

# Weak KAM Theory: the connection between Aubry-Mather theory and viscosity solutions of the Hamilton-Jacobi equation

Albert Fathi

**Abstract.** The goal of this lecture is to explain to the general mathematical audience the connection that was discovered in the last 20 or so years between the Aubry-Mather theory of Lagrangian systems, due independently to Aubry and Mather in low dimension, and to Mather in higher dimension, and the theory of viscosity solutions of the Hamilton-Jacobi equation, due to Crandall and Lions, and more precisely the existence of global viscosity solutions due to Lions, Papanicolaou, and Varadhan.

**Mathematics Subject Classification (2010).** Primary 37J50, 35F21; Secondary 70H20.

**Keywords.** Lagrangian, Hamiltonian, Hamilton-Jacobi, Aubry-Mather, weak KAM.

## 1. Introduction

This lecture is not intended for specialists, but rather for the general mathematical audience. Lagrangian Dynamical Systems have their origin in classical physics, especially in celestial mechanics. The Hamilton-Jacobi method is a way to obtain trajectories of a Lagrangian system through solutions of the Hamilton-Jacobi equation. However, solutions of this equation easily develop singularities. Therefore for a long time, only local results were obtained. Since the 1950's, several major developments both on the dynamical side, and the PDE side have taken place. In the 1980's, on the dynamical side there was the famous Aubry-Mather theory for twist maps, discovered independently by S. Aubry [2] and J.N. Mather [20], and its generalization to higher dimension by J.N. Mather [21, 22] in the framework of classical Lagrangian systems. On the PDE side, there was the viscosity theory of the Hamilton-Jacobi equation, due to M. Crandall and P.L. Lions [8], which introduces weak solutions for this equation, together with the existence of global solutions for the stationary Hamilton-Jacobi equation on the torus obtained by P.L. Lions, G. Papanicolaou, and S.R.S. Varadhan [18]. In 1996, the author found the connection between these apparently unrelated results: the Aubry and the Mather sets can be obtained from the global weak (=viscosity) solutions. Moreover, these sets serve as natural uniqueness sets for the stationary Hamilton-Jacobi equation, see [13]. Independently, a little bit later Weinan E [11] found the connection for twist maps, with some partial ideas for higher dimensions, and L.C. Evans and D. Gomes [12] showed how to obtain Mather measures from the PDE point of view.

In this introduction, we quickly explain some of these results.

---

■ Proceedings of the International Congress of Mathematicians, Seoul, 2014

Although all results are valid for Tonelli Hamiltonians defined on the cotangent space  $T^*M$  of a compact manifold, in this introduction, we will stick to the case where  $M = \mathbb{T}^k = \mathbb{R}^k/\mathbb{Z}^k$ .

A Tonelli Hamiltonian  $H$  on  $\mathbb{T}^k$  is a function  $H : \mathbb{T}^k \times \mathbb{R}^k \rightarrow \mathbb{R}$ ,  $(x, p) \mapsto H(x, p)$ , where  $x \in \mathbb{T}^k$  and  $p \in \mathbb{R}^k$ , which satisfies the following conditions:

- (i) The Hamiltonian  $H$  is  $C^r$ , with  $r \geq 2$ .
- (ii) (Strict convexity in the momentum) The second derivative  $\partial^2 H/\partial p^2(x, p)$  is positive definite, as a quadratic form, for every  $(x, p) \in \mathbb{T}^k \times \mathbb{R}^k$ .
- (iii) (Superlinearity)  $H(x, p)/\|p\|$  tends to  $+\infty$ , uniformly in  $x \in \mathbb{T}^k$ , as  $\|p\|$  tends to  $+\infty$ , where  $\|\cdot\|$  is the Euclidean norm.

In fact, it is more accurate to consider  $H$  as a function on the *cotangent* bundle  $\mathbb{T}^k \times (\mathbb{R}^k)^*$ , where  $(\mathbb{R}^k)^*$  is the vector space dual to  $\mathbb{R}^k$ . To avoid complications, in this introduction, we identify  $(\mathbb{R}^k)^*$  to  $\mathbb{R}^k$  in the usual way (using the canonical scalar product).

