WEAK KAM THEORETIC ASPECTS FOR NONREGULAR COMMUTING HAMILTONIANS

ANDREA DAVINI AND MAXIME ZAVIDOVIQUE

ABSTRACT. In this paper we introduce the notion of commutation for a pair of continuous and convex Hamiltonians, given in terms of commutation of their Lax–Oleinik semigroups. This is equivalent to the solvability of an associated multi–time Hamilton–Jacobi equation. We examine the weak KAM theoretic aspects of the commutation property and show that the two Hamiltonians have the same weak KAM solutions and the same Aubry set, thus generalizing a result recently obtained by the second author for Tonelli Hamiltonians. We make a further step by proving that the Hamiltonians admit a common critical subsolution, strict outside their Aubry set. This subsolution can be taken of class $C^{1,1}$ in the Tonelli case. To prove our main results in full generality, it is crucial to establish suitable differentiability properties of the critical subsolutions on the Aubry set. These latter results are new in the purely continuous case and of independent interest.

1. INTRODUCTION

In the last decades, the study of Hamiltonian systems has been dramatically impacted by a few new tools and methods. For general Hamiltonians, the framework of symplectic geometry led Gromov to his non–squeezing lemma [21], which gave rise to the key notion of symplectic capacity, now uniformly used in the field.

In the particular case of Tonelli (smooth, strictly convex, superlinear) Hamiltonians, some variational techniques led to tremendous improvements and results. John Mather led the way in this direction in [25, 26]. In the first paper he studied free time minimizers of the Lagrangian action functional, introducing the Aubry set, while in the second one he studied invariant minimizing measures, introducing what is now called the Mather set.

Later Fathi, through his weak KAM Theorem and Theory, showed the link between the variational sets introduced by Mather and the Hamilton–Jacobi equation. This allowed to simplify some proofs of Mather and establish new PDE results, such as the convergence of the solutions of the evolutive equation in the autonomous case [19]. This is all very well presented in [18].

The main challenge now seems to find analogues of the Aubry–Mather theory in wider settings. There are mainly two approaches to this problem. The first one is to lower the regularity of the Hamiltonians. This approach has been tackled successfully, let us mention for example [16, 20, 23]. The second, and much less fruitful approach, is to drop the convexity and coercivity assumptions, thus preventing from using traditional variational techniques.

2010 *Mathematics Subject Classification.* 35F21, 49L25, 37J50.
Key words and phrases. commuting Hamiltonians, viscosity solutions, weak KAM Theory.

In some sense, the study of commuting Hamiltonians makes a junction between these two worlds. Let us consider the multi–time Hamilton–Jacobi equation

$$\begin{cases} \dfrac{\partial u}{\partial t} + H(x, D_x u) = 0 & \text{in } (0,+\infty) \times (0,+\infty) \times M \\ \dfrac{\partial u}{\partial s} + G(x, D_x u) = 0 & \text{in } (0,+\infty) \times (0,+\infty) \times M \\ u(0,0,x) = u_0(x) & \text{on } M, \end{cases} \tag{1}$$

where M stands either for the Euclidean space $\mathbb{R}^N$ or the N–dimensional flat torus $\mathbb{T}^N$, H and G denote two real valued functions on $M \times \mathbb{R}^N$, and $u_0 : M \to \mathbb{R}$ is any given Lipschitz continuous initial datum. Equation (1) need not be solvable for general H and G. The first existence and uniqueness results appeared in [24] for Tonelli Hamiltonians independent of x via a representation formula for solutions of the Hamilton–Jacobi equation: the Hopf–Lax formula. Related problems were studied in [22].

A generalization of this result came much later in [4], where dependance in x is introduced (and the convexity hypothesis is kept). As a counterpart, the authors explain the necessity to impose the following commutation property on the Hamiltonians:

$$\langle D_p G,\, D_x H\rangle - \langle D_p H,\, D_x G\rangle = 0 \qquad \text{in } M \times \mathbb{R}^N. \tag{2}$$

Note that this condition is automatically satisfied when the Hamiltonians are independent of x. The proof involves an a priori different Hamilton–Jacobi equation with parameters and makes use of fine viscosity solution techniques. In [30], under stronger hypotheses, a more geometrical proof, following the original idea of Lions–Rochet, is given.

This equation was then studied under weaker regularity assumptions in [27]. The convexity is dropped in [8] in the framework of symplectic geometry and variational solutions. Finally, let us mention that in [29] the influence of first integrals (not necessarily of Tonelli type) on the dynamics of a Tonelli Hamiltonian is studied.

In [30] the second author has explored relation (2) for a pair of Tonelli Hamiltonians in the framework of weak KAM Theory to discover that the notions of Aubry set, of Peierls barrier and of weak KAM solution are invariants of the commutation property. Similar results were independently obtained in [9, 10].

This article deals with the first approach: we will consider purely continuous Hamiltonians, but we will keep the convexity and coercivity assumptions in the gradient variable.

The first difficulty is to define the notion of commutation for non–smooth Hamiltonians. Here, we will say two Hamiltonians H and G commute if the multi–time Hamilton–Jacobi equation (1) admits a viscosity solution for any Lipschitz initial datum. This is formulated in terms of commutation of their Lax–Oleinik semigroups, and is equivalent to (2) when the Hamiltonians are smooth enough, see [4] and Appendix B. We explore this definition by extending to this setting some properties holding in the regular case. We then generalize the main new results of [30] by proving

Theorem 1.1. *Let H and G be a pair of continuous, strictly convex and superlinear Hamiltonians on $\mathbb{T}^N \times \mathbb{R}^N$. If H and G commute, then they have the same weak KAM (or critical) solutions and the same Aubry set.*

Yet, the methods of [30] could not be adapted in a straightforward way. Actually, there was not even evidence the results would continue to hold by dropping the regularity. The proof in [30] is based on a careful study of the flows associated with H and G and exploits properties and tools developed in the framework of symplectic geometry and weak KAM Theory. The commutation hypothesis (2) entails a certain rigidity of the dynamics and of the underlying geometric frame of the equations. For all this, it is crucial to have strong regularity assumptions on the Hamiltonians. Even if in the purely continuous case some analogies can be drawn, all this rich structure seems lost.

The proof given here moves along the lines of [30], but the conclusion is reached through a simple remark on the time–dependent equations. We emphasize that this amounts to a considerable simplification of the argument even in the smooth case. Surprisingly, this allows us to obtain a new result for classical Tonelli Hamiltonians:

Theorem 1.2. *Let G and H be two commuting Tonelli Hamiltonians. Then they admit a $C^{1,1}$ critical subsolution which is strict outside their common Aubry set.*

In the end, our research reveals that the invariants observed in the framework of weak KAM Theory are consequence of the commutation of the Lax–Oleinik semigroups only, with no further reference to the Hamiltonian flows, that cannot be even defined in our setting. The only point where a kind of generalized dynamics plays a role is when we establish some differentiability properties of critical subsolutions on the Aubry set, which are crucial to state Theorem 1.1 in its full generality. These results are presented in Section 4, where we will prove a more precise version of the following

Theorem 1.3. *Let H be a continuous, strictly convex and superlinear Hamiltonian on $\mathbb{T}^N \times \mathbb{R}^N$. Then there exists a set $\mathcal{D} \subset \mathbb{T}^N$ such that any subsolution u of the critical Hamilton–Jacobi equation is differentiable on $\mathcal{D}$. Moreover, its gradient Du is independent of u on $\mathcal{D}$. Last, $\mathcal{D}$ is a uniqueness set for the critical equation, that is, if two weak KAM (or critical) solutions coincide on $\mathcal{D}$, then they are in fact equal.*

These latter results are new and we believe interesting *per se*. They generalize, in a weaker form, Theorem 7.8 in [20], and bring the hope of extending to the purely continuous case the results of [20] about the existence of a C^1 critical subsolution, strict outside the Aubry set. Such a generalization, however, seems out of reach without any further idea.

The article is organized as follows. In Section 2.1 we present the main notations and assumptions used throughout the paper, while in Section 2.2 we recall the definitions and the results about Hamilton–Jacobi equations that will be needed in the sequel. Section 3 consists in a brief overview of weak KAM Theory for non–regular Hamiltonians. Some proofs are postponed to Appendix A. In Section 4 we prove the differentiability properties of critical subsolutions above mentioned. Section 5 contains our main results for commuting, continuous Hamiltonians. More precisely, in Section 5.1 we give the definition of commutation in our setting and we explore its main properties. We also discuss its relation with the classical one, given in terms of cancellation of the Poisson bracket. The equivalence between these two definitions is proved in Appendix B in the case of Tonelli Hamiltonians. In Section 5.2 we examine the weak KAM theoretic aspects of the commutation property and we establish our main results for continuous and strictly convex Hamiltonians.

2. Preliminaries

2.1. **Notations and standing assumptions.** With the symbols $\mathbb{R}_+$ and $\mathbb{R}_-$ we will refer to the set of nonnegative and nonpositive real numbers, respectively. We say that a property holds *almost everywhere* (*a.e.* for short) on $\mathbb{R}^k$ if it holds up to a *negligible* subset, i.e. a subset of zero k–dimensional Lebesgue measure.

By modulus we mean a nondecreasing function from $\mathbb{R}_+$ to $\mathbb{R}_+$, vanishing and continuous at 0. A function $g : \mathbb{R}_+ \to \mathbb{R}$ will be termed *superlinear* if

$$\lim_{h \to +\infty} \frac{g(h)}{h} = +\infty.$$

Given a metric space X, we will write $\varphi_n \rightrightarrows \varphi$ on X to mean that the sequence of functions $(\varphi_n)_n$ uniformly converges to φ on compact subsets of X. Furthermore, we will denote by $\mathrm{Lip}(X)$ the family of Lipschitz–continuous real functions defined on X.

Throughout the paper, M will refer either to the Euclidean space $\mathbb{R}^N$ or to the N–dimensional flat torus $\mathbb{T}^N$, where N is an integer number. The scalar product in $\mathbb{R}^N$ will be denoted by $\langle \cdot\, , \cdot \rangle$, while the symbol $|\cdot|$ stands for the Euclidean norm. Note that the latter induces a norm on $\mathbb{T}^N$, still denoted by $|\cdot|$, defined as

$$|x| := \min_{\kappa \in \mathbb{Z}^N} |x + k| \qquad \text{for every } x \in \mathbb{T}^N.$$

We will denote by $B_R(x_0)$ and B_R the closed balls in M of radius R centered at x_0 and 0, respectively.

With the term *curve*, without any further specification, we refer to an absolutely continuous function from some given interval $[a, b]$ to M. The space of all such curves is denoted by $W^{1,1}([a, b]; M)$, while $\mathrm{Lip}_{x,y}([a, b]; M)$ stands for the family of Lipschitz–continuous curves γ joining x to y, i.e. such that $\gamma(a) = x$ and $\gamma(b) = y$, for any fixed x, y in M.

With the notation $\|g\|_\infty$ we will refer to the usual L^∞–norm of g, where the latter will be either a measurable real function on M or a vector–valued measurable map defined on some interval.

Given a continuous function u on M, we will denote by $D^+u(x)$ and $D^-u(x)$ the set of super and subdifferentials of u at $x \in M$, respectively. We recall that u is differentiable at x if and only if $D^+u(x)$ and $D^-u(x)$ are both nonempty. In this instance, $D^+u(x) = D^-u(x) = \{Du(x)\}$, where $Du(x)$ denotes the differential of u at x. We refer the reader to [7] for the precise definitions and proofs.

Throughout the paper, we will call *Hamiltonian* a function H satisfying the following set of assuptions:

(H1) $\quad H : M \times \mathbb{R}^N \to \mathbb{R} \qquad$ is continuous;

(H2) $\quad p \mapsto H(x, p) \qquad$ is convex on $\mathbb{R}^N$ for any $x \in M$;

(H3) $\quad$ there exist two superlinear functions $\alpha, \beta : \mathbb{R}_+ \to \mathbb{R}$ such that

$$\alpha(|p|) \leqslant H(x, p) \leqslant \beta(|p|) \qquad \text{for all } (x, p) \in M \times \mathbb{R}^N.$$

We define the *Fenchel transform* $L : M \times \mathbb{R}^N \to \mathbb{R}$ of H via

$$L(x,q) = H^*(x,q) := \sup_{p\in\mathbb{R}^N} \left\{ \langle p,q\rangle - H(x,p) \right\}. \tag{3}$$

The function L is called the Lagrangian associated with the Hamiltonian H; it satisfies the following properties, see Appendix A.2 in [7]:

(L1) $L : M \times \mathbb{R}^N \to \mathbb{R}$ is continuous;

(L2) $q \mapsto L(x,q)$ is convex on $\mathbb{R}^N$ for any $x \in M$;

(L3) there exist two superlinear functions $\alpha_*, \beta_* : \mathbb{R}_+ \to \mathbb{R}$ s.t.

$$\alpha_*(|q|) \leqslant L(x,q) \leqslant \beta_*(|q|) \qquad \text{for all } (x,q) \in M \times \mathbb{R}^N.$$

Remark 2.1. The functions α_*, β_* and α, β in (L3) and (H3), respectively, can be taken continuous (in fact, convex), without any loss of generality.

If (H2) is replaced by the following stronger assumption:

(H2)′ $p \mapsto H(x,p)$ is strictly convex on $\mathbb{R}^N$ for any $x \in M$,

then L enjoys

(L2)′ $q \mapsto L(x,q)$ is convex and of class C^1 on $\mathbb{R}^N$ for any $x \in M$.

Furthermore, the map $(x,q) \mapsto D_qL(x,q)$ is continuous in $M \times \mathbb{R}^N$. This fact will be exploited in the proof of Proposition 4.4. Here and in the sequel, $D_qL(x,q)$ and $D_xL(x,q)$ denote the partial derivative of L at (x,q) with respect to q and x, respectively. An analogous notation will be used for the Hamiltonian.

A *Tonelli Hamiltonian* is a particular kind of Hamiltonian satisfying conditions (H1), (H2)′ and (H3). It is additionally assumed of class C^2 in $M \times \mathbb{R}^N$ and condition (H2)′ is strengthen by requiring, for every $(x,p) \in M \times \mathbb{R}^N$, that

$$\frac{\partial^2 H}{\partial p^2}(x,p) \quad \text{is positive definite as a quadratic form.} \tag{4}$$

The associated Lagrangian has the same regularity as H and enjoys the analogous condition (4).

