{pp}H(x,\xi)\||Du(x+y)-Du(x)|,
\end{aligned}$$

for some $\xi\in[Du(x),Du(x+y)]$. Then, after observing that

$$\sup_{\substack{x\in\mathbb{T}^N\\ p\in B(0,|Du|_\infty)}} \|D^2_{pp}H(x,p)\|\le C$$

and by applying the estimate of Proposition 7.12 we obtain

$$\begin{aligned}
I_1&\le C\int_{\mathbb{T}^N\times\mathbb{T}^N} \varphi^2(x)|Du(x+y)-Du(x)|^2\eta_\epsilon(y)\,\mathrm{d}y\,\mathrm{d}\bar{\mu}_{\mathbb{T}^N}(x)\\
&\le C\int_{B(0,\epsilon/2)}|y|^2\eta_\epsilon(y)\,\mathrm{d}y\le C\epsilon^2.
\end{aligned}$$

We claim that I_2 is also bounded by $C\epsilon^2$. To see this notice that, using again the mean value theorem, we can write

$$I_2\le\int_{\mathbb{T}^N\times(\mathbb{T}^N\setminus\Sigma)} \varphi^2(x)\|D^2_{xp}H(z,D_xu(x+y))\|^2|y|^2\eta_\epsilon(y)\,\mathrm{d}y\,\mathrm{d}\bar{\mu}_{\mathbb{T}^N}(x)$$

for some $z\in[x,x+y], t\in[0,1]$. Observe also that

$$\sup_{\substack{z\in\mathbb{T}^N\setminus\Sigma_{\epsilon/2}\\ p\in B(0,|Du|_\infty)}} \|D^2_{xp}H(z,p)\|\le C,\quad \text{for any } \epsilon.$$

So $I_2\le C\epsilon^2$ as claimed. In conclusion we have

$$\int_{\mathbb{T}^N} \varphi^2(x)|v^\epsilon(x)-\mathbf{v}(x)|^2\ \mathrm{d}\bar{\mu}_{\mathbb{T}^N}(x)\le C\epsilon^2.$$

as desired. □

To prove Theorem 8.4 we need the following two technical results whose proofs are postponed in the Appendix 8.3.

Lemma 8.6. *Let $\bar{\mu}$ be a minimizing holonomic measure, v^ϵ as in Proposition 8.5 and $\phi(x,v)$ any smooth function compactly supported in $(\mathbb{T}^N \backslash \Sigma) \times \mathbb{R}^N$. Then*

$$\int \left\{ v_k \frac{\partial \phi}{\partial x_k}(x, v^\epsilon(x)) + \frac{\partial \phi}{\partial v_j}(x, v^\epsilon(x)) \left(\frac{\partial^2 L}{\partial v} \right)^{-1}_{js} (x, v^\epsilon(x)) \right.$$
$$\left. \times \left[\frac{\partial L}{\partial x_s}(x,v) - v_k \frac{\partial^2 L}{\partial x_k \partial v_s}(x, v^\epsilon(x)) \right] \right\} \mathrm{d}\bar{\mu}$$
$$= \int v_k \left(\frac{\partial L}{\partial v_s}(x, v^\epsilon(x)) - \frac{\partial L}{\partial v_s}(x,v) \right) \frac{\partial}{\partial x_k} \xi^\epsilon_s \, \mathrm{d}\bar{\mu}, \tag{8.5}$$

where

$$\xi^\epsilon_s(x) := \frac{\partial \phi}{\partial v_j}(x, v^\epsilon(x)) \left(\frac{\partial^2 L}{\partial v} \right)^{-1}_{js} (x, v^\epsilon(x)). \tag{8.6}$$

Proof. See Appendix 8.3 □

Lemma 8.7. *Assume (**H3**) holds. Let $\bar{\mu}$ be a minimizing holonomic measure and ξ^ϵ the vector field defined in (8.6). Then there exists $C > 0$ such that, for any $\delta > 0$ sufficiently small,*

$$\int_{\mathbb{T}^N \backslash \Sigma_\delta} \left| \frac{\partial}{\partial x_k} \xi^\epsilon(x) \right|^2 \, \mathrm{d}\bar{\mu}_{\mathbb{T}^N}(x) \leq C,$$

for any k, and any ϵ small enough.

Proof. See Appendix 8.3. □

Proof of Theorem 8.4. According to Definition 8.1 we must prove that

$$\int v_k \frac{\partial \phi}{\partial x_k} + \frac{\partial \phi}{\partial v_j} \left(\frac{\partial^2 L}{\partial v} \right)^{-1}_{js} \left[\frac{\partial L}{\partial x_s} - v_k \frac{\partial^2 L}{\partial x_k \partial v_s} \right] \mathrm{d}\bar{\mu} = 0 \tag{8.7}$$

for any smooth function $\phi(x,v)$ compactly supported in $(\mathbb{T}^N \backslash \Sigma_\delta) \times \mathbb{R}^N$, for any δ arbitrarily small.

Let v^ϵ be as in Proposition 8.5 and Lemma 8.6. The idea is to approximate the left hand side of (8.7) with the left hand side of (8.5) and then to prove that the right hand side of (8.5) converges to zero as $\epsilon \to 0$. To see this we first approximate separately each summand in the left hand side of (8.7) with the correspondent term in the left hand side of (8.5). To do this we use repeatedly Proposition 8.5.

Fix $\delta > 0$ and assume that the function ϕ used in the definition of ξ^ϵ, (8.6), is compactly supported in $(\mathbb{T}^N \backslash \Sigma_\delta) \times \mathbb{R}^N$. For any $\beta > 0$ we have

$$\left| \int_{(\mathbb{T}^N \backslash \Sigma_\delta) \times \mathbb{R}^N} v_k \left[\frac{\partial \phi}{\partial x_k}(x, v^\epsilon(x)) - \frac{\partial \phi}{\partial x_k}(x, \mathbf{v}(x)) \right] \mathrm{d}\bar{\mu} \right|$$
$$\leq \left| \mathrm{Lip} \frac{\partial \phi}{\partial x_k} \right| \int_{(\mathbb{T}^N \backslash \Sigma_\delta) \times \mathbb{R}^N} |v_k| \, |v^\epsilon - \mathbf{v}| \, \mathrm{d}\bar{\mu}$$
$$\leq \beta \left| \mathrm{Lip} \frac{\partial \phi}{\partial x_k} \right|^2 |v_k|^2_\infty + \frac{1}{\beta} \int_{(\mathbb{T}^N \backslash \Sigma_\delta) \times \mathbb{R}^N} |v^\epsilon - \mathbf{v}|^2 \, \mathrm{d}\bar{\mu}. \tag{8.8}$$

Remind (see Remark 7.19) that $|v_k|$ is bounded because $|\mathbf{v}|$ is so.

Analogously

$$\left|\int_{(\mathbb{T}^N\setminus\Sigma_\delta)\times\mathbb{R}^N} \frac{\partial L}{\partial x_s}(x,v)\left[\frac{\partial \phi}{\partial v_j}(x,v^\epsilon(x))\left(\frac{\partial^2 L}{\partial v}\right)^{-1}_{js}(x,v^\epsilon(x))\right.\right.$$
$$\left.\left. - \frac{\partial \phi}{\partial v_j}(x,\mathbf{v}(x))\left(\frac{\partial^2 L}{\partial v}\right)^{-1}_{js}(x,\mathbf{v}(x))\right] \, \mathrm{d}\bar{\mu}\right|$$
$$\leq \beta K_1^2 \int \left|\frac{\partial L}{\partial x_s}\right|^2 \, \mathrm{d}\bar{\mu} + \frac{1}{\beta}\int_{(\mathbb{T}^N\setminus\Sigma_\delta)\times\mathbb{R}^N} |v^\epsilon - \mathbf{v}|^2 \, \mathrm{d}\bar{\mu}, \tag{8.9}$$

where K_1 denotes the Lipschitz constant of the function

$$v \mapsto \frac{\partial \phi}{\partial v_j}(x,v)\left(\frac{\partial^2 L}{\partial v}\right)^{-1}_{js}(x,v),$$

and

$$\left|\int_{(\mathbb{T}^N\setminus\Sigma_\delta)\times\mathbb{R}^N} v_k\left[\frac{\partial \phi}{\partial v_j}(x,v^\epsilon(x))\left(\frac{\partial^2 L}{\partial v}\right)^{-1}_{js}(x,v^\epsilon(x))\frac{\partial^2 L}{\partial x_k \partial v_s}(x,v^\epsilon(x))\right.\right.$$
$$\left.\left. -\frac{\partial \phi}{\partial v_j}(x,\mathbf{v}(x))\left(\frac{\partial^2 L}{\partial v}\right)^{-1}_{js}(x,\mathbf{v}(x))\frac{\partial^2 L}{\partial x_k \partial v_s}(x,\mathbf{v}(x))\right] \, \mathrm{d}\bar{\mu}\right|$$
$$\leq \beta K_2^2 |v_k|_\infty^2 + \frac{1}{\beta}\int_{(\mathbb{T}^N\setminus\Sigma_\delta)\times\mathbb{R}^N} |v^\epsilon - \mathbf{v}|^2 \, \mathrm{d}\bar{\mu}, \tag{8.10}$$

where K_2 denotes the Lipschitz constant of the function

$$v \mapsto \left(\frac{\partial \phi}{\partial v_j}\left(\frac{\partial^2 L}{\partial v}\right)^{-1}_{js}\frac{\partial^2 L}{\partial x_k \partial v_s}\right)(x,v).$$

Putting together (8.8), (8.9) and (8.10), and passing to the limit first for $\epsilon \to 0$ and taking into account Proposition 8.5, and then for $\beta \to 0$, we see that the left hand side of (8.5) approximates the left hand side of (8.7) as $\epsilon \to 0$.