2.2. **Hamilton–Jacobi equations.** Let us consider a family of Hamilton–Jacobi equations of the kind

$$H(x, Du) = a \qquad \text{in } M, \tag{5}$$

where $a \in \mathbb{R}$. In the sequel, with the term *subsolution* (resp. *supersolution*) of (5) we will always refer to a continuous function u which is a subsolution (resp. a supersolution) of (5) in the viscosity sense, i.e. for every $x \in M$

$$\begin{aligned} & H(x,p) \leqslant a && \text{for any } p \in D^+u(x) \\ \big(\text{resp.}\quad & H(x,p) \geqslant a && \text{for any } p \in D^-u(x)\ \big). \end{aligned}$$

A function will be called a solution of (5) if it is both a subsolution and a supersolution.

Remark 2.2. Since H is coercive, i.e. satisfies the first inequality in (H3), it is well known that any continuous viscosity subsolution v of (5) is Lipschitz, see for instance [2]. In particular, v is an almost everywhere subsolution, i.e.

$$H(x, Dv(x)) \leqslant a \qquad \text{for a.e. } x \in M.$$

By the convexity assumption (H2), the converse holds as well: any Lipschitz, almost everywhere subsolution solves (5) in the viscosity sense, see [28]. In particular, v is a subsolution of (5) if and only if $-v$ is a subsolution of

$$H(x, -Du) = a \qquad \text{in } M.$$

We define the *critical value* c as

$$c = \min\{a \in \mathbb{R} \,:\, \text{equation (5) admits subsolutions}\}.$$

Following [20], we carry out the study of properties of subsolutions of (5), for $a \geqslant c$, by means of the semidistances S_a defined on $M \times M$ as follows:

$$S_a(x,y) = \inf\left\{ \int_0^1 \sigma_a(\gamma(s), \dot\gamma(s))\, \mathrm{d}s \,:\, \gamma \in \mathrm{Lip}_{x,y}([0,1]; M) \right\}, \tag{6}$$

where $\sigma_a(x,q)$ is the support function of the a–sublevel $Z_a(x)$ of H, namely

$$\sigma_a(x,q) := \sup\{\langle q, p\rangle \,:\, p \in Z_a(x)\} \tag{7}$$

and $Z_a(x) := \{p \in \mathbb{R}^N \,:\, H(x,p) \leqslant a\}$. The function $\sigma_a(x,q)$ is convex in q and upper semicontinuous in x (and even continuous at points such that $Z_a(x)$ has nonempty interior or reduces to a point), while S_a satisfies the following properties:

$$\begin{aligned} S_a(x,y) &\leqslant S_a(x,z) + S_a(z,y) \\ S_a(x,y) &\leqslant \kappa_a |x - y| \end{aligned}$$

for all $x, y, z \in M$ and for some positive constant κ_a. The following properties hold, see [20]:

Proposition 2.3. *Let $a \geqslant c$.*

(i) A function ϕ is a viscosity subsolution of (5) if and only if

$$\phi(x) - \phi(y) \leqslant S_a(y,x) \qquad \text{for all } x, y \in M.$$

In particular, all viscosity subsolutions of (5) are κ_a–Lipschitz continuous.

(ii) For any $y \in M$, the functions $S_a(y, \cdot)$ and $-S_a(\cdot, y)$ are both subsolutions of (5).

(iii) For any $y \in M$

$$S_a(y,x) = \sup\{v(x) \,:\, v \text{ is a subsolution to (5) with } v(y) = 0\}.$$

In particular, by maximality, $S_a(y, \cdot)$ is a viscosity solution of (5) in $M \setminus \{y\}$.

Definition 2.4. For $t > 0$ fixed, let us define the function $h^t : M \times M \to \mathbb{R}$ by

$$h^t(x,y) = \inf\left\{ \int_{-t}^0 L(\gamma, \dot\gamma)\, \mathrm{d}s \;:\; \gamma \in W^{1,1}([-t, 0]; M),\ \gamma(-t) = x,\ \gamma(0) = y \right\}. \tag{8}$$

It is well known, by classical result of Calculus of Variations, that the infimum in (8) is achieved. The curves that realize the minimum are called *Lagrangian minimizers*. The following more precise result will be needed in the sequel, see [1, 15]:

Proposition 2.5. *Let $x, y \in M$, $t > 0$ and $C \in \mathbb{R}$ such that $h^t(x,y) < t\,C$. Then any Lagrangian minimizer γ for $h^t(x,y)$ is Lipschitz continuous and satisfies $\|\dot\gamma\|_\infty \leqslant \kappa$, where κ is a constant only depending on C, α_*, β_*.*

We recall some properties of h^t, see for instance [15].

Proposition 2.6. *Let $t > 0$. Then h^t is locally Lipschitz continuous in $M \times M$. More precisely, for every $r > 0$ there exists $K = K(r, \alpha_*, \beta_*)$ such that the map*

$$(x, y, t) \mapsto h^t(x,y) \quad \textit{is } K\textit{–Lipschitz continuous in } C_r,$$

where $C_r := \{(x,y,t) \in M \times M \times (0,+\infty) \,:\, |x-y| < r\,t\}$.

We remark that, for every $a \geqslant c$, the following holds:

$$L(x,q) \geqslant \max_{p \in Z_a(x)} \langle p,\, q\rangle - H(x,p) \geqslant \sigma_a(x,q) - a \qquad \text{for every } (x,q) \in M \times \mathbb{R}^N, \tag{9}$$

yielding in particular $h^t(y,x) + a\,t \geqslant S_a(y,x)$ for every x, $y \in M$. The next result can be proved by making use of suitable reparametrization techniques, see [16, 20].

Lemma 2.7. *Let $a \geqslant c$. Then*

$$S_a(y,x) = \inf_{t>0} \Big(h^t(y,x) + at \Big) \qquad \textit{for every } x, y \in M,$$

and the infimum is always reached when $a > c$.

For every $t > 0$, we define a function on M as follows:

$$\big(\mathcal{S}(t)u\big)(x) = \inf\left\{ u(\gamma(0)) + \int_{-t}^{0} L(\gamma,\dot\gamma)\,\mathrm{d}s \;:\; \gamma \in W^{1,1}([-t,0];M),\ \gamma(0) = x \right\} \tag{10}$$

where $u : M \to \mathbb{R} \cup \{+\infty\}$ is an initial datum satisfying

$$u(\cdot) \geqslant a|\cdot| + b \qquad \text{on } M \tag{11}$$

for some a, $b \in \mathbb{R}$. Any function of this kind will be called *admissible initial datum* in the sequel.

The following properties hold:

Proposition 2.8.

(i) *For every admissible initial datum u, the map $(t,x) \mapsto \big(\mathcal{S}(t)u\big)(x)$ is finite valued and locally Lipschitz in $(0,+\infty) \times M$.*

(ii) *$\big(\mathcal{S}(t)\big)_{t>0}$ is a semigroup, i.e. for every admissible initial datum u*

$$\mathcal{S}(t)\big(\mathcal{S}(s)u\big) = \mathcal{S}(t+s)u \qquad \textit{for every } t,\ s > 0.$$

(iii) *$\mathcal{S}(t)$ is monotone and commutes with constants, i.e. for every admissible initial data u, v and any $a \in \mathbb{R}$ we have*

$$u \leqslant v \quad \Longrightarrow \quad \mathcal{S}(t)u \leqslant \mathcal{S}(t)v \qquad \textit{and} \qquad \mathcal{S}(t)(u+a) = \mathcal{S}(t)u + a.$$

In particular, $\mathcal{S}(t)$ is weakly contracting, i.e.

$$\|\mathcal{S}(t)u - \mathcal{S}(t)v\|_\infty \leqslant \|u - v\|_\infty.$$

(iv) *If $u \in \mathrm{Lip}(M)$, then the map $(t,x) \mapsto \big(\mathcal{S}(t)u\big)(x)$ is Lipschitz continuous in $[0,+\infty) \times M$ and*

$$\lim_{t \to 0^+} \|\mathcal{S}(t)u - u\|_\infty = 0.$$

The semigroup $\big(\mathcal{S}(t)\big)_{t>0}$ is called Lax–Oleinik semigroup and (10) is termed Lax–Oleinik formula. The relation with Hamilton–Jacobi equations is clarified by the next results.

Theorem 2.9. *Let H satisfy assumptions (H1), (H2), (H3) and let L be its Fenchel transform. Then, for every $u_0 \in \mathrm{Lip}(M)$, the Cauchy Problem*

$$\begin{cases} \partial_t u + H(x, Du) = 0 & \text{in } (0, +\infty) \times M \\ u(0, x) = u_0(x) & \text{on } M \end{cases} \tag{12}$$

admits a unique viscosity solution $u(t, x)$ in $\mathrm{Lip}([0, +\infty) \times M)$. Moreover,

$$u(t, x) = \big(\mathcal{S}(t)u_0\big)(x) \qquad \text{for every } (t, x) \in (0, +\infty) \times M.$$

With regard to the stationary equation (5), the following characterization holds:

Proposition 2.10. *Let u be a continuous function on M. The following facts hold:*

(i) u is a subsolution of (5) if and only if $t \mapsto \mathcal{S}(t)u + a\,t$ is non decreasing;

(ii) u is a solution of (5) if and only if $u \equiv S(t)u + a\,t$ for every $t > 0$.

Proof. *(i)* If u is a subsolution of (5), then for every x, $y \in M$

$$u(x) \leqslant u(y) + S_a(y, x) \leqslant u(y) + h^t(y, x) + a\,t \quad \text{for every } t > 0,$$

hence

$$u \leqslant \inf_{y \in M} \big(u(y) + h^t(y, \cdot) + a\,t\big) = \mathcal{S}(t)u + a\,t.$$

This readily implies, by monotonicity of the semi–group,

$$\mathcal{S}(h)u + a\,h \leqslant \mathcal{S}(t + h)u + a\,(t + h) \quad \text{for every } h > 0,$$

i.e. $t \mapsto \mathcal{S}(t)u + a\,t$ is non decreasing.

Conversely, if $t \mapsto \mathcal{S}(t)u + a\,t$ is non decreasing, then for every fixed x, $y \in M$ we have

$$u(x) \leqslant u(y) + h^t(y, x) + a\,t \qquad \text{for every } t > 0.$$

By taking the infimum in t of the right–hand side term, we obtain $u(x) - u(y) \leqslant S_a(y, x)$ for every x, $y \in M$ by Lemma 2.7, i.e. u is a subsolution of (5).

Assertion *(ii)* easily follows by noticing that u is a solution of (5) if and only if $u(x) - at$ is a solution of (12) with $u_0 := u$. □

We conclude this section with a theorem which was first proved in the regular case by Fathi [19] and subsequently generalized to this setting in [3, 16].

Theorem 2.11. *Let H satisfy conditions (H1), (H2)$'$, (H3). Then, for any continuous initial datum $u : \mathbb{T}^N \to \mathbb{R}$,*

$$\mathcal{S}(t)u + ct \underset{t \to +\infty}{\rightrightarrows} v \qquad \text{in } \mathbb{T}^N,$$

where v is the critical solution defined as

$$v(x) = \min_{y \in \mathcal{A}} \left(S(y, x) + \min_{z \in \mathbb{T}^N} \big(u(z) + S(z, y)\big) \right) \quad \text{for every } x \in \mathbb{T}^N. \tag{13}$$

Here, $\mathcal{A}$ denotes the Aubry set; its precise definition will be given at the beginning of the next section.

3. Nonregular weak KAM theory

The purpose of this Section is to present the main results of weak KAM Theory we are going to use in the sequel. This material is not new. It is well known for Tonelli Hamiltonians, see [18], while the extension to the non regular setting is either contained in other papers or can be easily recovered from the results proved therein. Nevertheless, it is less standard and it is not always possible to give precise references. For the reader's convenience, we provide here a brief presentation. Some proofs are postponed to the Appendix.

Throughout this Section, M stands either for $\mathbb{R}^N$ or for $\mathbb{T}^N$ and conditions (H1), (H2) and (H3) are assumed.

We focus our attention on the critical equation

$$H(x, Du) = c \qquad \text{in } M. \tag{14}$$

A subsolution, supersolution or solution of (14) will be termed critical in the sequel. To ease notations, we will moreover write S and σ in place of S_c and σ_c, respectively. Finally, by possibly considering $H - c$ instead of H, we will assume $c = 0$.

We define the *Aubry set* $\mathcal{A}$ as

$$\mathcal{A} := \{ y \in M \,:\, S(y, \cdot) \text{ is a critical solution} \}.$$

In the sequel, we will sometimes write S_y to denote the function $S(y, \cdot)$.

We will assume that the following holds

($\mathcal{A}$) $\quad$ $\mathcal{A}$ is nonempty.

This condition is always fulfilled when M is compact, but it may be false in the non compact case.

We define the set $\mathcal{E}$ of equilibrium points as

$$\mathcal{E} := \{ y \in M \,:\, \min_p H(y, p) = 0 \}.$$

This set may be empty, but if not it is a closed subset of the Aubry set $\mathcal{A}$.

Next, we define a family of curves, called *static*. In the next Section we will investigate the behavior of the critical subsolutions on such curves.

Definition 3.1. A curve γ defined on an interval J is called *static* if

$$S(\gamma(t_1), \gamma(t_2)) = \int_{t_1}^{t_2} L(\gamma, \dot\gamma)\, \mathrm{d}s = -S(\gamma(t_2), \gamma(t_1))$$

for every t_1, t_2 in J with $t_2 > t_1$.

We first show that static curves are always contained in the Aubry set.

Lemma 3.2. *Let γ be a static curve defined on some interval J. Then γ is contained in the Aubry set and satisfies*

$$L(\gamma(s), \dot\gamma(s)) = \sigma(\gamma(s), \dot\gamma(s)) \qquad \textit{for a.e. } s \in J. \tag{15}$$

Proof. The definition of the semidistance S and inequality (9) with $a = c$ readily implies that γ enjoys (15).

Let us prove that γ is contained in the Aubry set. If γ is a steady curve, i.e. $\gamma(t) = y$ for every $t \in J$, then for $(a, b) \subset J$ we get

$$(b - a)\, L(y, 0) = \int_a^b L(\gamma, \dot\gamma)\, \mathrm{d}s = S(y, y) = 0,$$

yielding that $y \in \mathcal{E} \subseteq \mathcal{A}$ for $L(y,0) = -\min_{\mathbb{R}^N} H(y,\cdot)$.

Let us then assume that γ is nonsteady. We want to prove that, for every fixed $t \in J$, the point $y := \gamma(t)$ belongs to $\mathcal{A}$, i.e. that $S(y,\cdot)$ is a critical solution on M. Of course, we just need to check that $S(y,\cdot)$ is a supersolution of (14) at y, by Proposition 2.3. To this purpose, choose a point $z \in \gamma(J)$ with $z \neq y$. Since γ is static, we have

$$S(y,z) + S(z,y) = 0.$$

This and the triangular inequality imply that the function $w(\cdot) = S(y,z) + S(z,\cdot)$ touches $S(y,\cdot)$ from above at y, hence $D^- S_y(y) \subseteq D^- w(y)$. Since w is a viscosity solution in $M \setminus \{z\}$ we derive

$$H(y,p) \geqslant 0 \qquad \text{for every } p \in D^- S_y(y),$$

that is, S_y is a supersolution of (14) at y and so a critical solution on M. □

The next result states that static curves fully cover the Aubry set.

Theorem 3.3. *Let $y \in \mathcal{A}$, then there exists a static curve η defined on $\mathbb{R}$ with $\eta(0) = y$.*

This result is proved in [16] by exploiting some ideas contained in [20]. A more concise and self–contained proof of this fact is proposed in the Appendix A.

We denote by $\mathcal{K}$ the family of all static curves defined on $\mathbb{R}$, and by $\mathcal{K}(y)$ the subset of $\mathcal{K}$ made up by those equaling y at $t = 0$.

The *Peierls barrier* is the function $h : M \times M \to \mathbb{R}$ defined by

$$h(x,y) = \liminf_{t \to +\infty} h^t(x,y). \tag{16}$$

The following holds:

Theorem 3.4. $\mathcal{A} = \{ y \in M \, : \, h(y,y) = 0 \}$.

Proof. Take $y \in M$ such that $h(y,y) = 0$ and set $u(\cdot) := S(y,\cdot)$. We want to prove that u is a critical solution in M; equivalently, by Proposition 2.3, that u is a critical supersolution at y. To this purpose, we first note that, since u is a critical subsolution on M, the functions $\mathcal{S}(t)u$ are increasing in t, see Proposition 2.10, and equi–Lipschitz in x for u is Lipschitz continuous, see Proposition 2.8. Let us set

$$v(x) = \sup_{t>0} \big(\mathcal{S}(t)u\big)(x) = \lim_{t \to +\infty} \big(\mathcal{S}(t)u\big)(x) \qquad \text{for every } x \in M.$$

According to what was remarked above, $v \geqslant u$. Furthermore, v is Lipschitz continuous provided it is finite everywhere, or, equivalently, at some point. We claim that $v(y) = u(y)$.

Indeed, let $(t_n)_{n \in \mathbb{N}}$ be a diverging sequence such that $\lim_n h^{t_n}(y,y) = h(y,y) = 0$. By definition of $\mathcal{S}(t)$ we have

$$\big(\mathcal{S}(t_n)u\big)(y) \leqslant u(y) + h^{t_n}(y,y) \qquad \text{for each } n \in \mathbb{N},$$

hence

$$v(y) = \lim_{n \to +\infty} \big(\mathcal{S}(t_n)u\big)(y) \leqslant \lim_{n \to +\infty} \big(u(y) + h^{t_n}(y,y)\big) = u(y),$$

as it was claimed. This also implies that v touches u from above at y, yielding $D^- u(y) \subseteq D^- v(y)$. Furthermore, v is a critical solution since it is a fixed point of the (continuous) semigroup $\mathcal{S}(t)$, see Proposition 2.10, in particular it is a critical supersolution at y. Collecting the information, we conclude that

$$H(y,p) \geqslant 0 \qquad \text{for every } p \in D^- u(y),$$

finally showing that u is a critical supersolution at y.

Let us prove the opposite inclusion. Take $y \in \mathcal{A}$. To prove that $h(y,y) = 0$, it will be enough, in view of Lemma 2.7, to find a diverging sequence $(t_n)_{n \in \mathbb{N}}$ such that $\liminf_n h^{t_n}(y,y) = 0$.

To this purpose, let $\eta \in \mathcal{K}(y)$. Then

$$-S(\eta(n), y) = \int_0^n L(\eta, \dot\eta)\, \mathrm{d}s = S(y, \eta(n))$$

for each $n \in \mathbb{N}$. By Lemma 2.7 there exist $s_n > 0$ such that

$$S(\eta(n), y) \leqslant h^{s_n}(\eta(n), y) < S(\eta(n), y) + \frac{1}{n}.$$

By definition of h^t we get

$$h^{n+s_n}(y,y) \leqslant h^n(y, \eta(n)) + h^{s_n}(\eta(n), y) < S(y, \eta(n)) + S(\eta(n), y) + \frac{1}{n} = \frac{1}{n},$$

and the assertion is proved by taking $t_n := n + s_n$. □

Here and in the remainder of the paper, by $\check{H}$ we will denote the Hamiltonian defined as

$$\check{H}(x,p) := H(x,-p) \qquad \text{for every } (x,p) \in M \times \mathbb{R}^N.$$

The following holds:

Proposition 3.5. *The Hamiltonians H and $\check{H}$ have the same critical value and the same Aubry set.*

Proof. The fact that H and $\check{H}$ have the same critical value immediately follows from the definition in view of Remark 2.2. Furthermore, the Peierls barrier $\check{h}$ associated with $\check{H}$ enjoys $\check{h}(x,y) = h(y,x)$ for every x, $y \in M$. Hence H and $\check{H}$ have the same Aubry set in view of Theorem 3.4. □

We end this section by proving some important properties of the Peierls barrier.

Proposition 3.6. *Under assumption ($\mathcal{A}$) the following properties hold:*

- *(i)* *h is finite valued and Lipschitz continuous.*
- *(ii)* *If v is a critical subsolution, then $h(y,x) \geqslant v(x) - v(y)$ for every $x, y \in M$.*
- *(iii)* *For every $x, y, z \in M$ and $t > 0$*

  $$h(y,x) \leqslant h(y,z) + h^t(z,x) \quad \textit{and} \quad h(y,x) \leqslant h^t(y,z) + h(z,x).$$

  In particular, $h(y,x) \leqslant h(y,z) + h(z,x)$.
- *(iv)* *$h(x,y) = S(x,y)$ if either x or y belong to $\mathcal{A}$.*
- *(v)* *$h(y,\cdot)$ is a critical solution for every fixed $y \in M$.*

Furthermore, when M is compact and condition (H2)′ is assumed, we have

$$h^t \underset{t \to +\infty}{\rightrightarrows} h \qquad \textit{in } M \times M.$$

Proof. *(i)* Let K_1 be the constant given by Proposition 2.6 with $r = 1$. It is easily seen that for every bounded open set $V \subset M \times M$ there exists t_V such that the functions $\{h^t : t \geqslant t_V\}$ are K_1–Lipschitz continuous in V. Moreover we already know, by Theorem 3.4, that $h(y, y) = 0$ for every $y \in \mathcal{A}$. This implies that h is finite valued and Lipschitz–continuous on the whole $M \times M$.

Items *(ii)* and *(iii)* follow directly from the definition of h and from assertion *(ii)* in Proposition 2.6.

(iv) Let us assume, for definiteness, that $y \in \mathcal{A}$. Let $(t_n)_n$ be a diverging sequence such that $0 \leqslant h^{t_n}(y, y) < 1/n$. Then for every $t > 0$ and $n \in \mathbb{N}$

$$S(x, y) \leqslant h^{t+t_n}(x, y) \leqslant h^t(x, y) + h^{t_n}(y, y) \leqslant h^t(x, y) + \frac{1}{n},$$

yielding

$$S(x, y) \leqslant \liminf_{t\to+\infty} h^t(x, y) \leqslant \inf_{t>0} h^t(x, y) = S(x, y)$$

in view of Lemma 2.7.

(v) By Proposition 2.10–*(ii)*, it suffices to prove that $\mathcal{S}(t)h_y = h_y$ for every fixed $t > 0$ and $y \in M$, where h_y denotes the function $h(y, \cdot)$. First notice that, by *(iii)* and Lemma 2.7,

$$h_y(x) - h_y(z) \leqslant \inf_{t>0} h^t(z, x) = S(z, x),$$

that is, h_y is a critical subsolution. By Proposition 2.10–*(i)*, that implies $\mathcal{S}(t)h_y \geqslant h_y$.

Let us prove the reverse inequality. For any fixed $x \in M$, pick a diverging sequence $(t_n)_{n\in\mathbb{N}}$ with $t_n > t$ for every $n \in \mathbb{N}$ and a family of curves $\gamma_n : [-t_n, 0] \to M$ connecting y to x such that $h^{t_n}(y, x) = \int_{-t_n}^{0} L(\gamma_n, \dot\gamma_n)\,\mathrm{d}s$ and

$$\lim_{n\to+\infty} \int_{-t_n}^{0} L(\gamma_n, \dot\gamma_n)\,\mathrm{d}s = h(y, x). \tag{17}$$

The functions h^{t_n} are equi–Lipschitz, see Proposition 2.6. This yields, by Proposition 2.5, that the curves γ_n are equi–Lipschitz. Up to extraction of a subsequence, we can then assume that there is a curve $\gamma : [-t, 0] \to M$ such that

$$\gamma_n \rightrightarrows \gamma \quad \text{in } [-t, 0] \qquad \text{and} \qquad \dot\gamma_n \rightharpoonup \dot\gamma \quad \text{in} \quad L^1\big([-t, 0]; \mathbb{R}^N\big).$$

Set $z = \gamma(-t)$. By a classical semi–continuity result of the Calculus of Variations [6], we have

$$\begin{aligned} h_y(x) = \liminf_{n\to+\infty} &\int_{-t_n}^{0} L(\gamma_n, \dot\gamma_n)\,\mathrm{d}s \\ &\geqslant \liminf_{n\to+\infty} \int_{-t_n}^{-t} L(\gamma_n, \dot\gamma_n)\,\mathrm{d}s \,+\, \liminf_{n\to+\infty} \int_{-t}^{0} L(\gamma_n, \dot\gamma_n)\,\mathrm{d}s \\ &\qquad\qquad \geqslant h(y, z) + \int_{-t}^{0} L(\gamma, \dot\gamma)\,\mathrm{d}s \geqslant \big(\mathcal{S}(t)h_y\big)(x). \end{aligned}$$

Last, let us show that h^t uniformly converges to h for $t \to +\infty$ when M is compact and condition (H2)$'$ is assumed. Let $y \in M$ be fixed. Then $h^t(y, \cdot)$ uniformly converges to $h(y, \cdot)$ in view of Theorem 2.11 and of the equality $h^t(y, \cdot) = \mathcal{S}(t-1)u$ with $u = h^1(y, \cdot)$. Since y was arbitrarily chosen in M, we have pointwise convergence of h^t to h. The assertion follows by Ascoli–Arzelà Theorem since the functions

$\{h^t : t \geqslant 1\}$ are equi–Lipschitz in $M \times M$ in view of Proposition 2.6. $\square$

4. Differentiability properties of critical subsolutions

The purpose of this Section is to prove some differentiability properties of critical subsolutions on the Aubry set. These results will be exploited in the subsequent section to obtain some information for commuting Hamiltonians.

Let us consider, for any fixed $t > 0$, the locally Lipschitz function defined on M as

$$\big(\mathcal{S}(t)u\big)(\cdot) = \inf_{z \in M} \big(u(z) + h^t(z, \cdot)\big),$$

where u is an admissible initial datum. If the latter is additionally assumed continuous, then the infimum is actually a minimum, and, as previously noticed, for every fixed $y \in M$ there exists a Lipschitz curve $\gamma : [-t, 0] \to M$ with $\gamma(0) = y$ such that

$$\big(\mathcal{S}(t)u\big)(y) = u(\gamma(-t)) + \int_{-t}^{0} L(\gamma, \dot\gamma)\,\mathrm{d}s.$$

As first step in our analysis, we prove some differentiability properties of $\mathcal{S}(t)u$ at y and of u at $\gamma(-t)$ in terms of γ, thus generalizing to this setting some known results in the regular case, see [18].

We start by dealing with the case when the Hamiltonian is independent of x. We need a lemma first.

Lemma 4.1. *Let H be independent of x and satisfy assumptions (H1), (H2)′ and (H3). Then any (Lipschitz) Lagrangian minimizer $\gamma : [-t, 0] \to M$ with $t > 0$ satisfies*

$$D_q L(\dot\gamma(s)) = D_q L\left(\frac{\gamma(0) - \gamma(-t)}{t}\right) \qquad \text{for a.e. } s \in [-t, 0], \tag{18}$$

with equality holding for every s if γ is of class C^1.

Remark 4.2. We remark for later use that, since equality (18) holds for almost every $s \in [-t, 0]$, then it holds in particular for every s that is both a differentiability point of γ and a Lebesgue point of $D_q L(\dot\gamma(\cdot))$ in $[-t, 0]$.

Proof. Let us set

$$v := \frac{\gamma(0) - \gamma(-t)}{t} \qquad \text{and} \qquad \eta(s) = \gamma(0) + sv.$$

It is easy to see, by the convexity of L and Jensen's inequality, that

$$\int_{-t}^{0} L(\dot\gamma)\,\mathrm{d}s \geqslant t\,L(v) = \int_{-t}^{0} L(\dot\eta)\,\mathrm{d}s,$$

while the converse inequality is true since γ is a Lagrangian minimizer. By exploiting the convexity of L again, we get

$$L(q) \geqslant L(v) + \langle D_q L(v), q - v\rangle \qquad \text{for every } q \in \mathbb{R}^N. \tag{19}$$

On the other hand,

$$\int_{-t}^{0} L(\dot\gamma(s))\,\mathrm{d}s = t\,L(v) = \int_{-t}^{0} \Big(L(v) + \langle D_q L(v), \dot\gamma(s) - v\rangle\Big)\,\mathrm{d}s,$$

meaning that we have an equality in (19) at $\dot\gamma(s)$ for a.e. $s \in [-t,0]$. Equality (18) follows by differentiability of L. □

Proposition 4.3. *Let H be independent of x and satisfy assumptions (H1), (H2)′ and (H3). Let u be an admissible initial datum and $\gamma : [-t,0] \to M$ a Lipschitz continuous curve such that $\gamma(0) = y$ and*

$$\big(\mathcal{S}(t)u\big)(y) = u(\gamma(-t)) + \int_{-t}^{0} L(\dot\gamma(s))\,\mathrm{d}s$$

for some $t > 0$ and $y \in M$. Then

$$D_q L\left(\frac{\gamma(0)-\gamma(-t)}{t}\right) \in D^+\big(\mathcal{S}(t)u\big)(y) \quad \textit{and} \quad D_q L\left(\frac{\gamma(0)-\gamma(-t)}{t}\right) \in D^- u(\gamma(-t)).$$

Proof. To ease notations, we set

$$v := \frac{\gamma(0)-\gamma(-t)}{t}$$

and denote by z the point $\gamma(-t)$. Let us first prove that $D_q L\,(v) \in D^+\big(\mathcal{S}(t)u\big)(y)$. According to the proof of Lemma 4.1, it is enough to prove the assertion when γ is the segment joining z to y. For every $x \in M$, we define a curve $\gamma_x : [-t,0] \to M$ joining z to x by setting $\gamma_x(s) = \gamma(s) + (s+t)(x-y)/t$. Let

$$\varphi(x) := u(z) + \int_{-t}^{0} L(\dot\gamma_x)\,\mathrm{d}s, \qquad x \in M.$$

Then $\big(\mathcal{S}(t)u\big)(\cdot) \leqslant \varphi(\cdot)$ with equality holding at y. It is easy to see, using the local Lipschitz character of L, that φ is locally Lipschitz continuous. We want to show that $D_q L(v) \in D^+\varphi(y)$, which clearly implies the assertion as $D^+\varphi(y) \subseteq D^+\big(\mathcal{S}(t)u\big)(y)$.