We claim now that the right hand side of (8.5) converges to zero as $\epsilon \to 0$. To see this, taking into account the very definition of ξ_s^ϵ, (8.6), notice that for any $\beta > 0$ we have

$$\left|\int_{(\mathbb{T}^N\setminus\Sigma_\delta)\times\mathbb{R}^N} v_k\left(\frac{\partial L}{\partial v_s}(x,v^\epsilon(x)) - \frac{\partial L}{\partial v_s}(x,v)\right)\frac{\partial}{\partial x_k}\xi_s^\epsilon \, \mathrm{d}\bar{\mu}\right|$$
$$\leq \left(\mathrm{Lip}\frac{\partial L}{\partial v_s}\right)\int_{(\mathbb{T}^N\setminus\Sigma_\delta)\times\mathbb{R}^N} |v_k||v^\epsilon(x) - \mathbf{v}(x)|\left|\frac{\partial}{\partial x_k}\xi_s^\epsilon\right| \, \mathrm{d}\bar{\mu}$$
$$\leq \beta\left(\mathrm{Lip}\frac{\partial L}{\partial v_s}\right)^2\int_{(\mathbb{T}^N\setminus\Sigma_\delta)\times\mathbb{R}^N} |v_k|^2\left|\frac{\partial}{\partial x_k}\xi_s^\epsilon\right|^2 \, \mathrm{d}\bar{\mu}$$
$$+\frac{1}{\beta}\int_{(\mathbb{T}^N\setminus\Sigma_\delta)\times\mathbb{R}^N} |v^\epsilon(x) - \mathbf{v}(x)|^2 \, \mathrm{d}\bar{\mu}.$$

By Lemma 8.7,

$$\int_{\mathbb{T}^N\setminus\Sigma_\delta}\left|\frac{\partial}{\partial x_k}\xi^\epsilon(x)\right|^2\ \mathrm{d}\bar{\mu}_{\mathbb{T}^N}(x)\le C,$$

for any ϵ small enough. Thus, since $|v_k|$ is bounded,

$$\left|\int v_k\left(\frac{\partial L}{\partial v_s}(x,v^\epsilon(x))-\frac{\partial L}{\partial v_s}(x,v)\right)\frac{\partial}{\partial x_k}\xi_s^\epsilon\ \mathrm{d}\bar{\mu}\right|$$
$$\le C\beta|v_k|^2_\infty+\frac{1}{\beta}\int|v^\epsilon(x)-\mathbf{v}(x)|^2\ \mathrm{d}\bar{\mu}.$$

The claimed convergence to zero of the right hand side of (8.5) then follows by sending first $\epsilon\to 0$ and using Proposition 8.5, and finally sending $\beta\to 0$.

The proof is then fulfilled, thanks to Lemma 8.6 and the arbitrariness of δ. □

8.3. Properties of minimizing measures on the singular set

For any $k>0$ small enough consider the signed distance from Σ in Σ_k, i.e. the function w_k solution to the problem

$$\begin{cases}|Dw_k|=1, & \text{in } \Sigma_k\setminus\Sigma\\ w_k=0, & \text{on } \Sigma.\end{cases}\tag{8.11}$$

For any k, w_k is smooth and, for any $x\in\Sigma, Dw_k(x)$ coincides with $\nu(x)$, the unit normal to Σ pointing towards Ω^-. Thus, w_k extends ν smoothly in an appropriate neighborhood of Σ. Furthermore, we can assume without loss of generality that

$$|Dw_k|\le Ck,\quad \text{in } \mathbb{T}^N\setminus\Sigma_k.\tag{8.12}$$

We then denote

$$\tilde{\nu}_k(x)=\begin{cases}\nu(x) & \text{if } x\in\Sigma\\ Dw_k(x) & \text{if } x\in\Sigma_k\setminus\Sigma.\end{cases}$$

Remind (see Remark 5.4) that a minimizing measure $\bar{\mu}(x,v)$ can be written as $\mathrm{d}\bar{\mu}(x,v)=\theta(\mathrm{d}x)\eta(\mathrm{d}v;x)$ for some θ and η, where θ is a probability measure and η, for θ-a.e. x, is a probability measure in v. Moreover, by Proposition 5.5 we have a representation for $\eta(\mathrm{d}v;x)$:

$$\bar{\mu}(x,v)=\delta_{\mathbf{v}(x)}(v)\theta(x).\tag{8.13}$$

We assume in this Section that

$$\begin{aligned}&\text{there exist } \alpha,\beta>0 \text{ such that} |\mathbf{v}(x)|\le\beta,\ \text{for any } x\in\mathbb{T}^N,\\ &\text{and for any } k \text{ large enough } \alpha\le\mathbf{v}(x)\cdot\tilde{\nu}_k(x) \text{ for any } x\in\Sigma_k.\end{aligned}\tag{8.14}$$

The main result of this Section is the following:

Theorem 8.8. *Assume that L satisfies assumption (**H3**). Let $\bar{\mu}(x,v)$ be a minimizing measure, and $\bar{\mu}_\#(x,p)$ its push-forward by the Legendre transform. Assume also that (8.14) holds. Then there exist measures σ^+ and σ^- on $\Sigma\times\mathbb{R}^N$ such that:*

1. *For any* $\phi \in C^1(\mathbb{T}^N \times \mathbb{R}^N)$,

$$\int_{(\mathbb{T}^N\setminus\Sigma)\times\mathbb{R}^N} \{\phi, H\} \mathrm{d}\bar{\mu}_{\#}(x,p) = \int_{\Sigma\times\mathbb{R}^N} \phi \,\mathrm{d}\sigma^+(x,p) - \int_{\Sigma\times\mathbb{R}^N} \phi \,\mathrm{d}\sigma^-(x,p). \tag{8.15}$$

2. *For any* $\varphi \in C^1(\mathbb{T}^N)$ *and any* $\psi \in C^1(\mathbb{R})$

$$\int_{\Sigma\times\mathbb{R}^N} \varphi(x)\ \mathrm{d}\sigma^+(x,p) = \int_{\Sigma\times\mathbb{R}^N} \varphi(x)\ \mathrm{d}\sigma^-(x,p), \tag{8.16}$$

$$\int_{\Sigma\times\mathbb{R}^N} \varphi(x)\psi(H(x,p))\ \mathrm{d}\sigma^+(x,p) = \int_{\Sigma\times\mathbb{R}^N} \varphi(x)\psi(H(x,p))\ \mathrm{d}\sigma^-(x,p). \tag{8.17}$$

3. *For any* C^1 *vector field* ξ*, tangent to* Σ*,*

$$\int_{\Sigma\times\mathbb{R}^N} p\cdot\xi(x)\ \mathrm{d}\sigma^+(x,p) = \int_{\Sigma\times\mathbb{R}^N} p\cdot\xi(x)\ \mathrm{d}\sigma^-(x,p). \tag{8.18}$$

Remark 8.9. Concerning the assumption (8.14) we observe the following:

1. The results in this Section are valid also if the second condition in (8.14) is replaced by $\mathbf{v}(x)\cdot\tilde{\nu}_k(x) < -\alpha$. Remind also that, as already pointed out in Remark 7.10, since $\mathbf{v}(x) = -D_pH(x, Du(x))$ and $|Du|_\infty$ is bounded, $|\mathbf{v}|$ is bounded. So the only nontrivial assumption on $\mathbf{v}$ is the lower bound.
2. We observed in Sect. 8.1 that, if $\bar{\mu}(\Sigma\times\mathbb{R}^N) > 0$, formula (8.15) may fail to hold. Notice now that condition (8.14) actually implies that the minimizing measure $\bar{\mu}$ gives no mass to the discontinuity set, that is $\bar{\mu}(\Sigma\times\mathbb{R}^N) = 0$. To check this fact, consider for any k the function w_k solving (8.11). By holonomy, (8.13) and (8.12), we get

$$0 = \int_{\mathbb{T}^N\times\mathbb{R}^N} vDw_k(x)\ \mathrm{d}\bar{\mu}(x,v) \geq \alpha\bar{\mu}(\Sigma_k\times\mathbb{R}^N) - k\int_{\mathbb{T}^N\setminus\Sigma_k} |\mathbf{v}|\,\mathrm{d}\theta(x).$$

Thus we have

$$\bar{\mu}(\Sigma_k\times\mathbb{R}^N) \leq Ck, \quad \text{for any } k$$

and by sending $k\to 0$ we get the conclusion. □

Remark 8.10. The properties listed in Theorem 8.8 rephrase in terms of minimizing measures the results obtained in Sect. 4 for minimizing trajectories.