By the standard result of differentiation under the integral sign, the function φ is in fact C^1 and we may compute its differential at y by the following formula :

$$D\varphi(y) = \left(\int_{-t}^{0} \frac{\partial}{\partial x} L(\dot\gamma_x)\,\mathrm{d}s\right)\Big|_{x=y} = D_q L\,(v)\,.$$

Let us now prove that $D_q L(v) \in D^- u(z)$.

For every $x \in M$, we define a curve $\eta_x : [-t,0] \to M$ joining x to y by setting $\eta_x(s) := \gamma(s) + s\,(z-x)/t$. Let

$$\psi(x) := -\int_{-t}^{0} L(\dot\eta_x)\,\mathrm{d}s + \big(\mathcal{S}(t)u\big)(y), \qquad x \in M.$$

Then $\psi(\cdot) \leqslant u(\cdot)$ with equality holding at z. We want to show that $D_q L(v) \in D^-\psi(z)$, which is enough to conclude as $D^-\psi(z) \subseteq D^- u(z)$. Arguing as above, we actually see that ψ is in fact C^1 and

$$D\psi(z) = D_q L\,(v)\,.$$

This conclude the proof. □

We proceed to show a more general version of the previous result.

Proposition 4.4. *Let H satisfy assumptions (H1), (H2)$'$ and (H3) and u be an admissible initial datum. Let $\gamma : [-t, 0] \to M$ be a Lipschitz continuous curve with $\gamma(0) = y$ such that*

$$\big(\mathcal{S}(t)u\big)(y) = u(\gamma(-t)) + \int_{-t}^{0} L(\gamma(s), \dot\gamma(s))\,\mathrm{d}s$$

for some $t > 0$ and $y \in M$. The following holds:

- *(i) if 0 is a differentiability point for γ and a Lebesgue point for $D_q L(\gamma(\cdot), \dot\gamma(\cdot))$, then $D_q L(\gamma(0), \dot\gamma(0)) \in D^+\big(\mathcal{S}(t)u\big)(y)$.*
- *(ii) Assume $u \in \mathrm{Lip}(M)$. If $-t$ is a differentiability point for γ and a Lebesgue point for $D_q L(\gamma(\cdot), \dot\gamma(\cdot))$, then $D_q L(\gamma(-t), \dot\gamma(-t)) \in D^- u(\gamma(-t))$.*

Proof. Let us choose an $R > 1$ sufficiently large in such a way that $\|\dot\gamma\|_\infty \leqslant R$ and $\gamma\big([-t, 0]\big) \subseteq B_R$. To ease notations, in the sequel we will call z the point $\gamma(-t)$.

Let $\omega : \mathbb{R}_+ \to \mathbb{R}_+$ be a modulus such that

$$|L(x, q) - L(y, q)| \leqslant \omega(|x - y|)^2 \qquad \text{for every } x, y \in B_{2R} \text{ and } q \in B_{2R}.$$

If $\omega(h) = O(h)$ then $L(x, q) = L(q)$ on $B_{2R} \times B_{2R}$, and the assertion follows from Proposition 4.3 when γ is the segment joining z to y, and from Remark 4.2 when γ is any Lipschitz continuous minimizer.

Let us then assume $\omega(h)/h$ is unbounded. Without loss of generality, we may require ω to be concave, in particular

$$\frac{\omega(h)}{h} \to +\infty \qquad \text{as } h \to 0^+.$$

Let $\delta : (0, +\infty) \to (0, +\infty)$ be such that

$$\delta(h)\,\omega(h) = h \qquad \text{for every } h > 0,$$

i.e.

$$\delta(h) := \frac{h}{\omega(h)} \qquad \text{for every } h > 0.$$

Since $\mathcal{S}(t)u$ is Lipschitz in M, to prove assertion *(i)* it is in fact enough to show that the following inequality holds for every $\xi \in \partial B_R$:

$$\big(\mathcal{S}(t)u\big)(y + h\xi) - \big(\mathcal{S}(t)u\big)(y) \leqslant h\,\langle D_q L(\gamma(0), \dot\gamma(0)), \xi\rangle + o(h) \qquad \text{for } h \to 0^+. \tag{20}$$

To this purpose, for every $h \in [0, 1]$ and for every $\xi \in \partial B_1$ we define a Lipschitz curve $\gamma_{h\xi} : [-t, 0] \to M$ joining z to $y + h\xi$ by setting

$$\gamma_{h\xi}(s) := \begin{cases} \gamma(s) & \text{if } s \in [-t, -\delta(h)] \\ \gamma(s) + \omega(h)\big(\delta(h) + s\big)\xi & \text{if } s \in [-\delta(h), 0]. \end{cases}$$

By definition of $\big(\mathcal{S}(t)u\big)$, we get

$$\big(\mathcal{S}(t)u\big)(y + h\xi) - \big(\mathcal{S}(t)u\big)(y) \leqslant \int_{-\delta(h)}^{0} \Big(L(\gamma_{h\xi}, \dot\gamma_{h\xi}) - L(\gamma, \dot\gamma)\Big)\,\mathrm{d}t \tag{21}$$

$$= \underbrace{\int_{-\delta(h)}^{0} \Big(L(\gamma_{h\xi}, \dot\gamma_{h\xi}) - L(\gamma, \dot\gamma_{h\xi})\Big)\,\mathrm{d}t}_{A} + \underbrace{\int_{-\delta(h)}^{0} \Big(L(\gamma, \dot\gamma_{h\xi}) - L(\gamma, \dot\gamma)\Big)\,\mathrm{d}t}_{B}.$$

For h small enough we have

$$
\begin{aligned}
|\gamma_{h\xi}(t) - \gamma(t)| \leqslant h < R \qquad & \text{for every } t \in [-\delta(h), 0], \\
|\dot\gamma_{h\xi}(t)| = |\dot\gamma(t) + \omega(h)\xi| < 2R \quad & \text{for a.e. } t \in [-\delta(h), 0],
\end{aligned}
$$

hence

$$
|L(\gamma_{h\xi}, \dot\gamma_{h\xi}) - L(\gamma, \dot\gamma_{h\xi})| \leqslant \omega(h)^2 \quad \text{for a.e. } t \in [-\delta(h), 0].
$$

This yields

$$
A \leqslant \delta(h)\, \omega(h)^2 = h\, \omega(h). \tag{22}
$$

To evaluate B, we use the Taylor expansion of $L(\gamma, \dot\gamma_{h\xi})$ to get

$$
L(\gamma, \dot\gamma + \omega(h)\xi) \leqslant L(\gamma, \dot\gamma) + \omega(h)\, \langle D_q L(\gamma, \dot\gamma), \xi\rangle + \omega(h)\, \Theta(\omega(h))
$$

for a.e. $t \in [-\delta(h), 0]$, where Θ is a continuity modulus for $D_q L$ on $B_{2R} \times B_{2R}$. From this we obtain

$$
\begin{aligned}
B \ &\leqslant\ \omega(h) \int_{-\delta(h)}^{0} \langle D_q L(\gamma, \dot\gamma), \xi\rangle \,\mathrm{d}t + \delta(h)\, \omega(h) \Theta(\omega(h)) \\
&\leqslant\ h\, \langle D_q L(\gamma(0), \dot\gamma(0)), \xi\rangle + h \fint_{-\delta(h)}^{0} |D_q L(\gamma, \dot\gamma) - D_q L(\gamma(0), \dot\gamma(0))| \,\mathrm{d}t + h\, \Theta(\omega(h)),
\end{aligned}
$$

i.e.

$$
B \leqslant h\, \langle D_q L(\gamma(0), \dot\gamma(0)), \xi\rangle + o(h) \tag{23}
$$

by recalling that $t = 0$ is a Lebesgue point for $D_q L(\gamma(\cdot), \dot\gamma(\cdot))$. Relations (22) and (23) together with (21) finally give (20).

To prove *(ii)*, it suffices to show, by the Lipschitz character of u, that for every fixed $\xi \in \partial B_1$

$$
u(y + h\xi) - u(y) \geqslant h\, \langle D_q L(\gamma(0), \dot\gamma(0)), \xi\rangle + o(h) \qquad \text{for } h \to 0^+. \tag{24}
$$

To this purpose, for every $h \in [0, 1]$ and for every $\xi \in \partial B_1$ we define a Lipschitz curve $\eta_{h\xi} : [-t, 0] \to M$ joining $z + h\xi$ to y by setting

$$
\eta_{h\xi}(s) := \begin{cases} \gamma(s) + \omega(h)\big(\delta(h) - t - s\big)\xi & \text{if } s \in [-t,\, -t + \delta(h)] \\ \gamma(s) & \text{if } s \in [-t + \delta(h),\, 0]. \end{cases}
$$

By definition of $\big(\mathcal{S}(t)u\big)(y)$, we get

$$
\begin{aligned}
u(z + h\xi) - u(z) &\geqslant \int_{-t}^{-t+\delta(h)} \Big(L(\gamma, \dot\gamma) - L(\eta_{h\xi}, \dot\eta_{h\xi})\Big) \,\mathrm{d}t \\
&= \underbrace{\int_{-t}^{-t+\delta(h)} \Big(L(\gamma, \dot\gamma) - L(\gamma, \dot\eta_{h\xi})\Big) \,\mathrm{d}t}_{A'} + \underbrace{\int_{-\delta(h)}^{0} \Big(L(\gamma, \dot\eta_{h\xi}) - L(\eta_{h\xi}, \dot\eta_{h\xi})\Big) \,\mathrm{d}t}_{B'}.
\end{aligned}
$$

To evaluate B', we argue as above to get $B' \geqslant -h\omega(h)$. To evaluate A', we use the Taylor expansion of $L(\gamma, \dot\eta_{h\xi})$ to get

$$
L(\gamma, \dot\gamma - \omega(h)\xi) \leqslant L(\gamma, \dot\gamma) - \omega(h)\, \langle D_q L(\gamma, \dot\gamma), \xi\rangle + \omega(h)\, \Theta(\omega(h))
$$

for a.e. $t \in [-\delta(h), 0]$. Arguing as above we finally get

$$
A' \geqslant h\, \langle D_q L(\gamma(0), \dot\gamma(0)), \xi\rangle + o(h),
$$

and (24) follows. □

We now exploit the information gathered to deduce some differentiability properties of critical subsolutions. In what follows, we stress the fact that we have assumed the critical value c equal to 0, which is not restrictive up to the addition of a constant to the Hamiltonian.

We start by recalling some results proved in previous works. We underline that the compactness of M, which is assumed in these papers, does not actually play any role for the results we are about to state. The first one has been proved in [20].

Proposition 4.5. *Let H satisfy assumptions (H1), (H2) and (H3). For every $y \in M \setminus \mathcal{A}$ the set $Z_0(y)$ has nonempty interior and*

$$D^- S_y(y) = Z_0(y).$$

In particular, S_y is not differentiable at y.

Therefore, critical subsolutions are in general not differentiable outside the Aubry set. The situation is quite different on it. A fine result proved in [20] shows that, when H is locally Lipschitz–continuous in x and condition (H2)$'$ is assumed, all critical subsolutions are (strictly) differentiable at any point of the Aubry set, and have the same gradient. These results are based upon some semiconcavity estimates which, in turn, depend essentially on the Lipschitz character of the Hamiltonian in x. Something analogous still survives in the case of a purely continuous and convex Hamiltonian by looking at the behavior of the critical subsolutions on static curves, see [16].

Theorem 4.6. *Let H satisfy assumptions (H1), (H2) and (H3) and $\gamma \in \mathcal{K}$. Then there exists a negligible set $\Sigma \subset \mathbb{R}$ such that, for any critical subsolution u, the map $u \circ \gamma$ is differentiable on $\mathbb{R} \setminus \Sigma$ and satisfies*

$$\frac{\mathrm{d}}{\mathrm{d}t} (u \circ \gamma)(t_0) = \sigma(\gamma(t_0), \dot\gamma(t_0)) \qquad \textit{whenever } t_0 \in \mathbb{R} \setminus \Sigma. \tag{25}$$

Here we want to strengthen Theorem 4.6 by proving that, when condition (H2)$'$ is assumed, any critical subsolution is actually differentiable at $\mathcal{H}^1$–a.e. point of $\gamma(\mathbb{R})$. We give a definition first.

Definition 4.7. Let γ be an absolutely continuous curve defined on $\mathbb{R}$. We will denote by Σ_γ the negligible subset of $\mathbb{R}$ such that $\mathbb{R} \setminus \Sigma_\gamma$ is the following set:

$$\{t \in \mathbb{R} : t \text{ is a differentiability point of } \gamma \text{ and a Lebesgue point of } D_q L\big(\gamma(\cdot), \dot\gamma(\cdot)\big)\}.$$

Theorem 4.8. *Let H satisfy conditions (H1),(H2)$'$ and (H3). Then, for any $\gamma \in \mathcal{K}$, every critical subsolution u is differentiable at $\gamma(t_0)$ for any $t_0 \in \mathbb{R} \setminus \Sigma_\gamma$, and we have*

$$Du(\gamma(t_0)) = D_q L(\gamma(t_0), \dot\gamma(t_0)) \qquad \textit{for every } t_0 \in \mathbb{R} \setminus \Sigma_\gamma. \tag{26}$$

Proof. Fix $t_0 \in \mathbb{R} \setminus \Sigma$. As u is a critical subsolution, it is easily seen that

$$\big(\mathcal{S}(t_0)u\big)(x) \geqslant u(x) \qquad \text{for every } x \in M,$$

with equality holding at $\gamma(t_0)$ since

$$\big(\mathcal{S}(t_0)u\big)(\gamma(t_0)) \leqslant u(\gamma(0)) + \int_0^{t_0} L(\gamma, \dot\gamma)\, \mathrm{d}s = u(\gamma(t_0)).$$

By this and by Proposition 4.4 we obtain

$$D_q L(\gamma(t_0), \dot\gamma(t_0)) \in D^+\big(\mathcal{S}(t_0)u\big)(\gamma(t_0)) \subseteq D^+u(\gamma(t_0))$$

Analogously

$$\big(\mathcal{S}(t_0+1)u\big)(\gamma(t_0+1)) = u(\gamma(t_0)) + \int_{t_0}^{t_0+1} L(\gamma, \dot\gamma)\,\mathrm{d}s,$$

and by Proposition 4.4 we have

$$D_q L(\gamma(t_0), \dot\gamma(t_0)) \in D^- u(\gamma(t_0)).$$

Then u is differentiable at $\gamma(t_0)$ and $Du(\gamma(t_0)) = D_q L(\gamma(t_0), \dot\gamma(t_0))$, as it was to be shown. □

Let us denote by $\mathfrak{S}\mathfrak{S}$ the set of critical subsolutions for H, i.e. the subsolutions of equation (14). We define the set

$$\mathcal{D} := \bigcap_{v \in \mathfrak{S}\mathfrak{S}} \{ y \in M \,:\, v \text{ and } S_y \text{ are differentiable at } y,\ Dv(y) = DS_y(y) \}, \qquad (27)$$

where S_y stands for the function $S(y, \cdot)$. The following holds:

Proposition 4.9. *Let us assume (H1),(H2)′ and (H3). Then $\mathcal{D}$ is a dense subset of $\mathcal{A}$. When M is compact, we have in particular that $\mathcal{D}$ is a uniqueness set for the critical equation, i.e. if two critical solutions agree on $\mathcal{D}$, then they agree on the whole M.*

Proof. It is clear by Proposition 4.5 that $\mathcal{D}$ is contained in $\mathcal{A}$. Pick $y \in \mathcal{A}$ and choose a static curve $\gamma \in \mathcal{K}$ passing through y. According to Theorem 4.8, there exists a sequence of points $y_n \in \gamma(\mathbb{R}) \cap \mathcal{D}$ converging to y. This proves that $\mathcal{D}$ is dense in $\mathcal{A}$.