First of all observe that, by taking $\varphi\equiv 1$ in formula (8.16),

$$\int_{\Sigma\times\mathbb{R}^N} \mathrm{d}\sigma^+(x,p) = \int_{\Sigma\times\mathbb{R}^N} \mathrm{d}\sigma^-(x,p);$$

that is the two measures σ^+ and σ^- give the same mass to $\Sigma\times\mathbb{R}^N$. We interpret formula (8.18) as conservation of the tangential momentum.

Moreover, we read formula (8.17) as conservation of energy. To see this, fix $x_0\in\Sigma$ and $\rho_0>0$ and consider, for any $\rho>0$, a $C^1(\mathbb{T}^N)$ function $\varphi(x)\geq 0$, compactly supported in Σ, with $\varphi\equiv 0$ in $\Sigma\setminus B(x_0,\rho_0+\rho)$ and $\varphi\equiv 1$ in $\Sigma\cap B(x_0,\rho_0)$. Let also $\psi(z)$ be a positive function in $C^1(\mathbb{R})$, with

$\psi \equiv 1$ if $z \le \rho_0$ and $\psi \equiv 0$ if $z \ge \rho_0 + \rho$. Then the support of the function $\varphi(x)\psi(H(x_0,p))$ is the subset of $\Sigma \times \mathbb{R}^N$

$$U_\rho := \bar{B}(x_0, \rho_0 + \rho) \times \left\{ p \in \mathbb{R}^N : |H(x_0,p)| \le \rho_0 + \rho \right\}.$$

By (8.17), for any ρ we have

$$\begin{aligned}\sigma^+(U_\rho) &\ge \int_{U_\rho} \varphi(x)\psi(H(x_0,p))\ \mathrm{d}\sigma^+(x,p)\\ &= \int_{U_\rho} \varphi(x)\psi(H(x_0,p))\ \mathrm{d}\sigma^-(x,p) \ge \sigma^-(U_0)\end{aligned}$$

and

$$\begin{aligned}\sigma^-(U_\rho) &\ge \int_{U_\rho} \varphi(x)\psi(H(x_0,p))\ \mathrm{d}\sigma^-(x,p)\\ &= \int_{U_\rho} \varphi(x)\psi(H(x_0,p))\ \mathrm{d}\sigma^+(x,p) \ge \sigma^+(U_0).\end{aligned}$$

Thus, since $U_0 = \cap_{\rho>0} U_\rho$, passing to the limit as $\rho \to 0$ in the previous inequalities we get $\sigma^+(U_0) = \sigma^-(U_0)$.

By taking ϕ compactly supported outside Σ in formula (8.15) we immediately see that a minimizing holonomic measure is invariant under the Euler Lagrange flow away from the singular set. On the other hand, to prove formula (8.15) we will use (see Proposition 8.12 below) invariance of minimizing measure in $\mathbb{T}^N \setminus \Sigma$. Then the analysis performed in the previous Sect. 8.2 cannot be avoided. □

We postpone the proof of Theorem 8.8 at the end of the Section, and establish now some preliminary technical result.

Lemma 8.11. *Under the assumption (8.14), there exists a constant $C > 0$ such that, for any k small enough,*

$$\theta(\Sigma_k) \le Ck. \tag{8.19}$$

Proof. By holonomy we have

$$0 = \int_{\mathbb{T}^N} \mathbf{v}(x) D w_k(x)\ \mathrm{d}\theta(x).$$

By (8.12) and (8.14),

$$\alpha\theta(\Sigma_k) \le \int_{\Sigma_k} \mathbf{v}(x)\cdot Dw_k(x)\ \mathrm{d}\theta(x) = -\int_{\mathbb{T}^N\setminus\Sigma_k} \mathbf{v}(x)\cdot Dw_k(x)\ \mathrm{d}\theta(x) \le C\beta k.$$

□

Proposition 8.12. *Let $\bar{\mu}(x,v)$ be a minimizing measure, and $\bar{\mu}_\#(x,p)$ its push-forward by the Legendre transform. Assume that (8.14) holds. Then there exists $C > 0$ such that*

$$\left| \int_{(\mathbb{T}^N\setminus\Sigma)\times\mathbb{R}^N} \{\phi, H\}\ \mathrm{d}\bar{\mu}_\#(x,p) \right| \le C|\phi|_\infty,$$

for any $\phi \in C^1(\mathbb{T}^N \times \mathbb{R}^N)$.

Proof. For any k small enough let $\Sigma_k := \{x : d(x,\Sigma) < k\}$. Then $\cap_k \Sigma_k = \Sigma$. Let $\theta(\mathrm{d}x)$ be as in (8.13). By Remark 8.9, assumption (8.14) implies $\bar{\mu}_\#(\Sigma \times \mathbb{R}^N) = 0$. Therefore,

$$0 = \theta(\Sigma) = \lim_{k\to\infty} \theta(\Sigma_k).$$

For any k, let $\varphi_k \in C^1(\mathbb{T}^N)$ such that $\varphi_k \equiv 1$ in $\Sigma_{k/2}$ and $\varphi_k \equiv 0$ in $\mathbb{T}^N \backslash \Sigma_k$. Then $|D\varphi_k| \leq \frac{C}{k}$ for some constant $C > 0$. Since $(1-\varphi_k)\phi$ is supported away from Σ, and $\bar{\mu}$ is invariant under the Euler–Lagrange flow in $(\mathbb{T}^N\backslash\Sigma)\times\mathbb{R}^N$, by Theorem 8.4 we have, by Remark 8.2,

$$\begin{aligned}
&\int_{(\mathbb{T}^N\backslash\Sigma)\times\mathbb{R}^N} (1-\varphi_k) D_p H D_x \phi \,\mathrm{d}\bar{\mu}_\#(x,p) \\
&\quad = \int_{(\mathbb{T}^N\backslash\Sigma)\times\mathbb{R}^N} [D_p H D_x((1-\varphi_k)\phi) - \phi D_p H D_x(1-\varphi_k)] \,\mathrm{d}\bar{\mu}_\#(x,p) \\
&\quad = \int_{(\mathbb{T}^N\backslash\Sigma)\times\mathbb{R}^N} [(1-\varphi_k) D_x H D_p \phi - \phi D_p H D_x(1-\varphi_k)] \,\mathrm{d}\bar{\mu}_\#(x,p).
\end{aligned}$$

Then, for any k,

$$\int_{(\mathbb{T}^N\backslash\Sigma)\times\mathbb{R}^N} (1-\varphi_k)\{\phi, H\} \,\mathrm{d}\bar{\mu}_\#(x,p) = \int_{\mathbb{T}^N\times\mathbb{R}^N} \phi D_p H D_x (1-\varphi_k) \,\mathrm{d}\bar{\mu}_\#(x,p). \tag{8.20}$$

By taking the absolute value, and passing to the limit as $k \to 0$ we obtain

$$\left| \int_{(\mathbb{T}^N\backslash\Sigma)\times\mathbb{R}^N} \{\phi, H\} \,\mathrm{d}\bar{\mu}_\#(x,p) \right| \leq \lim_{k\to 0} |\phi|_\infty |D_p H|_\infty \int_{\mathbb{T}^N} |D\varphi_k| \,\mathrm{d}\theta(x)$$

which gives the desired estimate after observing that for any k, $|D\varphi_k| \leq C/k$ and that, by Lemma 8.11, $\theta(\Sigma_k) \leq Ck$, for k small enough. □

Proof of Theorem 8.8. By Proposition 8.12 the mapping from $C^1(\mathbb{T}^N \times \mathbb{R}^N)$ to $\mathbb{R}$

$$\phi \mapsto \int_{(\mathbb{T}^N\backslash\Sigma)\times\mathbb{R}^N} \{\phi, H\} \,\mathrm{d}\bar{\mu}_\#(x,p)$$

is a zero-order distribution. Furthermore, by Theorem 8.4 and Remark 8.2 this distribution is supported on $\Sigma \times \mathbb{R}^N$. Formula 8.15 then follows by the Riesz representation theorem.

Formula (8.16) is obtained by putting $\phi = \varphi(x)$ in (8.15) and observing that, by holonomy,

$$\int_{(\mathbb{T}^N\backslash\Sigma)\times\mathbb{R}^N} \{\phi, H\} \,\mathrm{d}\bar{\mu}_\#(x,p) = -\int_{(\mathbb{T}^N\backslash\Sigma)\times\mathbb{R}^N} D_p H D_x \varphi \,\mathrm{d}\bar{\mu}_\#(x,p) = 0.$$

To show (8.17), it is enough to take $\phi(x,p) = \varphi(x)\psi(H(x,p))$ and observe that

$$\{\phi, H\} = \psi(H(x,p))\{\varphi, H\}$$

Moreover, by Remark 7.11, $\psi(H(x,p)) = \psi(\bar{H})$, $\bar{\mu}_{\#}$-a.e. Then, by holonomy and (8.15) we have

$$0 = \psi(\bar{H}) \int_{(\mathbb{T}^N \setminus \Sigma) \times \mathbb{R}^N} \{\varphi, H\} \ \mathrm{d}\bar{\mu}_{\#}(x,p) \tag{8.21}$$

$$= \int_{\Sigma \times \mathbb{R}^N} \varphi(x)\psi(H(x,p)) \ \mathrm{d}\sigma^{+}(x,p) - \int_{\Sigma \times \mathbb{R}^N} \varphi(x)\psi(H(x,p)) \ \mathrm{d}\sigma^{-}(x,p). \tag{8.22}$$

We now prove (8.18). Remind that, by Proposition 5.5 $\bar{\mu}$ is supported on the graph of a function $\mathbf{v}$ and that, by Proposition 7.9 and Remark (7.10),

$$p = Du(x),\ \bar{\mu}_{\#}\text{-a.e.}, \quad \mathbf{v}(x) = -D_p H(x, D_x u(x)), \text{ Lebesgue a.e. in } \mathbb{T}^N;$$

here u is a solution to (7.6). Moreover

$$Du(x) = D_v L(x,v), \quad \bar{\mu}\text{-a.e. in } (\mathbb{T}^N \setminus \Sigma) \times \mathbb{R}^N.$$

Taking into account the previous results, using (8.15) and holonomy we compute:

$$\begin{aligned}
&\int_{\Sigma \times \mathbb{R}^N} p \cdot \xi(x) \ \mathrm{d}\sigma^{-}(x,p) - \int_{\Sigma \times \mathbb{R}^N} p \cdot \xi(x) \ \mathrm{d}\sigma^{+}(x,p) \\
&\quad = - \int_{(\mathbb{T}^N \setminus \Sigma) \times \mathbb{R}^N} \{p \cdot \xi(x), H\} \ \mathrm{d}\bar{\mu}_{\#}(x,p) \\
&\quad = - \int_{(\mathbb{T}^N \setminus \Sigma) \times \mathbb{R}^N} \{p \cdot \xi(x), H\} \ \mathrm{d}\bar{\mu}_{\#}(x,p) \\
&\qquad -2 \int_{(\mathbb{T}^N \setminus \Sigma) \times \mathbb{R}^N} D_x(p \cdot \xi(x)) D_p H(x,p) \ \mathrm{d}\bar{\mu}_{\#}(x,p) \\
&\quad = \int_{(\mathbb{T}^N \setminus \Sigma) \times \mathbb{R}^N} [-\xi(x) \cdot D_x H(x,p) - p \cdot D\xi(x) D_p H(x,p)] \ \mathrm{d}\bar{\mu}_{\#}(x,p) \\
&\quad = \int_{(\mathbb{T}^N \setminus \Sigma) \times \mathbb{R}^N} [\xi(x) \cdot D_x L(x,v) + v \cdot D_v L(x,v) D\xi(x)] \ \mathrm{d}\bar{\mu}(x,v) = 0,
\end{aligned}$$

where in the last identity we used Proposition 5.12. □

Remark 8.13. Arguing as in the proof of formula (8.20) (see the proof of Proposition 8.12 above) it is possible to write a representation formula for $\sigma^{\pm}$. Let $\varphi^{\delta}(x)$ be a $C^1(\mathbb{T}^N)$ function compactly supported on a strip of size δ around Σ. Then, for any $\phi(x,p) \in C^1(\mathbb{T}^N \times \mathbb{R}^N)$,

$$\int_{(\mathbb{T}^N\setminus\Sigma)\times\mathbb{R}^N} \varphi^\delta \{\phi, H\} \, \mathrm{d}\bar{\mu}_\#(x,p) = \int_{\times\mathbb{R}^N} \phi D_p H D_x \varphi^\delta \, \mathrm{d}\bar{\mu}_\#(x,p),$$

Then, by taking the limit as $\delta \to 0$,

$$\int_{\Sigma\times\mathbb{R}^N} \phi(x,p) \, \mathrm{d}\sigma^+(x,p) - \int_{\Sigma\times\mathbb{R}^N} \phi(x,p) \, \mathrm{d}\sigma^-(x,p)$$
$$= \lim_{\delta\to 0} \int_{\mathbb{T}^N\times\mathbb{R}^N} \phi(x,p) D_p H(x,p) D_x \varphi^\delta \, \mathrm{d}\bar{\mu}(x,p).$$

□

Appendix A

A.1. Proof of Proposition 4.3

Since Σ is smooth, a sufficiently small neighborhood of $\mathbf{x}(t_0)$ in Σ is diffeomorphic to the set $\{x \in \mathbb{R}^N : x_1 = \mathbf{x}_1(t_0)\} \cap \mathcal{N}$, for a certain open set $\mathcal{N}$.

Let $\delta > 0$ be as in Definition 4.2, and $\varphi : \mathbb{R} \to \mathbb{R}$ be C^∞ and compactly supported on $(\mathbf{x}_1(t_0) - \delta, \mathbf{x}_1(t_0) + \delta)$, such that $\varphi \equiv 1$ on $(\mathbf{x}_1(t_0) - \delta/2, \mathbf{x}_1(t_0) + \delta/2)$. Consider, for any fixed $i = 2, \ldots, N$, and any positive ϵ the following variation of $\mathbf{x}(t)$:

$$\mathbf{x}^\epsilon(t) := \mathbf{x}(t) + \epsilon\varphi(\mathbf{x}_1(t)) e_i.$$

Since $\mathbf{x}(t)$ minimizes the action, taking into account (4.2), a direct computation gives

$$0 = \frac{\mathrm{d}}{\mathrm{d}\epsilon}\Big|_{\epsilon=0} \int_0^T L(\mathbf{x}^\epsilon(t), \dot{\mathbf{x}}^\epsilon(t)) \, \mathrm{d}t = -[\![D_v L(\mathbf{x}(t), \dot{\mathbf{x}}(t)) \cdot e_i]\!]_{t_0}$$

Hence, for any $i = 2, \ldots, N$. This implies (4.3), because when flattening Σ near $x_0, T_{x_0}\Sigma$ is mapped into $\mathbb{R}^{N-1}$ and any $\xi \in T_{x_0}\Sigma$ can be written as $\xi = \sum_{i=2}^N \xi_i e_i$ for appropriate $\xi_i \in \mathbb{R}$.

A.2. Proof of Proposition 4.4

Let us consider for any positive ϵ the following variation of $\mathbf{x}(t)$:

$$\mathbf{x}^\epsilon(t) := \mathbf{x}(t) + \epsilon\varphi(t)\nu(\mathbf{x}(t)),$$

where φ is a C^∞ function, positive, and compactly supported on (t_1, t_2). As $\mathbf{x}^\epsilon$, for any $\epsilon > 0$, does not decrease the action (4.1), we have:

$$0 \le \frac{\mathrm{d}}{\mathrm{d}\epsilon}\Big|_{\epsilon=0} \int_{t_1}^{t_2} L(\mathbf{x}^\epsilon, \dot{\mathbf{x}}^\epsilon) \, \mathrm{d}t$$
$$= \int_{t_1}^{t_2} \left[\varphi(t)\nu(\mathbf{x}(t)) D_x L(\mathbf{x}^\epsilon, \dot{\mathbf{x}}^\epsilon) + \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\mathrm{d}}{\mathrm{d}\epsilon}\mathbf{x}^\epsilon(t)\right) D_v L(\mathbf{x}^\epsilon, \dot{\mathbf{x}}^\epsilon)\right] \, \mathrm{d}t\Big|_{\epsilon=0}$$
$$= \int_{t_1}^{t_2} \varphi(t) \left[D_x L(\mathbf{x}^\epsilon, \dot{\mathbf{x}}^\epsilon) - \frac{\mathrm{d}}{\mathrm{d}t} D_v L(\mathbf{x}^\epsilon, \dot{\mathbf{x}}^\epsilon)\right] \cdot \nu(\mathbf{x}(t)) \, \mathrm{d}t\Big|_{\epsilon=0}.$$

Because φ is nonnegative and arbitrary, we get (4.4). We should note that because, by (**H2**)-3., the action may be discontinuous for $\epsilon < 0$ we cannot use these variations together with differentiation to extract additional information.