The fact that $\mathcal{D}$ is a uniqueness set is now a direct consequence of the fact that $\mathcal{A}$ is a uniqueness set, see [20]. □

Remark 4.10. We underline for later use that, by definition of $\mathcal{D}$, any two critical subsolutions u and v are differentiable on $\mathcal{D}$ and have same gradient.

5. Commuting Hamiltonians

In this Section we give the definition of commutation for continuous Hamiltonians and we explore its main properties. Afterward, we restrict to the case when M is compact and we study the critical equations associated with two commuting Hamiltonians. We discover in the end that the commutation property gives very strong information on the corresponding critical equations.

Throughout this section H and G will denote a pair of Hamiltonians satisfying assumptions (H1), (H2) and (H3). The following notations will be assumed

- L_H and L_G are the Lagrangians associated through the Fenchel transform with H and G, respectively.
- $\mathcal{S}_H$ and $\mathcal{S}_G$ denote the Lax–Oleinik semigroups associated with H and G, respectively.
- h_H^t and h_G^t will denote, for every $t > 0$, the functions associated via (8) with H and G, respectively.

- h_H and h_G are the Peierls barriers associated with H and G, respectively.

5.1. **Definition and preliminary facts.** We will say that two Hamiltonians H and G satisfying (H1), (H2), (H3) *commute* if

$$\mathcal{S}_G(s)\big(\mathcal{S}_H(t)\,u\big)(x) = \mathcal{S}_H(t)\big(\mathcal{S}_G(s)\,u\big)(x) \qquad \text{for every } s,\, t > 0 \text{ and } x \in M, \tag{28}$$

and for every function admissible initial datum $u : M \to \mathbb{R} \cup \{+\infty\}$.

Remark 5.1. It is not difficult to show that condition (28) is equivalent to saying that the multi–time Hamilton–Jacobi equation (1) admits a solution for every Lipschitz continuous initial datum. Also note that, when M is compact, any continuous function is an admissible initial datum.

Note that a Hamiltonian function H always commute with itself.

In the literature, the notion of commutativity appears differently, in terms of cancellation of the Poisson bracket. In [4], the following results were proved:

Theorem 5.2. *Assume G and H are two C^1 Hamiltonians on $\mathbb{R}^N \times \mathbb{R}^N$ which are convex in p and verify the following hypothesis:*

(1) for any $R > 0$ there exists a constant K_R such that

$$\max\big\{|G(x,p)|,\,|H(x,p)|\big\} < K_R, \quad \max\big\{|D_pG(x,p)|,|D_pH(x,p)|\big\} < K_R(1+|x|)$$

for every $(x,p) \in \mathbb{R}^N \times B_R$;

(2) $\min\big\{G(x,p),\, H(x,p)\big\} \to +\infty$ *as $|p| \to +\infty$, uniformly in x;*

(3) $\{G,H\} := \langle D_xG,\, D_pH\rangle - \langle D_xH,\, D_pG\rangle = 0$ *in $\mathbb{R}^N \times \mathbb{R}^N$ for every $(x,p) \in \mathbb{R}^N \times \mathbb{R}^N$.*

Then (28) *holds for any $u \in \mathrm{Lip}(\mathbb{R}^N)$.*

Theorem 5.3. *Assume G and H are locally Lipschitz continuous in $\mathbb{R}^N \times \mathbb{R}^N$ and verify hypotheses (1), (2) and (3) of Theorem 5.2 in the almost everywhere sense. Furthermore, assume one of the following two conditions:*

- *the Hamiltonian G is C^1;*
- *there exist two sequences of Hamiltonians $(H_n)_n$ and $(G_n)_n$ satisfying the hypotheses of Theorem 5.2 for any fixed n such that $H_n \rightrightarrows H$ and $G_n \rightrightarrows G$ as $n \to +\infty$.*

Then (28) *holds for any $u \in \mathrm{Lip}(\mathbb{R}^N)$.*

These results have been subsequently generalized in [27] in a control theoretical framework, and improved in [8] in the framework of variational solutions via symplectic arguments.

Remark 5.4. The definition of commutation given above via (28) is actually equivalent to the cancellation of the Poisson bracket when the Hamiltonians are additionally assumed of class C^1. The proof of this fact is sketched in the introduction of [4], and is detailed in Appendix B in the case of Tonelli Hamiltonians. This equivalence will be used to establish Theorem 5.15, see the proof of Lemma 5.16.

Let us proceed to explore the properties of commutation of Hamiltonians as defined above via (28).

Proposition 5.5. *The Hamiltonians H and G commute if and only if*

$$\min_{z\in M}\big(h^t_H(y,z)+h^s_G(z,x)\big)=\min_{z\in M}\big(h^s_G(y,z)+h^t_H(z,x)\big) \tag{29}$$

for every $x, y \in M$ and $t,\ s > 0$.

Remark 5.6. Formula (29) holds with minima even when M is non compact. Indeed,

$$\tau\,\alpha_*\left(\frac{|z-\zeta|}{\tau}\right)\leqslant h^\tau_H(z,\zeta)\leqslant \tau\,\beta_*\left(\frac{|z-\zeta|}{\tau}\right),$$

and the same is valid for $h^\tau_G(z,\zeta)$. Hence for every fixed $x,\ y \in M$ and $t, s > 0$ there exists a compact set K, only depending on α_* and β_* (equivalently, on α and β), such that

$$\inf_{z\in M}\big(h^t_H(y,z)+h^s_G(z,x)\big)=\min_{z\in K}\big(h^t_H(y,z)+h^s_G(z,x)\big)$$

and

$$\inf_{z\in M}\big(h^s_G(y,z)+h^t_H(z,x)\big)=\min_{z\in K}\big(h^s_G(y,z)+h^t_H(z,x)\big).$$

Proof. Let $u : M \to \mathbb{R}\cup\{+\infty\}$ be an admissible initial datum. Using the definitions and the commutation of two nested infima we get, for every $x \in M$ and $t, s > 0$

$$\begin{aligned}\mathcal{S}_G(s)\big(\mathcal{S}_H(t)\,u\big)(x) &= \inf_{z\in M}\inf_{\zeta\in M}\Big(h^t_H(\zeta,z)+h^s_G(z,x)+u(\zeta)\Big)\\ &= \inf_{\zeta\in M}\left(\inf_{z\in M}\ \big(h^t_H(\zeta,z)+h^s_G(z,x)\big)\ +u(\zeta)\right),\end{aligned} \tag{30}$$

$$\begin{aligned}\mathcal{S}_H(t)\big(\mathcal{S}_G(s)\,u\big)(x) &= \inf_{z\in M}\inf_{\zeta\in M}\Big(h^s_G(\zeta,z)+h^t_H(z,x)+u(\zeta)\Big)\\ &= \inf_{\zeta\in M}\left(\inf_{z\in M}\ \big(h^s_G(\zeta,z)+h^t_H(z,x)\big)\ +u(\zeta)\right).\end{aligned} \tag{31}$$

Now, if H and G commute, then (29) follows by plugging in the above equalities as u the function equal to 0 at y and $+\infty$ elsewhere, for every fixed $y \in M$. Conversely, if (29) holds true, then (30) and (31) are equal for any admissible u, so H and G commute. □

As a straightforward consequence we get the following results:

Corollary 5.7. *Let G and H be a pair of Hamiltonians satisfying (H1), (H2), (H3).*

(i) If H and G commute, then $\check{H}$ and $\check{G}$ commute.

(ii) Let $(H_n)_n$ and $(G_n)_n$ be two sequences of Hamiltonians satisfying (H1), (H2), (H3) such that $H_n \rightrightarrows H$ and $G_n \rightrightarrows G$ as $n \to +\infty$. If H_n commutes with G_n for any fixed n, then H commutes with G.

Proof. Assertion *(i)* immediately follows from Proposition 5.5 by simply noticing that $h^t_{\check{H}}(y,z) = h^t_H(z,y)$ and $h^s_{\check{G}}(z,x) = h^s_G(x,z)$ for every $x,\ y,\ z \in M$ and $t, s > 0$. For assertion *(ii)*, it suffices to show, in view of Remark 5.6, that

$$h^t_{H_n} \rightrightarrows h^t_H \quad \text{and} \quad h^s_{G_n} \rightrightarrows h^s_G \qquad \text{as } n \to +\infty.$$

For definiteness, let us show that $h^t_{H_n} \rightrightarrows h^t_H$. The fact that $H_n \rightrightarrows H$ implies $L_{H_n} \rightrightarrows L_H$ in $M \times \mathbb{R}^N$, see for instance Appendix A.2 in [7]. From Proposition 2.6 we know that the functions $h^t_{H_n}$ and h^t_H are locally equi–Lipschitz, so it suffices to show that the convergence holds pointwise. Let us fix $x,\, y \in M$. By the dominated convergence Theorem we easily get

$$\limsup_n h^t_{H_n}(x,y) \leqslant h^t_H(x,y).$$

To get the liminf inequality, we choose, for every $n \in \mathbb{N}$, a curve $\gamma_n : [-t, 0] \to M$ joining x to y such that

$$\int_{-t}^{0} L_{H_n}(\gamma_n, \dot\gamma_n)\,\mathrm{d}s = h^t_{H_n}(x,y).$$

By Proposition 2.5 we know that such curves are equi–Lipschitz, in particular the images of the maps $s \mapsto (\gamma_n(s),\, \dot\gamma_n(s))$ are all contained in a compact subset of $M \times \mathbb{R}^N$. Then

$$\liminf_{n\to+\infty} h^t_{H_n}(x,y) = \liminf_{n\to+\infty} \int_{-t}^{0} L_{H_n}(\gamma_n, \dot\gamma_n)\,\mathrm{d}s = \liminf_{n\to+\infty} \int_{-t}^{0} L_{H}(\gamma_n, \dot\gamma_n)\,\mathrm{d}s \geqslant h^t_H(x,y),$$

as it was to be shown. □

Proposition 5.8. *Assume H_1 and H_2 are two commuting Hamiltonians satisfying (H1), (H2), (H3), and set*

$$G(x,p) = \max\{H_1(x,p),\, H_2(x,p)\} \qquad \text{for every } (x,p) \in M \times \mathbb{R}^N.$$

Then G commutes both with H_1 and H_2.

We set $L(x,q) := \min\{L_{H_1}(x,q),\, L_{H_2}(x,q)\}$ for all $(x,q) \in M \times \mathbb{R}^N$. To ease notations, in the sequel we will write L_i, h^t_i in place of L_{H_i}, $h^t_{H_i}$. We recall that L^* denotes the Fenchel transform of L, defined according to (3).

We need two lemmas first.

Lemma 5.9. *For every $x,\, y \in M$ and $t > 0$*

$$h^t_G(x,y) = \inf\left\{ \int_0^t L(\gamma,\dot\gamma)\,\mathrm{d}s \,:\, \gamma \in C^1([0,t];M),\, \gamma(0)=x,\, \gamma(t)=y \right\}. \tag{32}$$

Proof. By classical results of Calculus of Variations, see for instance [6, 11], we know that the infimum appearing in (32) agrees with

$$\inf\left\{ \int_0^t L^{**}(\gamma,\dot\gamma)\,\mathrm{d}s \,:\, \gamma \in W^{1,1}([0,t];M),\, \gamma(0)=x,\, \gamma(t)=y \right\},$$

so to conclude we only need to prove that $L^{**} = L_G$. From the inequalities $G \geqslant H_i$ we derive $L_G \leqslant L_i$ for $i \in \{1,\, 2\}$, so $L_G \leqslant L$. By duality

$$G = L_G^* \geqslant L^* \geqslant L_i^* = H^i, \qquad i \in \{1,\, 2\},$$

so $G \geqslant L^* \geqslant \max\{H_1,\, H_2\} = G$. Hence $G = L^*$ and consequently $L_G = L^{**}$. □

Lemma 5.10. *For every $x,\, y \in M$ and $t > 0$*

$$h^t_G(x,y) = \inf\left\{ \sum_{i=1}^n h^{t_i}_{\sigma(i)}(x_{i-1},x_i) \,:\, x_0 = x,\, x_n = y,\, \sum_i t_i = t,\, \sigma \in \{1,\, 2\}^n,\, n \in \mathbb{N} \right\}.$$

Proof. The fact that the right–hand side term of the above equality is non smaller than $h^t_G(x,y)$ is an immediate consequence of the inequalities $L_i \geqslant L_G$ for $i \in \{1,\,2\}$. To prove the opposite inequality, in view of Lemma 5.9, it suffices to show that for every $\varepsilon > 0$ and for every curve $\gamma : [0,t] \to M$ of class C^1 joining x to y we have

$$\varepsilon + \int_0^t L(\gamma,\dot\gamma)\,\mathrm{d}s \geqslant \sum_{i=1}^n \int_{t_{i-1}}^{t_i} L_{\sigma(i)}(\gamma,\dot\gamma)\,\mathrm{d}s$$

for a suitable choice of $n \in \mathbb{N}$, $\{t_i\,:\,0 \leqslant i \leqslant n\}$ and $\sigma \in \{1,2\}^n$.

To this purpose, choose a sufficiently large positive number R such that $\|\dot\gamma\|_\infty < R$ and $\gamma([0,t]) \subseteq B_R$. Denote by ω a continuity modulus for L_1 and L_2 in $B_R \times B_R$. Let r be an arbitrarily chosen positive number and choose $n \in \mathbb{N}$ large enough in such a way that

$$|\gamma(s)-\gamma(\tau)| + |\dot\gamma(s)-\dot\gamma(\tau)| < r \qquad \text{for any } s,\, \tau \in [0,t] \text{ with } |s-\tau| < \tfrac{t}{n}.$$

Let $t_i := i\,t/n$ for $0 \leqslant i \leqslant n$ and define $\sigma \in \{1,2\}^n$ in such a way that

$$L_{\sigma(i)}(\gamma(t_i),\dot\gamma(t_i)) = L(\gamma(t_i),\dot\gamma(t_i)) \qquad \text{for every } 1 \leqslant i \leqslant n.$$

For every $s \in [t_{i-1},\,t_i]$ we get

$$\begin{aligned}\big|L_{\sigma(i)}(\gamma(s),\dot\gamma(s)) - L(\gamma(s),\dot\gamma(s))\big| &\leqslant \big|L_{\sigma(i)}(\gamma(s),\dot\gamma(s)) - L_{\sigma(i)}(\gamma(t_i),\dot\gamma(t_i))\big| \\ &\quad + |L(\gamma(t_i),\dot\gamma(t_i)) - L(\gamma(s),\dot\gamma(s))| \leqslant 2\,\omega(r).\end{aligned}$$

Then

$$\sum_{i=1}^n \int_{t_{i-1}}^{t_i} L_{\sigma(i)}(\gamma,\dot\gamma)\,\mathrm{d}s \leqslant \int_0^t L(\gamma,\dot\gamma)\,\mathrm{d}s + 2t\,\omega(r),$$

and the assertion follows by choosing r small enough. □

Proof of Proposition 5.8. Let us prove that G commutes with H_i, where i has been fixed, say $i = 1$ for definitiveness. In view of Proposition 5.5, we need to show that (29) holds with H_1 and G in place of H and G, respectively. Let us fix $t, s > 0$ and $x,\, y \in M$. Let us show that

$$\min_{z\in M}\big(h^t_G(x,z) + h^s_{H_1}(z,y)\big) \geqslant \min_{z\in M}\big(h^s_{H_1}(x,z) + h^t_G(z,y)\big). \tag{33}$$

In view of Lemma 5.10, it will be enough to prove that, for every $n \in \mathbb{N}$, the following inequality holds:

$$\sum_{i=1}^n h^{t_i}_{\sigma(i)}(x_{i-1},x_i) + h^s_{H_1}(x_n,y) \geqslant \min_{z\in M}\big(h^s_{H_1}(x,z) + h^t_G(z,y)\big) \tag{34}$$

for every $\sigma \in \{1,2\}^n$, $\{x_i\,:\,0 \leqslant i \leqslant n\}$ with $x_0 = x$, $\{t_i\,:\,0 \leqslant i \leqslant n\}$ with $\sum_i t_i = t$.

The proof will be by induction on n.

For $n = 1$ inequality (34) holds true for

$$h^t_{\sigma(1)}(x,x_1) + h^s_{H_1}(x_1,y) \geqslant \min_{z\in M}\big(h^t_{\sigma(1)}(x,z) + h^s_{H_1}(z,y)\big),$$

and we conclude since $H_{\sigma(1)}$ and H_1 commute and $h^t_{\sigma(1)} \geqslant h^t_G$, for every $\sigma(1) \in \{1,\,2\}$.

Let us now assume that (34) holds for n and let us show it holds for $n+1$. Let

$\sigma \in \{1,2\}^{n+1}$, $\{x_i : 0 \leqslant i \leqslant n+1\}$ with $x_0 = x$, $\{t_i : 0 \leqslant i \leqslant n+1\}$ with $\sum_i t_i = t$. We have

$$
\begin{aligned}
&\sum_{i=1}^{n} h^{t_i}_{\sigma(i)}(x_{i-1}, x_i) + h^{t_{n+1}}_{\sigma(n+1)}(x_{n+1}, y) + h^{s}_{H_1}(x_n, y) \\
&\geqslant \sum_{i=1}^{n} h^{t_i}_{\sigma(i)}(x_{i-1}, x_i) + \min_{z \in M} \Big(h^{t_{n+1}}_{\sigma(n+1)}(x_n, z) + h^{s}_{H_1}(z, y) \Big) \\
&= \sum_{i=1}^{n} h^{t_i}_{\sigma(i)}(x_{i-1}, x_i) + \min_{z \in M} \Big(h^{s}_{H_1}(x_n, z) + h^{t_{n+1}}_{\sigma(n+1)}(z, y) \Big),
\end{aligned}
$$

where we used the fact that H_1 and $H_{\sigma(n+1)}$ commute. Let us denote by $\overline{z}$ a point realizing the minimum in the last row of the above expression. By making use of the inductive hypothesis we get

$$
\begin{aligned}
&\sum_{i=1}^{n} h^{t_i}_{\sigma(i)}(x_{i-1}, x_i) + h^{s}_{H_1}(x_n, \overline{z}) + h^{t_{n+1}}_{\sigma(n+1)}(\overline{z}, y) \\
&\geqslant \min_{\zeta \in M} \Big(h^{s}_{H_1}(x, \zeta) + h^{t - t_{n+1}}_{G}(\zeta, \overline{z}) \Big) + h^{t_{n+1}}_{\sigma(n+1)}(\overline{z}, y) \\
&= \min_{\zeta \in M} \Big(h^{s}_{H_1}(x, \zeta) + h^{t - t_{n+1}}_{G}(\zeta, \overline{z}) + h^{t_{n+1}}_{\sigma(n+1)}(\overline{z}, y) \Big) \\
&\geqslant \min_{\zeta \in M} \Big(h^{s}_{H_1}(x, \zeta) + h^{t}_{G}(\zeta, y) \Big).
\end{aligned}
$$

The proof of the opposite inequality in (33) is analogous. The proof is complete. □

5.2. **Critical equations.** In this subsection we will specialize to the case $M = \mathbb{T}^N$. We will denote by c_H and c_G the corresponding critical values of H and G, respectively. Up to adding a constant to the Hamiltonians, we will assume that $c_H = c_G = 0$. Note that this does not affect the commutation property. The symbols S_H, S_G and $\mathcal{A}_H$, $\mathcal{A}_G$ refer to the critical semidistance and the Aubry set associated with H and G, respectively.

We will also denote by $\mathfrak{S}\mathfrak{S}_H$ and $\mathfrak{S}_H$ the set of subsolutions and solutions of the critical equations $H = 0$, respectively, and by $\mathfrak{S}\mathfrak{S}_G$ and $\mathfrak{S}_G$ the analogous objects for the critical equation $G = 0$.

We start with two results which exploit the fact that H and G commute. Actually, the first result is a direct consequence of the monotonicity of the semigroups and does not require M to be compact. The second one uses the fact that the Lax–Oleinik semigroups are weakly contracting for the infinity norm and the proof is done applying DeMarr's theorem on existence of common fixed points for commuting weakly contracting maps of Banach spaces [17]. The compactness of M is crucial to assure that such common fixed points are critical solutions for both the Hamiltonians.

The proofs of these results may be found in [30] and will be omitted.

Proposition 5.11. *Let H and G be a pair of commuting Hamiltonians satisfying assumptions (H1), (H2), (H3). Then, for every $t > 0$, we have*

$$
\begin{aligned}
&\mathcal{S}_H(t)u \in \mathfrak{S}\mathfrak{S}_G && \text{for every } u \in \mathfrak{S}\mathfrak{S}_G, \\
&\mathcal{S}_H(t)u \in \mathfrak{S}_G && \text{for every } u \in \mathfrak{S}_G.
\end{aligned}
$$

Proposition 5.12. *Let H and G be a pair of commuting Hamiltonians satisfying assumptions (H1), (H2), (H3). Then there exists $u_0 \in \mathfrak{S}_H \cap \mathfrak{S}_G$. In particular,*

$$H(x, Du_0(x)) = G(x, Du_0(x)) = 0$$

at any differentiability point x of u_0.

We now assume strict convexity of the Hamiltonians and we exploit the differentiability properties of critical subsolutions established in Section 4 to prove the following

Theorem 5.13. *Let H and G satisfy assumptions (H1), (H2)′, (H3). If H and G commute, then $\mathfrak{S}_H = \mathfrak{S}_G$.*

Proof. It suffices to show that $\mathfrak{S}_G \subseteq \mathfrak{S}_H$, since the opposite inclusion follows by interchanging the roles of H and G.

Take $u \in \mathfrak{S}_G$. To prove that $u \in \mathfrak{S}_H$, it suffices to show, in view of Proposition 2.10–*(ii)*, that

$$\mathcal{S}_H(t)u = u \quad \text{on } \mathbb{T}^N \qquad \text{for every } t > 0.$$

Since $\mathcal{S}_H(t)u \in \mathfrak{S}_G$, according to Proposition 4.9 it suffices to prove that

$$\mathcal{S}_H(t)u = u \quad \text{on } \mathcal{D}_G \qquad \text{for every } t > 0.$$

Let $u_0 \in \mathfrak{S}\mathfrak{S}_H \cap \mathfrak{S}\mathfrak{S}_G$, and pick a point $y \in \mathcal{D}_G$. By definition of $\mathcal{D}_G$, the function u_0 is differentiable at y. Moreover, see Remark 4.10, for every $v \in \mathfrak{S}\mathfrak{S}_G$

$$v \text{ is differentiable at } y \qquad \text{and} \qquad Dv(y) = Du_0(y).$$

Then the function $w(t, x) := \big(\mathcal{S}_H(t)u\big)(x)$ is differentiable at y for every $t > 0$ and

$$H(y, D_x w(t, y)) = H(y, Du_0(y)) = 0 \qquad \text{for every } t > 0 \tag{35}$$

by Proposition 5.12. Moreover w is a solution of the evolutive equation

$$\frac{\partial w}{\partial t} + H(x, D_x w) = 0 \qquad \text{in } (0, +\infty) \times \mathbb{T}^N.$$

Since the map $t \mapsto w(t, y)$ is Lipschitz continuous, it is differentiable for a.e. $t > 0$ and we get, taking also into account (35),

$$\frac{\partial w}{\partial t}(t, y) = \frac{\partial w}{\partial t}(t, y) + H(y, D_x w(t, y)) = 0 \qquad \text{for a.e. } t > 0,$$

yielding that $w(\cdot, y)$ is constant. Hence

$$\big(\mathcal{S}_H(t)u\big)(y) = \lim_{\tau \to 0^+} \big(\mathcal{S}_H(\tau)u\big)(y) = u(y) \qquad \text{for every } t > 0,$$

as it was to be shown. □

Theorem 5.13 has very strong consequences from the weak KAM theoretic viewpoint. Indeed, we have

Theorem 5.14. *Let H and G be a pair of commuting Hamiltonians satisfying assumptions (H1), (H2)′, (H3). Then*

(i) $h_H = h_G$ *on* $\mathbb{T}^N \times \mathbb{T}^N$*;*

(ii) $\mathcal{A}_H = \mathcal{A}_G$*;*

(iii) $S_H(x, y) = S_G(x, y)$ *if either x or y belong to* $\mathcal{A}_H = \mathcal{A}_G$.

Proof. *(i)* Let us arbitrarily fix $y \in \mathbb{T}^N$. By Proposition 3.6, $h_H(y,\cdot)$ and $h_G(y,\cdot)$ both belong to $\mathfrak{S}_H = \mathfrak{S}_G$, so

$$\Big(\mathcal{S}_G(s)\, h_H(y,\cdot)\Big) = h_H(y,\cdot), \qquad\qquad \Big(\mathcal{S}_H(t)\, h_G(y,\cdot)\Big) = h_G(y,\cdot)$$

for every $s, t > 0$. Moreover

$$h_H^t \underset{t\to+\infty}{\rightrightarrows} h_H \qquad \text{and} \qquad h_G^s \underset{s\to+\infty}{\rightrightarrows} h_G \qquad\qquad \text{in } \mathbb{T}^N \times \mathbb{T}^N.$$

Let us denote by u the function equal to 0 at y and $+\infty$ elsewhere. For every $s > 0$ we have

$$\begin{aligned} h_H(y,\cdot) &= \Big(\mathcal{S}_G(s)\, h_H(y,\cdot)\Big) = \lim_{t\to+\infty} \Big(\mathcal{S}_G(s)\, h_H^t(y,\cdot)\Big) \\ &= \lim_{t\to+\infty} \Big(\mathcal{S}_G(s)\, \mathcal{S}_H(t) u\Big) = \lim_{t\to+\infty} \Big(\mathcal{S}_H(t)\, \mathcal{S}_G(s) u\Big) = \lim_{t\to+\infty} \Big(\mathcal{S}_H(t)\, h_G^s(y,\cdot)\Big). \end{aligned}$$

We derive

$$\begin{aligned} \|h_H(y,\cdot) - h_G(y,\cdot)\|_\infty = \lim_{t\to+\infty} \big\|\mathcal{S}_H(t)\big(\, h_G^s(y,\cdot)\big) - \mathcal{S}_H(t)\big(h_G(y,\cdot)\big)\big\|_\infty \\ \leqslant \|h_G^s(y,\cdot) - h_G(y,\cdot)\|_\infty, \end{aligned}$$

and the assertion follows sending $s \to +\infty$.

Assertions *(ii)* and *(iii)* are a direct consequence of *(i)* in view of Theorem 3.4 and of Proposition 3.6, respectively. □

Next, we show that H and G admit a common strict subsolution.

Theorem 5.15. *Let H and G be a pair of commuting Hamiltonians satisfying (H1), (H2)′, (H3), and let $\mathcal{A}$ denote $\mathcal{A}_H = \mathcal{A}_G$. Then there exists $v \in \mathfrak{S}\mathfrak{S}_H \cap \mathfrak{S}\mathfrak{S}_G$ which is C^∞ and strict in $\mathbb{T}^N \setminus \mathcal{A}$ both for H and for G, i.e.*

$$H(x, Dv(x)) < 0 \quad \textit{and} \quad G(x, Dv(x)) < 0 \qquad \textit{for every } x \in \mathbb{T}^N \setminus \mathcal{A}. \tag{36}$$

If H and G are locally Lipschitz continuous in $\mathbb{T}^N \times \mathbb{R}^N$, then v can be additionally chosen in $C^1(\mathbb{T}^N)$.