We now prove (4.5). Fix $t_0 \in (t_1, t_2)$. Since Σ is smooth, a sufficiently small neighborhood of $\mathbf{x}(t_0)$ in Σ is diffeomorphic to the set $\Sigma' := \{x \in \mathbb{R}^N : x_1 = \mathbf{x}_1(t_0)\} \cap \mathcal{N}$, for a certain open set $\mathcal{N}$. Thus $T_{\mathbf{x}(t_0)}\Sigma$ is mapped into $\mathbb{R}^{N-1}$ and any $\xi \in T_{\mathbf{x}(t_0)}\Sigma$ can be written as $\xi = \sum_{i=2}^N \xi_i e_i$ for appropriate $\xi_i \in \mathbb{R}$. Let $\delta > 0$ and $\varphi(t)$ a C^∞ function, compactly supported $(\mathbf{x}_1(t_0) - \delta, \mathbf{x}_1(t_0) + \delta)$, $0 \le \varphi \le 1$, and $\epsilon > 0$. Consider for any $\epsilon > 0$ and any $i = 2, \ldots N$ the following variation of $\mathbf{x}$:

$$\mathbf{x}^\epsilon(t) := \mathbf{x}(t) + \epsilon\varphi(t)e_i.$$

If ϵ is sufficiently small, $\mathbf{x}^\epsilon$ belongs to the flat set Σ', for any $t \in (\mathbf{x}_1(t_0) - \delta/2, \mathbf{x}_1(t_0) + \delta/2)$. Consequently, at $\epsilon = 0$ we get:

$$\left[D_x L(\mathbf{x}(t_0), \dot{\mathbf{x}}(t_0)) - \frac{\mathrm{d}}{\mathrm{d}t} D_v L(\mathbf{x}(t_0), \dot{\mathbf{x}}(t_0))\right] \cdot e_i = 0.$$

A.3. Proof of Proposition 4.6

We sketch the proof, as it goes as that of Proposition 4.3. After flattening Σ around x_0, we consider for any $i = 2, \ldots, N$ any $\epsilon > 0$ the variation

$$\mathbf{x}^\epsilon(t) := \mathbf{x}(t) + \epsilon\varphi(t)e_i;$$

φ is a C^∞ function, compactly supported on $(t_0 - \delta, t_0 + \delta)$, $\varphi \equiv 1$ on $(t_0 - \delta/2, t_0 + \delta/2)$, $\delta > 0$ being as in Definition 4.5. Then, integrating by parts and using (4.2) we have

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon} \int_{t_0-\delta}^{t_0} L(\mathbf{x}^\epsilon(t), \dot{\mathbf{x}}^\epsilon(t))\ \mathrm{d}t = \lim_{t \to t_0^-} \partial_{v_i} L(\mathbf{x}^\epsilon(t), \dot{\mathbf{x}}^\epsilon(t)).$$

Moreover, taking into account (4.5),

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon} \int_{t_0}^{t_0+\delta} L(\mathbf{x}^\epsilon(t), \dot{\mathbf{x}}^\epsilon(t))\ \mathrm{d}t = -\lim_{t \to t_0^+} \partial_{v_i} L(\mathbf{x}^\epsilon(t), \dot{\mathbf{x}}^\epsilon(t)).$$

Thus, at $\epsilon = 0$, since $\mathbf{x}$ is a minimizer, we get the conclusion.

A.4. Proof of Proposition 4.7

If $\mathbf{x}$ does not intersect Σ there is nothing to prove, as this is a well known fact. Then we assume that there exist $t_1, t_2, 0 < t_1 \le t_2 < T$ such that $\mathbf{x}(t) \in \Sigma$ for any $t \in [t_1, t_2]$, and $\mathbf{x}(t) \notin \Sigma$ if $t \notin [t_1, t_2]$. Because $\mathbf{x}$ is piecewise C^2 all the other cases can be handled similarly.

Consider for any $\epsilon > 0$ the variation of $\mathbf{x}$:

$$\mathbf{x}^\epsilon(t) := \mathbf{x}(t + \epsilon\varphi(t)); \tag{.23}$$

here φ is a C^∞ positive function compactly supported on a certain open interval I, such that $\bar{I} \subset (0, T)$. Thus $\mathbf{x}^\epsilon$ belongs to $\mathbf{X}_{a,b}^T$. From time to time, in order to extract different informations, we will precise the subinterval of $(0, T)$ where $\bar{I}$ is contained. Observe preliminarily that:

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon}\Big|_{\epsilon=0} \int_0^T L(\mathbf{x}^\epsilon, \dot{\mathbf{x}}^\epsilon)\ \mathrm{d}t$$
$$= \int_I \varphi \left[\dot{\mathbf{x}}\, D_x L(\mathbf{x}, \dot{\mathbf{x}}) + \ddot{\mathbf{x}}\, D_v L(\mathbf{x}, \dot{\mathbf{x}}) - \frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{x}\, D_v L(\mathbf{x}, \dot{\mathbf{x}}))\right]\ \mathrm{d}t$$
$$+\varphi\, \dot{\mathbf{x}}\, D_v L(\mathbf{x}, \dot{\mathbf{x}})\Big|_{\partial I}. \tag{.24}$$

Since $\mathbf{x}$ is a minimizer the left hand side of the previous inequality is 0. Then, by the arbitrariness of φ,

$$\left[\dot{\mathbf{x}}\, D_x L(\mathbf{x}, \dot{\mathbf{x}}) + \ddot{\mathbf{x}}\, D_v L(\mathbf{x}, \dot{\mathbf{x}}) - \frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{x}\, D_v L(\mathbf{x}, \dot{\mathbf{x}}))\right] = 0, \quad \text{for a.e. } t \in (0, T). \tag{.25}$$

Case 1. Suppose first that $t_1 < t_2$; $\mathbf{x}(t)$ stays in Σ for $t_1 < t < t_2$. Consider the variation (.23) and take I such that $\bar{I} \subset (0, t_2)$, and φ such that $\varphi \equiv 1$ in a small neighborhood of t_1. Since $\mathbf{x}$ is a minimizer, by (.24) and (.25) we get

$$0 = \frac{\mathrm{d}}{\mathrm{d}\epsilon}\Big|_{\epsilon=0} \int_0^T L(\mathbf{x}^\epsilon, \dot{\mathbf{x}}^\epsilon)\ \mathrm{d}t = [\![E_{\mathbf{x}}(t)]\!]_{t_1}.$$

Arguing analogously, for $\bar{I} \subset (t_1, T)$, and φ such that $\varphi \equiv 1$ in a small neighborhood of t_2, we obtain $[\![E_{\mathbf{x}}(t)]\!]_{t_2} = 0$. Then there is no dissipation of energy when entering or exiting the discontinuity locus.

Integrating by parts and taking into account (4.2), we compute:

$$\frac{\mathrm{d}}{\mathrm{d}h} \int_h^{t_1+h} L(\mathbf{x}(t-h), \dot{\mathbf{x}}(t-h))\ \mathrm{d}t = E_{\mathbf{x}}(0) - \lim_{t\to t_1^-} E_{\mathbf{x}}(t)$$

and

$$\frac{\mathrm{d}}{\mathrm{d}h} \int_{t_1+h}^{t_2+h} L(\mathbf{x}(t-h), \dot{\mathbf{x}}(t-h))\ \mathrm{d}t = \lim_{t\to t_1^+} E_{\mathbf{x}}(t) - \lim_{t\to t_2^-} E_{\mathbf{x}}(t),$$
$$\frac{\mathrm{d}}{\mathrm{d}h} \int_{t_2+h}^{T+h} L(\mathbf{x}(t-h), \dot{\mathbf{x}}(t-h))\ \mathrm{d}t = \lim_{t\to t_2^+} E_{\mathbf{x}}(t) - E_{\mathbf{x}}(T).$$

Observing that $\int_h^{T+h} L(\mathbf{x}(t-h), \dot{\mathbf{x}}(t-h))\ \mathrm{d}t$ is independent by h, and taking into account that $[\![E_{\mathbf{x}}(t)]\!]_{t_1} = [\![E_{\mathbf{x}}(t)]\!]_{t_2} = 0$ we conclude the proof in Case 1.
Case 2. We suppose now that $t_1 = t_2 =: t_0$; thus $\mathbf{x}(t)$ crosses Σ at $t = t_0$. Arguing as in the first part of the proof it is easy to verify that

$$\frac{\mathrm{d}}{\mathrm{d}h}\left[\int_h^{t_0+h} L(\mathbf{x}(t-h), \dot{\mathbf{x}}(t-h))\ \mathrm{d}t\right] = E_{\mathbf{x}}(0) - \lim_{t\to {t_0}^-} E_{\mathbf{x}}(t),$$
$$\frac{\mathrm{d}}{\mathrm{d}h}\left[\int_{t_0+h}^{T+h} L(\mathbf{x}(t-h), \dot{\mathbf{x}}(t-h))\ \mathrm{d}t\right] = \lim_{t\to {t_0}^+} E_{\mathbf{x}}(t) - E_{\mathbf{x}}(T).$$

Then,

$$0 = \frac{\mathrm{d}}{\mathrm{d}h} \int_h^{T+h} L(\mathbf{x}(t-h), \dot{\mathbf{x}}(t-h))\ \mathrm{d}t = E_{\mathbf{x}}(0) - E_{\mathbf{x}}(T).$$

A.5. Proof of Proposition 5.5

By Remark 5.4, if $\bar{\mu}$ is supported on a graph, there exists $\mathbf{v}(x)$ such that $\eta(\mathrm{d}v; x) = \delta_{\mathbf{v}(x)}(\mathrm{d}v)$.