Finally, if H and G are Tonelli, then v can be chosen in $C^{1,1}(\mathbb{T}^N)$.

Proof. Let us set

$$F(x,p) := \max\{H(x,p),\, G(x,p)\} \qquad \text{for every } (x,p) \in \mathbb{T}^N \times \mathbb{R}^N.$$

This new Hamiltonian still satisfies (H1), (H2)′ and (H3). Moreover any $u \in \mathfrak{S}_H = \mathfrak{S}_G$ solves the equation

$$F(x, Du) = 0 \qquad \text{in } \mathbb{T}^N$$

in the viscosity sense, as it is easily seen by definition of F. This yields $c_F = 0$ and, according to Proposition 5.8 and Theorem 5.14, $\mathcal{A}_F = \mathcal{A}$.

We now invoke the results proved [20]: by Theorem 6.2, there exists a critical subsolution v for F which is strict and smooth in $\mathbb{T}^N \setminus \mathcal{A}$. If H and G are locally Lipschitz, the same holds for F, so v can be additionally chosen of class C^1 on the whole $\mathbb{T}^N$ in view of Theorem 8.1. The inequalities (36) follow since $F \geqslant H, G$.

If now H and G are Tonelli Hamiltonians, the commutation property is equivalent to the fact that the Poisson bracket $\{H, G\} = 0$ everywhere, as explained in Appendix B. Starting with a C^1 (or in fact any) common strict subsolution v, it is possible to realize, as in [5], a Lasry–Lions regularization v_0 of v, using alternatively the positive and negative semigroups of H. More precisely,

$$v_0 = \mathcal{S}_H(t)\big(\mathcal{S}_H^+(s)\,v\big)$$

for s and t suitably chosen, where the positive Lax–Oleinik semigroup is defined as follows:

$$\mathcal{S}_H^+(s)v = -\big(\mathcal{S}_{\check{H}}(-v)\big).$$

Remark 2.2, Corollary 5.7 and Proposition 5.11 yield that v_0 remains a subsolution, both of G and of H. The fact that it is strict in $\mathbb{T}^N \setminus \mathcal{A}$ is proved in the next lemma. The fact that v_0 is $C^{1,1}$ for t and s small enough is proved in [5]. □

We recall that a Lipschitz subsolution $v \in \mathfrak{S}\mathfrak{S}_G$ is said to be *strict* in an open set $U \subset \mathbb{T}^N$ if for any $x_0 \in U$ there is a neighborhood V of x_0 and a constant $\varepsilon > 0$ such that $G(x, Dv(x)) < -\varepsilon$ almost everywhere in V.

Note that if v is C^1, it is strict on U if and only if $G(x, Dv(x)) < 0$ for any $x \in U$.

Lemma 5.16. *Let G and H be two commuting Tonelli Hamiltonians. Assume v is a critical subsolution for G which is strict outside $\mathcal{A}$. Then, for all $t > 0$, both $\mathcal{S}_G(t)\,v$ and $\mathcal{S}_H(t)\,v$ are critical subsolution for G, strict outside $\mathcal{A}$.*

Proof. We will only prove the result for $\mathcal{S}_H(t)\,v$. The result for $\mathcal{S}_G(t)\,v$ is then a consequence for $G = H$. We already know by Proposition 5.11 that $\mathcal{S}_H(t)\,v$ is a critical subsolution of G. It is only left to prove the strict part. This is done in two steps: in a first one, we prove a point wise strictness at differentiability points of $\mathcal{S}_H(t)\,v$. In a second one, we extend this result using Clarke's gradient to any point before concluding.

Let $x \in \mathbb{T}^N \setminus \mathcal{A}$. Consider a curve γ verifying that $\gamma(0) = x$ and

$$\big(\mathcal{S}_H(t)\,v\big)(x) = v(\gamma(-t)) + \int_{-t}^{0} L_H(\gamma(s), \dot\gamma(s))\mathrm{d}s.$$

The curve $(\gamma, \dot\gamma)$ is then a piece of trajectory of the Euler–Lagrange flow of H. It is also known (see [18] or Proposition 4.4) that $D_q L_H(\gamma(-t), \dot\gamma(-t)) \in D^- v(\gamma(-t))$ and

$$D_q L_H(\gamma(s), \dot\gamma(s)) \in D^+\big(\mathcal{S}_H(t+s)\,v\big)(\gamma(s)) \qquad \text{for every } s \in (-t, 0].$$

Moreover, the curve γ does not intersect $\mathcal{A}$. Indeed, if this were not the case, the curve $(\gamma, \dot\gamma)$ would be included in the lifted Aubry set, which is invariant by the Euler–Lagrange flow of H, see [18], while $x = \gamma(0) \notin \mathcal{A}$. We therefore deduce that $\gamma(-t) \notin \mathcal{A}$ and, since v is strict,

$$G\big(\gamma(-t), D_q L_H(\gamma(-t), \dot\gamma(-t))\big) < 0.$$

Now G and H commute; since they are Tonelli, this means their Poisson bracket is null, see Proposition B.1. Otherwise stated, G is constant on the integral curves of the Hamiltonian flow of H, in particular on $s \mapsto (\gamma(s), D_q L_H(\gamma(s), \dot\gamma(s))$. Thus

$$G(x, D_q L_H(x, \dot\gamma(0))) < 0,$$

from which we infer that $G\left(x, D\big(\mathcal{S}_H(t)\,v\big)(x)\right) < 0$ whenever $\mathcal{S}_H(t)\,v$ is differentiable at x. But this is not sufficient to conclude since the function $\mathcal{S}_H(t)\,v$ is Lipschitz continuous in $\mathbb{T}^N$, hence differentiable almost everywhere only. We will prove the following:

Claim. Let $x \notin \mathcal{A}$. Then

$$G(x,p) < 0 \qquad \text{for every } p \in \partial^* \left(\mathcal{S}_H(t)\, v\right)(x),$$

where $\partial^* \left(\mathcal{S}_H(t)\, v\right)(x)$ denotes the set of *reachable gradients* of $\mathcal{S}_H(t)\, v$ at x, defined as

$$\partial^* \left(\mathcal{S}_H(t)\, v\right)(x) = \left\{ \lim_{x_n \to x} D\left(\mathcal{S}_H(t)\, v\right)(x_n) \,:\, \mathcal{S}_H(t)\, v \text{ differentiable at } x_n \right\}.$$

Note that since $\mathcal{S}_H(t)\, v$ is Lipschitz, the set of reachable gradients, $\partial^* \left(\mathcal{S}_H(t)\, v\right)(x)$, is compact.

Let $p \in \partial^* \left(\mathcal{S}_H(t)\, v\right)(x)$ and consider $x_n \to x$ a sequence of differentiability points for $\mathcal{S}_H(t)\, v$ such that $D(\mathcal{S}_H(t)\, v)(x_n) \to p$. For each n, choose a curve $\gamma_n : [-t, 0] \to \mathbb{T}^N$ (which is in fact unique) such that

$$\big(\mathcal{S}_H(t)\, v\big)(x_n) = v(\gamma_n(-t)) + \int_{-t}^{0} L_H(\gamma_n(s), \dot\gamma_n(s))\, \mathrm{d}s.$$

For each n, the curve $(\gamma_n, \dot\gamma_n)$ is the (only) trajectory of the Euler–Lagrange flow with initial condition verifying $D_q L_H(x_n, \dot\gamma_n(0)) = D(\mathcal{S}_H(t)\, v)(x_n)$. By continuity of this flow, they uniformly converge, along with their derivatives, to a curve γ. By continuity, we obtain

$$\big(\mathcal{S}_H(t)\, v\big)(x) = v(\gamma(-t)) + \int_{-t}^{0} L_H(\gamma(s), \dot\gamma(s))\mathrm{d}s.$$

Moreover, by passing to the limit in the equalities $D_q L_H(x_n, \dot\gamma_n(0)) = D(\mathcal{S}_H(t)\, v)(x_n)$, we obtain

$$D_q L_H(x, \dot\gamma(0)) = p.$$

By arguing as above and by exploiting the fact that $x \notin \mathcal{A}$, we obtain $G(x,p) < 0$. Since G is convex, we infer

$$G(x,p) < 0 \qquad \text{for every } p \in \partial(\mathcal{S}_H(t)\, v)(x),$$

where $\partial\big(\mathcal{S}_H(t)\, v\big)(x)$ denotes the *Clarke differential* of $\mathcal{S}_H(t)\, v$ at x, defined as the convex hull of $\partial^*\big(\mathcal{S}_H(t)\, v\big)(x)$. We now exploit the fact that the Clarke differential is upper semi–continuous with respect to the inclusion and point wise compact, see [12]. Let $x_0 \notin \mathcal{A}$ and choose $\varepsilon > 0$ in such a way that

$$G(x_0,p) < -2\varepsilon \qquad \text{for every } p \in \partial\big(\mathcal{S}_H(t)\, v\big)(x_0).$$

Then there exists a neighborhood V of x_0 such that

$$G(x,p) < -\varepsilon \qquad \text{for every } p \in \partial\big(\mathcal{S}_H(t)\, v\big)(x) \text{ and } x \in V.$$

In particular, $G(x, Dv(x)) < -\varepsilon$ for almost every $x \in V$. The proof is complete. □

Appendix A

The purpose of this Section is to give a self–contained proof of Theorem 3.3. We prove two lemmas first. Recall that we are assuming that the critical value c is equal to 0.

Lemma A.1. *Let* $\gamma : [a,b] \to M$ *such that*

$$S(\gamma(b), \gamma(a)) + \int_a^b L(\gamma, \dot\gamma)\,\mathrm{d}s = 0. \tag{37}$$

Then γ *is a static curve.*

Proof. Let s, t be points of $[a,b]$ with $s < t$. We want to prove that

$$-S(\gamma(t), \gamma(s)) = \int_s^t L(\gamma, \dot\gamma)\,\mathrm{d}\tau = S(\gamma(s), \gamma(t)). \tag{38}$$

We set $y := \gamma(b)$ and observe that equality (37) can be equivalently written as

$$S(y, \gamma(b)) - S(y, \gamma(a)) = \int_a^b L(\gamma, \dot\gamma)\,\mathrm{d}s.$$

Since $S(y, \cdot)$ is a critical subsolution, the following hold:

$$S(y, \gamma(b)) - S(y, \gamma(t)) \leqslant \int_t^b L(\gamma, \dot\gamma)\,\mathrm{d}s.$$

and

$$S(y, \gamma(t)) - S(y, \gamma(a)) \leqslant \int_a^t L(\gamma, \dot\gamma)\,\mathrm{d}s.$$

Both inequalities are in fact equalities (summing them up gives an equality) and we obtain

$$-S(y, \gamma(t)) = S(y, \gamma(b)) - S(y, \gamma(t)) = \int_t^b L(\gamma, \dot\gamma)\,\mathrm{d}s$$

for any $t \in [a,b]$. We infer

$$0 = S(y, \gamma(t)) + \int_t^b L(\gamma, \dot\gamma)\,\mathrm{d}\tau \geqslant S(y, \gamma(t)) + S(\gamma(t), y) \geqslant 0,$$

so

$$S(\gamma(t), y) = \int_t^b L(\gamma, \dot\gamma)\,\mathrm{d}\tau = -S(y, \gamma(t)).$$

In particular for every $a \leqslant s < t \leqslant b$

$$S(\gamma(s), y) - S(\gamma(t), y) = \int_s^t L(\gamma, \dot\gamma)\,\mathrm{d}\tau.$$

The second equality in (38) then follows since

$$S(\gamma(s), y) - S(\gamma(t), y) \leqslant S(\gamma(s), \gamma(t)) \leqslant \int_s^t L(\gamma, \dot\gamma)\,\mathrm{d}\tau.$$

Let us now prove the other equality in (38). By making use of what was just proved, we have

$$\begin{aligned} \int_s^t L(\gamma, \dot\gamma)\,\mathrm{d}\tau &= S(\gamma(s), y) - S(\gamma(t), y) \\ &= -\Big(S(y, \gamma(s)) + S(\gamma(t), y)\Big) \leqslant -S(\gamma(t), \gamma(s)), \end{aligned}$$

and the assertion follows for

$$\int_s^t L(\gamma, \dot\gamma)\,\mathrm{d}\tau + S(\gamma(t), \gamma(s)) \geqslant S(\gamma(s), \gamma(t)) + S(\gamma(t), \gamma(s)) \geqslant 0.$$

□

Lemma A.2. *There exists a real number $R > 0$ such that*

$$\bigcup_{x\in M} \{q \in \mathbb{R}^N \,:\, L(x,q) = \sigma(x,q)\} \subseteq B_R.$$

Proof. By assumption (H3) there exists a constant κ such that $Z_0(x) \subseteq B_\kappa$ for every $x \in M$, so $\sigma(x,q) \leqslant \kappa|q|$ for every $(x,q) \in M \times \mathbb{R}^N$. By (L3) and by the superlinear and continuous character of α_*, see Remark 2.1, there exists a constant $\alpha_0 > 0$ such that

$$(\kappa+1)|q| - \alpha_0 \leqslant \alpha_*(|q|) \leqslant L(x,q) \qquad \text{for every } (x,q) \in M \times \mathbb{R}^N.$$

The assertion follows by choosing $R := \alpha_0$. □

Proof of Theorem 3.3. Fix $y \in \mathcal{A}$ and set $u(\cdot) = S(y,\cdot)$. The function $w(x,t) = u(x)$ is a solution of the equation

$$\partial_t w(x,t) + H(x, D_x w(x,t)) = 0, \tag{39}$$

hence $\mathcal{S}(t)u = u$ for every $t > 0$. In particular, for each $n \in \mathbb{N}$ there exists a curve $\gamma_n : [-n,0] \to M$ with $\gamma_n(0) = y$ such that

$$u(y) = u(\gamma_n(-n)) + \int_{-n}^{0} L(\gamma_n, \dot\gamma_n)\,\mathrm{d}s.$$

Now $u(y) = 0$ and $u(\gamma_n(-n)) = S(y, \gamma_n(-n))$, so by Lemma A.1 we derive that γ_n is a static curve. Lemma A.2 guarantees that the curves γ_n are equi–Lipschitz continuous, in particular there exists a Lipschtiz curve $\gamma : \mathbb{R}_- \to M$ such that, up to subsequences,

$$\gamma_n \rightrightarrows \gamma \quad \text{in } \mathbb{R}_- \qquad \text{and} \qquad \dot\gamma_n \rightharpoonup \dot\gamma \quad \text{in} \quad L^1_{loc}(\mathbb{R}_-; \mathbb{R}^N).$$

By a classical semi–continuity result of the Calculus of Variations [6] we have

$$\liminf_{n\to+\infty} \int_a^b L(\gamma_n, \dot\gamma_n)\,\mathrm{d}s \geqslant \int_a^b L(\gamma, \dot\gamma)\,\mathrm{d}s$$

for every $a < b \leqslant 0$, yielding in particular that γ is static too.