Assume by contradiction that $\bar{\mu}$ is not supported on a graph and set for any x

$$\mathbf{v}(x) := \int_{\mathbb{R}^N} v\,\eta(\mathrm{d}v; x).$$

We further define the measure $\tilde{\eta}(\mathrm{d}v; x) := \delta_{\mathbf{v}(x)}(v)$. We finally set $\mathrm{d}\tilde{\mu} := \tilde{\eta}(\mathrm{d}v; x)\theta(\mathrm{d}x)$. Observe that $\tilde{\mu}$ is readily a nonnegative probability measure, as both $\tilde{\eta}$ and θ are so. Moreover, for any $\varphi \in C^1(\mathbb{T}^N)$, by definition of $\tilde{\eta}$ and $\mathbf{v}(x)$,

$$\int_{\mathbb{R}^N} vD\varphi(x)\tilde{\eta}(\mathrm{d}v; x) = \mathbf{v}(x)D\varphi(x) = \int_{\mathbb{R}^N} vD\varphi(x)\eta(\mathrm{d}v; x),$$

for any $x \in \mathbb{T}^N$. Thus,

$$\begin{aligned}\int_{\mathbb{T}^N\times\mathbb{R}^N} vD\varphi(x)\mathrm{d}\tilde{\mu} &= \int_{\mathbb{T}^N\times\mathbb{R}^N} vD\varphi(x)\tilde{\eta}(\mathrm{d}v; x)\theta(\mathrm{d}x)\\ &= \int_{\mathbb{T}^N\times\mathbb{R}^N} vD\varphi(x)\eta(\mathrm{d}v; x)\theta(\mathrm{d}x) = \int_{\mathbb{T}^N\times\mathbb{R}^N} vD\varphi(x)\ \mathrm{d}\bar{\mu} = 0,\end{aligned}$$

because $\bar{\mu}$ is holonomic. Then $\tilde{\mu}$ belongs to $\mathcal{M}_{\mathrm{hol}} \cap \mathcal{M}_1^+$.

Now, since $L(x, v)$ is strictly convex in v, we have by Jensen's inequality,

$$\int_{\mathbb{R}^N} L(x, v)\tilde{\eta}(\mathrm{d}v; x) < \int_{\mathbb{R}^N} L(x, v)\eta(\mathrm{d}v; x),$$

for any $x \in \mathbb{T}^N$ such that $\eta(\mathrm{d}v; x) \neq \tilde{\eta}(\mathrm{d}v; x)$. Hence, integrating against θ, we get

$$\int L(x, v)\ \mathrm{d}\tilde{\mu} < \int L(x, v)\ \mathrm{d}\bar{\mu},$$

which is a contradiction with $\bar{\mu}$ being a minimizer.

A.6. Proof of Lemma 5.10

Since h is convex, for any $\mu \in \mathcal{M}$ we have:

$$\hat{h}(\mu) = \sup_{\psi\in\mathcal{C}_0^\gamma}\left(\int \psi\ \mathrm{d}\mu - h(\psi)\right) = \sup_{\psi\in\mathcal{C}}\int \psi\ \mathrm{d}\mu.$$

Thus, if $\mu \in \mathcal{M}_{\mathrm{hol}}$, then $\hat{h}(\mu) = \int \psi\ \mathrm{d}\mu = 0$. If instead $\mu \notin \mathcal{M}_{\mathrm{hol}}$, there exist $\bar{\psi} \in \mathcal{C}$ such that $\int \bar{\psi}\ \mathrm{d}\mu =: \delta > 0$. Hence,

$$\hat{h}(\mu) = \sup_{\psi\in\mathcal{C}}\int \psi\ \mathrm{d}\mu \geq \sup_{\lambda>0}\lambda\int \bar{\psi}\ \mathrm{d}\mu = +\infty.$$

The expression of $\hat{h}$ is then proved.

Since g is concave, for any $\mu \in \mathcal{M}$ we have:

$$\hat{g}(\mu) = \inf_{\psi\in\mathcal{C}_0^\gamma}\left(\int \psi\ \mathrm{d}\mu - g(\psi)\right).$$

Assume that $\mu \not\geq 0$, that is, there exists $\bar{\psi} \geq 0$ such that $\int \bar{\psi}\ \mathrm{d}\mu =: -\delta < 0$. Set, for any n, $\psi_n := n\bar{\psi}$. Since $g(\psi_n) \geq 0$ for any n, we have

$$\hat{g}(\mu) \leq \inf_n \left(\int \psi_n\ \mathrm{d}\mu - g(\psi_n) \right) = -\infty.$$

Assume now that $\mu \geq 0$. Let $L_k(x, v)$ be sequence in $\mathcal{C}_0^\gamma$ converging pointwise to $L(x, v)$ from below. Note that any function $\psi \in \mathcal{C}_0^\gamma$ can be written as $\psi = \phi - L_k$, for some $\phi \in \mathcal{C}_0^\gamma$. Then, for any k,

$$\begin{aligned}
\hat{g}(\mu) &= \inf_{\phi \in \mathcal{C}_0^\gamma} \left(\int (\phi - L_k)\ \mathrm{d}\mu - g(\phi - L_k) \right) \\
&= -\int L_k\ \mathrm{d}\mu + \inf_{\phi \in \mathcal{C}_0^\gamma} \left(\int \phi\ \mathrm{d}\mu - g(\phi - L_k) \right) \\
&= -\int L_k\ \mathrm{d}\mu + \inf_{\phi \in \mathcal{C}_0^\gamma} \left(\int \phi\ \mathrm{d}\mu - \min_{x,v}(L - L_k + \phi) \right) \\
&\leq -\int L_k\ \mathrm{d}\mu + \inf_{\phi \in \mathcal{C}_0^\gamma} \left(\int \phi\ \mathrm{d}\mu - \min_{x,v}(L - L_k) - \min_{x,v} \phi \right) \\
&\leq -\int L_k\ \mathrm{d}\mu + \inf_{\phi \in \mathcal{C}_0^\gamma} \left(\int \phi\ \mathrm{d}\mu - \min_{x,v} \phi \right).
\end{aligned}$$

By applying the monotone convergence theorem to the right hand side of the last inequality, we obtain

$$\hat{g}(\mu) \leq -\int L\ \mathrm{d}\mu + \inf_{\phi \in \mathcal{C}_0^\gamma} \left(\int \phi\ \mathrm{d}\mu - \min_{x,v} \phi \right) \leq -\int L\ \mathrm{d}\mu, \tag{.26}$$

by choosing $\phi \equiv 0$.

If $\mu > 0$, but $\int \mathrm{d}\mu < 1$ the previous computation shows that

$$\begin{aligned}
\hat{g}(\mu) &\leq -\int L\ \mathrm{d}\mu + \inf_{\phi \in \mathcal{C}_0^\gamma} \left(\int \phi\ \mathrm{d}\mu - \min_{x,v} \phi \right) \\
&\leq -\int L\ \mathrm{d}\mu + \inf_{\lambda > 0} \lambda \left(\int\ \mathrm{d}\mu - 1 \right) = -\infty,
\end{aligned}$$

by choosing $\phi \equiv \lambda$. Similarly, $\hat{g}(\mu) = -\infty$ if $\int \mathrm{d}\mu > 1$, by taking the infimum over $\lambda < 0$.

Finally, if $\int L\ \mathrm{d}\mu = +\infty$, by (.26) $\hat{g}(\mu) = -\infty$. If instead $\int \mathrm{d}\mu = 1$ and $\int L\ \mathrm{d}\mu < +\infty$, we have

$$\begin{aligned}
\hat{g}(\mu) &= \inf_{\psi \in \mathcal{C}_0^\gamma} \left(\int \psi\ \mathrm{d}\mu - g(\psi) \right) \\
&= -\int L\ \mathrm{d}\mu + \inf_{\psi \in \mathcal{C}_0^\gamma} \left(\int (L + \psi)\ \mathrm{d}\mu - \min_{x,v}(L + \psi) \right) \geq -\int L\ \mathrm{d}\mu;
\end{aligned} \tag{.27}$$

in fact, if $\int \mathrm{d}\mu = 1, \min_{x,v}(L + \psi) \leq \int (L + \psi)\mathrm{d}\mu$. The inequalities (.26) and (.27) conclude the proof.

A.7. Proof of Lemma 8.6

This is an application of the holonomy preserving variations; see Sect. 5.3 and Proposition 5.12. Observe first that

$$\begin{aligned}
&\int \Big\{ v_k \frac{\partial \phi}{\partial x_k}(x, v^\epsilon(x)) \\
&\quad + \frac{\partial \phi}{\partial v_j}(x, v^\epsilon(x)) \left(\frac{\partial^2 L}{\partial v^2}\right)^{-1}_{js} (x, v^\epsilon(x)) \left[\frac{\partial L}{\partial x_s}(x, v)\right. \\
&\quad \left.\left. - v_k \frac{\partial^2 L}{\partial x_k \partial v_s}(x, v^\epsilon(x))\right]\right\} \, \mathrm{d}\bar{\mu} \\
&= \int v_k \frac{\partial}{\partial x_k}(\phi(x, v^\epsilon(x))) \, \mathrm{d}\bar{\mu} - \int v_k \frac{\partial}{\partial x_k}\left(\frac{\partial L}{\partial v_s}(x, v^\epsilon)\, \xi^\epsilon_s\right) \, \mathrm{d}\bar{\mu} \\
&\quad + \int v_k \left(\frac{\partial L}{\partial v_s}(x, v^\epsilon(x)) - \frac{\partial L}{\partial v_s}(x, v)\right) \frac{\partial}{\partial x_k} \xi^\epsilon_s \, \mathrm{d}\bar{\mu}. \qquad (.28)
\end{aligned}$$

To confirm this notice that

$$\begin{aligned}
&\frac{\partial \phi}{\partial x_k}(x, v^\epsilon(x)) = \frac{\partial}{\partial x_k}\left(\phi(x, v^\epsilon(x))\right) - \frac{\partial \phi}{\partial v_j}(x, v^\epsilon(x)) \frac{\partial v^\epsilon_j}{\partial x_k}(x) \\
&= \frac{\partial}{\partial x_k}\left(\phi(x, v^\epsilon(x))\right) - \frac{\partial \phi}{\partial v_j}(x, v^\epsilon(x)) \left(\frac{\partial^2 L}{\partial v^2}\right)^{-1}_{js} \frac{\partial^2 L}{\partial v_s \partial v_q}(x, v^\epsilon(x)) \frac{\partial v^\epsilon_q}{\partial x_k}(x) \\
&= \frac{\partial}{\partial x_k}\left(\phi(x, v^\epsilon(x))\right) - \xi^\epsilon_s(x) \frac{\partial^2 L}{\partial v_s \partial v_q}(x, v^\epsilon(x)) \frac{\partial v^\epsilon_q}{\partial x_k}(x).
\end{aligned}$$

Then use the previous identity to realize that, after some cancellation, (.28) is equivalent to the following

$$\begin{aligned}
&\int \xi^\epsilon_s \left(\frac{\partial L}{\partial x_s}(x, v) - v_k \left(\frac{\partial^2 L}{\partial x_k \partial v_s}(x, v^\epsilon(x)) + \frac{\partial^2 L}{\partial v_s \partial v_q}(x, v^\epsilon(x)) \frac{\partial v^\epsilon_q}{\partial x_k}(x)\right)\right) \, \mathrm{d}\bar{\mu} \\
&= - \int v_k \frac{\partial}{\partial x_k}\left(\frac{\partial L}{\partial v_s}(x, v^\epsilon(x)) \xi^\epsilon_s\right) \, \mathrm{d}\bar{\mu} \\
&\quad + \int v_k \left(\frac{\partial L}{\partial v_s}(x, v^\epsilon(x)) - \frac{\partial L}{\partial v_s}(x, v)\right) \frac{\partial}{\partial x_k}(\xi^\epsilon_s) \, \mathrm{d}\bar{\mu}
\end{aligned}$$

that we can further rewrite by using the chain rule:

$$\begin{aligned}
&\int \xi^\epsilon_s \left(\frac{\partial L}{\partial x_s}(x, v) - v_k \frac{\partial}{\partial x_k}\left(\frac{\partial L}{\partial v_s}(x, v^\epsilon(x))\right)\right) \, \mathrm{d}\bar{\mu} \\
&\quad = - \int v_k \xi^\epsilon_s \frac{\partial}{\partial x_k}\left(\frac{\partial L}{\partial v_s}(x, v^\epsilon(x))\right) \, \mathrm{d}\bar{\mu} - \int v_k \frac{\partial L}{\partial v_s}(x, v) \frac{\partial}{\partial x_k}(\xi^\epsilon_s) \, \mathrm{d}\bar{\mu}.
\end{aligned}$$

After the cancellation of the term $\int v_k \xi^\epsilon_s \frac{\partial}{\partial x_k}\left(\frac{\partial L}{\partial v_s}(x, v^\epsilon(x))\right) \mathrm{d}\bar{\mu}$ in both sides of the previous identity, we can finally write it as

$$\int \left[\xi^\epsilon_s(x) \frac{\partial L}{\partial x_s}(x, v) + v_k \frac{\partial L}{\partial v_s}(x, v) \frac{\partial \xi^\epsilon_s}{\partial x_k}(x)\right] \, \mathrm{d}\bar{\mu} = 0. \qquad (.29)$$

Now, since ϕ is compactly supported in $(\mathbb{T}^N \backslash \Sigma) \times \mathbb{R}^N$, $\frac{\partial \phi}{\partial v_j}(x,v) \equiv 0$ in a certain neighborhood of $\Sigma \times \mathbb{R}^N$. This in turn implies that the vector field ξ^ϵ defined in (8.6) is tangent to Σ. We are then in position to invoke Proposition 5.12 to affirm that (.29) holds, and consequently (.28) is satisfied.

To conclude the proof observe that, by holonomy, the first and the second integral in the right and side of in (.28) are actually zero. Formula (8.5) is then established.

A.8. Proof of Lemma 8.7

To ease the notation put

$$F_s(x,v) := \frac{\partial \phi}{\partial v_j}(x,v) \left(\frac{\partial^2 L}{\partial v} \right)^{-1}_{js} (x,v) \quad (s = 1, \ldots, N);$$

thus, taking into account Remark 7.10, $\xi^\epsilon_s = F_s(x, \eta_\epsilon * D_p H(x, Du(x)))$ where u is a Lipschitz viscosity solution to the Eq. (7.6). Set

$$C_0 := \max \left\{ |D^2_{xp} H|_{L^\infty(\mathbb{T}^N \times B(0,|Du|_\infty))}; |D^2_{pp} H|_{L^\infty(\mathbb{T}^N \times B(0,|Du|_\infty))} \right\}$$

We compute

$$\begin{aligned} \frac{\partial}{\partial x_k} F_s(x, \eta_\epsilon * D_p H(x, Du(x))) &= \frac{\partial F_s}{\partial x_k}(x, \eta_\epsilon * D_p H(x, Du(x))) \\ &+ \frac{\partial F_s}{\partial v_k}(x, \eta_\epsilon * D_p H(x, Du(x))) \left(\frac{\partial \eta_\epsilon}{\partial x_k} * D_p H(x, Du(x)) \right). \end{aligned} \tag{.30}$$

Observe that $F(x,v)$ is Lipschitz continuous, then $F| \leq C$ for some positive constant C, thus the first term in the previous identity is bounded uniformly in ϵ:

$$\left| \frac{\partial}{\partial x_k} F_s(x, \eta_\epsilon * D_p H(x, Du(x))) \right|^2 \leq C + C \left| \frac{\partial \eta_\epsilon}{\partial x_k} * D_p H(x, Du(x)) \right|^2.$$

We now concentrate on the last term in the right hand side of (.30). Note that $|D_x \eta_\epsilon| \leq \frac{\tilde{\eta}_\epsilon}{\epsilon}$ for some function $\tilde{\eta}_\epsilon$ satisfying

$$\left| \int \tilde{\eta}_\epsilon(x-y) \, \mathrm{d}y \right| \leq C, \text{uniformly in } \epsilon. \tag{.31}$$

Thus,

$$\begin{aligned} &\int_{\mathbb{T}^N \backslash \Sigma_\delta} \left| \frac{\partial \eta_\epsilon}{\partial x_k} * D_p H(x, Du(x)) \right|^2 \, \mathrm{d}\bar{\mu}_{\mathbb{T}^N} \\ &= \int_{\mathbb{T}^N \backslash \Sigma_\delta} \left| \int \frac{\partial \eta_\epsilon}{\partial x_k}(x-y) D_p H(y, Du(y)) \, \mathrm{d}y \right|^2 \, \mathrm{d}\bar{\mu}_{\mathbb{T}^N} \\ &= \int_{\mathbb{T}^N \backslash \Sigma_\delta} \left| \int \frac{\partial \eta_\epsilon}{\partial x_k}(x-y) [D_p H(y, Du(y)) - D_p H(x, Du(x))] \, \mathrm{d}y \right|^2 \, \mathrm{d}\bar{\mu}_{\mathbb{T}^N} \end{aligned}$$

$$\leq C\left(\int_{\mathbb{T}^N\setminus\Sigma_\delta}\left|\int\left|\frac{\partial\eta_\epsilon}{\partial x_k}(x-y)\right|\left|D_pH(y,Du(y))-D_pH(x,Du(y))\right|\,\mathrm{d}y\right|^2\mathrm{d}\bar{\mu}_{\mathbb{T}^N}\right.$$
$$\left.+\int_{\mathbb{T}^N\setminus\Sigma_\delta}\left|\int\left|\frac{\partial\eta_\epsilon}{\partial x_k}(x-y)\right|\left|D_pH(x,Du(y))-D_pH(x,Du(x))\right|\,\mathrm{d}y\right|^2\,\mathrm{d}\bar{\mu}_{\mathbb{T}^N}\right). \tag{.32}$$