We now consider the Hamiltonian $\check H(x,p) = H(x,-p)$. By Proposition 3.5, we know that the critical value and the Aubry set of $\check H$ agree with 0 (i.e. the critical value of H) and $\mathcal{A}$. We can apply the previous argument with $\check S$ and $\check L$ in place of S and L to obtain a curve $\xi : \mathbb{R}_- \to M$ which is static for $\check H$. We define a curve $\eta : \mathbb{R} \to M$ by setting

$$\eta(s) := \begin{cases} \xi(-s) & \text{if } s \geqslant 0 \\ \gamma(s) & \text{if } s \leqslant 0. \end{cases}$$

We claim that η is the static curve we were looking for. To prove this, it will be enough, in view of Lemma A.1, to show

$$S(\eta(b), \eta(a)) + \int_a^b L(\eta, \dot\eta)\,\mathrm{d}s = 0 \tag{40}$$

for any fixed $a < 0 < b$. Indeed, by noticing that $\check L(x,q) = L(x,-q)$ and $\check S(x,y) = S(y,x)$, we obtain

$$\int_0^b L(\eta, \dot\eta)\,\mathrm{d}s = \int_{-b}^0 \check L(\xi, \dot\xi)\,\mathrm{d}s = -\check S(\xi(0), \xi(-b))) = -S(\eta(b), \eta(0)).$$

Hence

$$\begin{aligned}\int_a^b L(\eta,\dot\eta)\,\mathrm{d}s &= \int_a^0 L(\eta,\dot\eta)\,\mathrm{d}s + \int_0^b L(\eta,\dot\eta)\,\mathrm{d}s \\ &= -\big(S(\eta(b),\eta(0)) + S(\eta(0),\eta(a)\big) \leqslant -S(\eta(b),\eta(a))\end{aligned}$$

and (40) follows since the opposite inequality is always true. □

Appendix B

In this Appendix, we justify our definition of commutation by proving that it reduces to the classical one, given in terms of cancellation of the Poisson bracket, when the Hamiltonians are regular enough. In [4], [30], it is proved that for two convex C^1–Hamiltonians, G and H, having null Poisson bracket, i.e.

$$\{G,\, H\} := \langle D_pG,\, D_xH\rangle - \langle D_pH,\, D_xG\rangle = 0 \qquad \text{in } M\times\mathbb{R}^N,$$

the multi–time Hamilton–Jacobi equation (1) admits a (unique) viscosity solution for any Lipschitz initial datum u_0. This amounts to saying that the Lax–Oleinik semigroups commute in the sense of (28). In [4], the question of the reciprocal statement is treated by a heuristic argument. We feel natural to give a neat proof of this fact, at least in the case of Tonelli Hamiltonians. For clarity of the exposition, we will place ourself in the case of $M=\mathbb{T}^N$, but the results remain true if $M=\mathbb{R}^N$.

Proposition B.1. *Let G and H be two Tonelli Hamiltonians on $\mathbb{T}^N\times\mathbb{R}^N$. Assume that*

$$\mathcal{S}_G(s)\big(\mathcal{S}_H(t)\,u\big)(x) = \mathcal{S}_H(t)\big(\mathcal{S}_G(s)\,u\big)(x) \qquad \textit{for every } s,\, t>0 \textit{ and } x\in\mathbb{T}^N, \tag{41}$$

and for every admissible initial datum $u:\mathbb{T}^N\to\mathbb{R}\cup\{+\infty\}$. Then the following relation is identically verified:

$$\langle D_pG,\, D_xH\rangle - \langle D_pH,\, D_xG\rangle = 0 \qquad \textit{in } \mathbb{T}^N\times\mathbb{R}^N.$$

In order to prove this, we will use some results about the behavior of solutions of the Hamilton–Jacobi equation with smooth initial datum. We introduce some notations. If $f:\mathbb{T}^N\to\mathbb{R}$ is differentiable, then $\Gamma(f)\subset\mathbb{T}^N\times\mathbb{R}^N$ will denote the graph of its differential. We will denote by ϕ_G (resp. ϕ_H) the Hamiltonian flow of G (resp. H), that is, the flow generated by the vectorfield

$$X_G(x,p) = \Big(x,p,\, D_pG(x,p),\, -D_xG(x,p)\Big), \qquad (x,p)\in\mathbb{T}^N\times\mathbb{R}^N,$$

(resp. $X_H(x,p) = \big(x,p,\, D_pH(x,p),\, -D_xH(x,p)\big)$).

The following is a reformulation of Lemma 3 in [5]:

Proposition B.2. *For any C^2 function $u:\mathbb{T}^N\to\mathbb{R}$, there is an $\varepsilon>0$ such that for any $s,t<\varepsilon$, the functions $u_{s,t}=\mathcal{S}_G(s)\big(\mathcal{S}_H(t)\,u\big)$ and $u^{s,t}=\mathcal{S}_H(t)\big(\mathcal{S}_G(s)\,u\big)$ are C^2. Moreover, $\Gamma(u_{s,t})=\phi_G^s\circ\phi_H^t\,\Gamma(u)$ and $\Gamma(u^{s,t})=\phi_H^t\circ\phi_G^s\,\Gamma(u)$.*

Finally, for $t<\varepsilon$ fixed (resp. $s<\varepsilon$ fixed), the function $(s,x)\mapsto u_{s,t}(x)$ (resp. $(t,x)\mapsto u^{s,t}(x)$) is a classical solution to the Hamilton Jacobi equation

$$\frac{\partial u_{s,t}}{\partial s} + G(x, D_xu_{s,t}) = 0 \qquad \textit{in } (0,+\infty)\times\mathbb{T}^N,$$

(resp. $\dfrac{\partial u^{s,t}}{\partial t} + H(x,D_xu^{s,t}) = 0 \quad \textit{in } (0,+\infty)\times\mathbb{T}^N$*).*

Proof of Proposition B.1. Let us use Proposition B.2, differentiating various times the Hamilton-Jacobi equation, to compute a Taylor expansion of $u_{s,t}$ for small times, smooth initial datum:

$$u_{s,t}(x) = u_{0,t}(x) - sG(x, D_x u_{0,t}(x)) - \frac{s^2}{2}\left\langle D_pG(x, D_x u_{0,t}(x)),\ \frac{\partial}{\partial s} D_x u_{0,t}(x)\right\rangle + o(s^2)$$

$$= u_{0,t}(x) - sG(x, D_x u_{0,t}(x)) + \frac{s^2}{2}\left\langle D_pG(x, D_x u_{0,t}(x)),\ D_xG(x, D_x u_{0,t}(x))\right\rangle + o(s^2).$$

Notice now that similarly,

$$u_{0,t}(x) = u(x) - tH(x, Du(x)) + \frac{t^2}{2}\left\langle D_pH(x, Du(x)),\ D_xH(x, Du(x))\right\rangle + o(t^2)$$

and

$$D_x u_{0,t}(x) = Du(x) - t\left[(D_xH(x, Du(x)) + D^2u(x)\, D_pH(x, Du(x))\right] + o(t).$$

By substitution, we obtain the following identity on $\mathbb{T}^N$:

$$\begin{aligned} u_{t,t} = u - tH(x, Du) + \frac{t^2}{2}\left\langle D_pH(x, Du),\ D_xH(x, Du)\right\rangle \\ - t\Big(G(x, Du) - t\left\langle D_pG(x, Du),\ D_xH(x, Du) + D^2\, u\, D_pH(x, Du)\right\rangle\Big) \\ + \frac{t^2}{2}\left\langle D_pG(x, Du),\ D_xG(x, Du)\right\rangle + o(t^2), \end{aligned}$$

that is,

$$\begin{aligned} u_{t,t} = u - \quad t \quad & \Big(H(x, Du) + G(x, Du)\Big) + t^2\left\langle D_pG(x, Du),\ D^2\, u\, D_pH(x, Du)\right\rangle \\ + \quad \frac{t^2}{2} \quad & \Big(\left\langle D_pH(x, Du),\ D_xH(x, Du)\right\rangle + \left\langle D_pG(x, Du),\ D_xG(x, Du)\right\rangle\Big) \\ t^2 \quad & \left\langle D_pG(x, Du),\ D_xH(x, Du)\right\rangle + o(t^2), \end{aligned}$$

We now make the symmetrical computation for $u^{t,t}$ and we subtract to get

$$u_{t,t} - u^{t,t} = t^2\Big(\left\langle D_pG(x, Du),\ D_xH(x, Du)\right\rangle - \left\langle D_pH(x, Du),\ D_xG(x, Du)\right\rangle\Big) + o(t^2).$$

The left–hand side term is 0 by the commutation hypothesis, so the assertion follows by letting $t \to 0$ and by exploiting the fact that u, and hence Du, is arbitrary. $\square$

References

[1] L. Ambrosio, O. Ascenzi, G. Buttazzo, Lipschitz Regularity for Minimizers of Integral Functionals with Highly discontinuous Integrands. *J. Math. Anal. Appl.* **142**, no. 2 (1989), 301–316.

[2] G. Barles, Solutions de viscosité des équations de Hamilton–Jacobi. Mathématiques & Applications, 17. Springer–Verlag, Paris, 1994.

[3] G. Barles, P.E. Souganidis, On the large time behavior of solutions of Hamilton–Jacobi equations, *SIAM J. Math. Anal.* **31**, no. 4 (2000), 925–939

[4] G. Barles, A. Tourin, Commutation properties of semigroups for first-order Hamilton-Jacobi equations and application to multi-time equations. *Indiana Univ. Math. J.*. **50**, no. 4 (2001), 1523–1544.

[5] P. Bernard, Existence of $C^{1,1}$ critical sub-solutions of the Hamilton-Jacobi equation on compact manifolds, *Ann. Sci. École Norm. Sup. (4)*, **40**, no. 3 (2007), 445–452.

[6] G. Buttazzo, M. Giaquinta, S. Hildebrandt, One–dimensional variational problems. An introduction. Oxford Lecture Series in Mathematics and its Applications, **15**. The Clarendon Press, Oxford University Press, New York, 1998.

[7] P. Cannarsa, C. Sinestrari, Semiconcave functions, Hamilton–Jacobi Equations, and Optimal Control. Progress in Nonlinear Differential Equations and their Applications, 58. Birkhuser Boston, Inc., Boston, MA, 2004.

[8] F. Cardin, C. Viterbo, Commuting Hamiltonians and Hamilton-Jacobi multi-time equations. *Duke Math. J.* **144**, no. 2 (2008), 235284.

[9] X. Cui, J. Li, On commuting Tonelli Hamiltonians: Autonomous case. *preprint* (2009).

[10] X. Cui, J. Li, On commuting Tonelli Hamiltonians: Time–Periodic case. *preprint* (2009).

[11] A. Cellina, The Classical Problem of the Calculus of Variations in the Autonomous Case: Relaxation and Lipschitzianity of Solutions. *Trans. Amer. Math. Soc.*, **356**, no. 1 (2004), pp. 415–426.

[12] F.H. Clarke, Optimization and nonsmooth analysis. Wiley, New York, 1983.

[13] F.H. Clarke, R.B. Vinter, Regularity properties of solutions to the basic problem in the Calculus of Variations. *Trans. Amer. Math. Soc.* **289** (1985), 73–98.

[14] G. Contreras, R. Iturriaga, Global Minimizers of Autonomous Lagrangians. 22nd Brazilian Mathematics Colloquium, IMPA, Rio de Janeiro, 1999.

[15] A. Davini, Bolza Problems with discontinuous Lagrangians and Lipschitz continuity of the value function. *SIAM J. Control Optim.* **46**, no. 5 (2007), 1897–1921.

[16] A. Davini, A. Siconolfi, A generalized dynamical approach to the large time behavior of solutions of Hamilton–Jacobi equations. *SIAM J. Math. Anal.*, Vol. 38, , no. 2 (2006), 478–502.

[17] R. DeMarr, Common fixed points for commuting contraction mappings, *Pacific J. Math.*, Vol. 13, (1963), 1139–1141.

[18] A. Fathi, Weak Kam Theorem in Lagrangian Dynamics. Cambridge University Press, *to appear.*

[19] A. Fathi, Sur la convergence du semi-groupe de Lax-Oleinik. *C. R. Acad. Sci. Paris Sér. I Math.*, Vol. 327, no. 3 (1998), 267–270.

[20] A. Fathi, A. Siconolfi, PDE aspects of Aubry–Mather theory for continuous convex Hamiltonians. *Calc. Var. Partial Differential Equations* **22**, no. 2 (2005) 185–228.

[21] M. Gromov, Pseudoholomorphic curves in symplectic manifolds. *Invent. Math.*, **82**, (1985), no. 2, 307–347.

[22] C. Imbert, M. Volle, On vectorial Hamilton–Jacobi equations. *Control Cybernet.*, **31**, no. 3 (2002), 493–506

[23] H. Ishii, Weak KAM aspects of convex HamiltonJacobi equations with Neumann type boundary conditions. *J. Math. Pures Appl.*, **95**, no. 1 (2011), 99–135

[24] P.–L. Lions, J.–C. Rochet, Hopf formula and multitime Hamilton-Jacobi equations, *Proc. Amer. Math. Soc.*, **96**, no. 1 (1986), 79–84.

[25] J.N. Mather, Action minimizing invariant measures for positive definite Lagrangian systems, *Math. Z.* , **207**, no. 2 (1991), 169–207.

[26] J.N. Mather, Variational construction of connecting orbits, *Ann. Inst. Fourier (Grenoble),* **43**, no. 5 (1993), 1349–1386.

[27] M. Motta, F. Rampazzo, Nonsmooth multi-time Hamilton-Jacobi systems. *Indiana Univ. Math. J..* **55**, no. 5 (2006), 1573–1614.

[28] A. Siconolfi, Hamilton–Jacobi equations and weak KAM Theory. Encyclopedia of Complexity and Systems Science. Springer–Verlag, (2009), 4540–4561.

[29] A. Sorrentino, On the integrability of Tonelli Hamiltonians, *Trans. Amer. Math. Soc.*, to appear. (Preprint (2009) ArXiv: 0903.4300)

[30] M. Zavidovique, Weak KAM for commuting Hamiltonians. *Nonlinearity*, **23**, no. 4 (2010), 793–808.

Dip. di Matematica, Università di Roma "La Sapienza", P.le Aldo Moro 2, 00185 Roma, Italy
E-mail address: davini@mat.uniroma1.it

UMPA (ENS de Lyon), 46 allée d'Italie, 69364 Lyon cedex 07
E-mail address: mzavidov@umpa.ens-lyon.fr