We now prove that both terms in the right hand side of the previous inequality are bounded uniformly in ϵ. In fact, observing that $|x-y|\leq\epsilon$ and taking into account (.31) we get:

$$\int_{\mathbb{T}^N\setminus\Sigma_\delta}\left|\int\left|\frac{\partial\eta_\epsilon}{\partial x_k}(x-y)\right|\left|D_pH(y,Du(y))-D_pH(x,Du(y))\right|\,\mathrm{d}y\right|^2\,\mathrm{d}\bar{\mu}_{\mathbb{T}^N}(x)$$
$$\leq C_0\int_{\mathbb{T}^N\setminus\Sigma_\delta}\left|\int\frac{\tilde{\eta}_\epsilon(y-x)}{\epsilon}|y-x|\,\mathrm{d}y\right|^2\,\mathrm{d}\bar{\mu}_{\mathbb{T}^N}(x)\leq C,$$

uniformly in ϵ. We now deal with the other term in (.32). Notice first that

$$\int\left|\frac{\partial\eta_\epsilon}{\partial x_k}(x-y)\right|\left|D_pH(x,Du(y))-D_pH(x,Du(x))\right|\,\mathrm{d}y$$
$$\leq C_0\int\left|\frac{\partial\eta_\epsilon}{\partial x_k}(x-y)\right|\left|Du(y)-Du(x)\right|\,\mathrm{d}y$$

and that, by (.31), $\int|\frac{\partial\eta_\epsilon}{\partial x_k}(x-y)|\,\mathrm{d}y=C/\epsilon$, for some constant C independent by ϵ. Let $\varphi\geq 0$ be a C^∞ cutoff function compactly supported in $\mathbb{T}^N\setminus\Sigma_{2\delta}$, $\varphi\equiv 1$ in $\mathbb{T}^N\setminus\Sigma_\delta$. Using Jensen inequality and observing again that $|x-y|\leq\epsilon$ we have, by Proposition 7.12, for any ϵ sufficiently small,

$$\int_{\mathbb{T}^N\setminus\Sigma_\delta}\left|\int\left|\frac{\partial\eta_\epsilon}{\partial x_k}(x-y)\right|\left|D_pH(x,Du(y))-D_pH(x,Du(x))\right|\,\mathrm{d}y\right|^2\,\mathrm{d}\bar{\mu}_{\mathbb{T}^N}(x)$$
$$\leq C_0\int_{\mathbb{T}^N\setminus\Sigma_\delta}\left|\int\left|\frac{\partial\eta_\epsilon}{\partial x_k}(x-y)\right|\left|Du(y)-Du(x)\right|\,\mathrm{d}y\right|^2\,\mathrm{d}\bar{\mu}_{\mathbb{T}^N}(x)$$
$$\leq C_0\int_{\mathbb{T}^N\setminus\Sigma_\delta}\frac{C}{\epsilon}\int\left|\frac{\partial\eta_\epsilon}{\partial x_k}(x-y)\right|\left|Du(y)-Du(x)\right|^2\,\mathrm{d}y\,\mathrm{d}\bar{\mu}_{\mathbb{T}^N}(x)$$
$$\leq\frac{C}{\epsilon^2}\int\int_{\mathbb{T}^N\setminus\Sigma_\delta}|\tilde{\eta}_\epsilon(x-y)||Du(y)-Du(x)|^2\,\mathrm{d}\bar{\mu}_{\mathbb{T}^N}(x)\,\mathrm{d}y$$
$$\leq\frac{C}{\epsilon^2}\int|\tilde{\eta}_\epsilon(h)|\left(\int_{\mathbb{T}^N}\varphi^2(x)|Du(x+h)-Du(x)|^2\,\mathrm{d}\bar{\mu}_{\mathbb{T}^N}(x)\right)\,\mathrm{d}h$$
$$\leq\frac{C}{\epsilon^2}\int|\tilde{\eta}_\epsilon(h)||h^2|\,\mathrm{d}h\leq C\int|\tilde{\eta}_\epsilon(h)|\,\mathrm{d}h\leq C.$$

Acknowledgements

The authors are grateful to the Referee for many helpful suggestions for improvement of Sects. 6 and 7 of this paper.

References

[1] Bardi, M., Capuzzo-Dolcetta, I.: Optimal Control and Viscosity Solutions of Hamilton–Jacobi–Bellman Equations. Systems and Control: Foundations and Applications. Birkhäuser Boston Inc., Boston (1997) (With appendices by Maurizio Falcone and Pierpaolo Soravia)

[2] Barles, G.: Solutions de viscosité des équations de Hamilton–Jacobi. Mathématiques and Applications (Berlin) [Mathematics and Applications], vol. 17. Springer, Paris (1994)

[3] Barles, G., Briani, A., Chasseigne, E.: A Bellman approach for two-domains optimal control problems in $\mathbb{R}^n$. COCV. http://hal.archives-ouvertes.fr/hal-00652406

[4] Biryuk, A., Gomes, D.A.: An introduction to the Aubry-Mather theory. São Paulo J. Math. Sci. **4**(1), 17–63 (2010)

[5] Bolza, O.: Lectures on the Calculus of Variations, 2nd edn. Chelsea Publishing Co., New York (1961)

[6] Briani, A., Davini, A.: Monge solutions for discontinuous Hamiltonians. ESAIM Control Optim. Calc. Var. **11**(2), 229–251 (2005, electronic)

[7] Caffarelli, L., Crandall, M.G., Kocan, M., Swięch, A.: On viscosity solutions of fully nonlinear equations with measurable ingredients. Commun. Pure Appl. Math. **49**(4), 365–397 (1996)

[8] Camilli, F.: An Hopf-Lax formula for a class of measurable Hamilton–Jacobi equations. Nonlinear Anal. **57**(2), 265–286 (2004)

[9] Camilli, F., Siconolfi, A.: Time-dependent measurable Hamilton–Jacobi equations. Commun. Partial Differ. Equ. **30**(4–6), 813–847 (2005)

[10] Dal Maso, G., Frankowska, H.: Autonomous integral functionals with discontinuous nonconvex integrands: Lipschitz regularity of minimizers, DuBois-Reymond necessary conditions, and Hamilton–Jacobi equations. Appl. Math. Optim. **48**(1), 39–66 (2003)

[11] Davini, A.: Bolza problems with discontinuous Lagrangians and Lipschitz-continuity of the value function. SIAM J. Control Optim. **46**(5), 1897–1921 (2007)

[12] Evans, L.C.: Weak Convergence Methods for Nonlinear Partial Differential Equations. CBMS Regional Conference Series in Mathematics, vol. 74. Published for the Conference Board of the Mathematical Sciences, Washington, DC (1990)

[13] Evans, L.C., Gomes, D.: Effective Hamiltonians and averaging for Hamiltonian dynamics. I. Arch. Ration. Mech. Anal. **157**(1), 1–33 (2001)

[14] Fathi, A., Siconolfi, A.: Existence of C^1 critical subsolutions of the Hamilton–Jacobi equation. Invent. Math. **155**(2), 363–388 (2004)

[15] Fathi, A., Siconolfi, A.: PDE aspects of Aubry-Mather theory for quasiconvex Hamiltonians. Calc. Var. Partial Differ. Equ. **22**(2), 185–228 (2005)

[16] Figalli, A., Mandorino, V.: Fine properties of minimizers of mechanical lagrangians with sobolev potentials. Discrete Contin. Dyn. Syst. Ser. A **31**(4), 1325–1346 (2011)

[17] Gomes, D.: Viscosity solutions of Hamilton–Jacobi equations. Publicações Matemáticas do IMPA. [IMPA Mathematical Publications]. Instituto Nacional de Matemática Pura e Aplicada (IMPA), Rio de Janeiro, 2009. 27o Colóquio Brasileiro de Matemática. [27th Brazilian Mathematics Colloquium]

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

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

[20] Newcomb, R.T., Su, J.: Eikonal equations with discontinuities. Differ. Integral Equ. **8**(8), 1947–1960 (1995)

[21] Soner, H.M.: Optimal control with state-space constraint I & II. SIAM J. Control Optim., 24(3 & 6): 552–561, 1110–1122 (1986)

[22] Soravia, P.: Boundary value problems for Hamilton–Jacobi equations with discontinuous Lagrangian. Indiana Univ. Math. J **51**(2), 451–477 (2002)

[23] Soravia, P.: Degenerate eikonal equations with discontinuous refraction index. ESAIM Control Optim. Calc. Var. **12**(2), 216–230 (2006, electronic)

[24] Villani, C.: Topics in Optimal Transportation. Graduate Studies in Mathematics, vol. 58. American Mathematical Society, Providence (2003)

Diogo A. Gomes and Gabriele Terrone
Departamento de Matemática
Center for Mathematical Analysis, Geometry and Dynamical Systems
Instituto Superior Técnico
1049-001 Lisboa
Portugal
e-mail: gterrone@math.ist.utl.pt

Diogo A. Gomes
CSMSE Division
King Abdullah University of Science and Technology (KAUST)
Thuwal 23955-6900
Saudi Arabia
e-mail: dgomes@math.ist.utl.pt

Received: 2 March 2012.
Accepted: 15 July 2013.