There is a flow  $\phi_t^*$  associated to the Hamiltonian. This flow is given by the ODE

$$\begin{aligned} \dot{x} &= \frac{\partial H}{\partial p}(x, p) \\ \dot{p} &= -\frac{\partial H}{\partial x}(x, p). \end{aligned} \tag{1.1}$$

It is easy to see that the Hamiltonian  $H$  is constant along solutions of the ODE. Since the level sets of the function  $H$  are compact by the superlinearity condition (iii), this flow  $\phi_t^*$  is defined for all  $t \in \mathbb{R}$ , and therefore is a genuine dynamical system.

The following theorem is due to John Mather [21, 22] with a contribution by Mário Jorge Dias Carneiro [9].

**Theorem 1.1.** *There exists a convex superlinear function  $\alpha : \mathbb{R}^k \rightarrow \mathbb{R}$  such that for every  $P \in \mathbb{R}^k$ , we can find a non-empty compact subset  $\tilde{\mathcal{A}}^*(P) \subset \mathbb{T}^k \times \mathbb{R}^k$ , called the Aubry set of  $H$  for  $P$ , satisfying:*

- 1) *The set  $\tilde{\mathcal{A}}^*(P)$  is non-empty and compact.*
- 2) *The set  $\tilde{\mathcal{A}}^*(P)$  is invariant by the flow  $\phi_t^*$ .*
- 3) *The set  $\tilde{\mathcal{A}}^*(P)$  is a graph on the base  $\mathbb{T}^k$ , i.e. the restriction of the projection  $\pi : \mathbb{T}^k \times \mathbb{R}^k \rightarrow \mathbb{T}^k$  to  $\tilde{\mathcal{A}}^*(P)$  is injective.*
- 4) *The set  $\tilde{\mathcal{A}}^*(P)$  is included in the level set  $\{(x, p) \in \mathbb{T}^k \times \mathbb{R}^k \mid H(x, p) = \alpha(P)\}$ .*

John Mather [21, 22] gave also a characterization of the probability measures invariant by  $\phi_t^*$  whose support is included in the Aubry set  $\tilde{\mathcal{A}}^*(P)$ .

**Theorem 1.2.** *For every  $P \in \mathbb{R}^k$ , and every Borel probability measure  $\tilde{\mu}$  on  $\mathbb{T}^k \times \mathbb{R}^k$  which is invariant by the flow  $\phi_t^*$ , we have*

$$-\alpha(P) \leq \int_{\mathbb{T}^k \times \mathbb{R}^k} \frac{\partial H}{\partial p}(x, p)[p - P] - H(x, p) d\tilde{\mu}(x, p),$$

*with equality if and only if  $\tilde{\mu}(\tilde{\mathcal{A}}^*(P)) = 1$ , i.e. the support of  $\tilde{\mu}$  is contained in  $\tilde{\mathcal{A}}^*(P)$ .*

Part 4) Theorem 1.1 is the contribution of Mário Jorge Dias Carneiro. It leads us to a connection with the Hamilton-Jacobi equation. In fact, there is a well-known way to obtain invariant sets which are both graphs on the base and contained in a level set of  $H$ . It is given by the Hamilton-Jacobi theorem.

**Theorem 1.3** (Hamilton-Jacobi). *Let  $u : \mathbb{T}^k \rightarrow \mathbb{R}$  be a  $C^2$  function. If  $P \in \mathbb{R}^k$ , the graph*

$$\text{Graph}(P + \nabla u) = \{(x, P + \nabla u(x)) \mid x \in \mathbb{T}^k\}$$

*is invariant under the Hamiltonian flow of  $H$  if and only if  $H$  is constant on  $\text{Graph}(P + \nabla u)$ . i.e. if and only if  $u$  is a solution of the (stationary) Hamilton-Jacobi equation*

$$H(x, P + \nabla u(x)) = c, \text{ for every } x \in \mathbb{T}^k,$$

*where  $c$  is a constant independent of  $x$ .*

Therefore, it is tempting to try to obtain the Aubry-Mather sets from invariant graphs. This cannot be done with  $u$  smooth, since this would give too many invariant tori in general Hamiltonian systems, see the explanations in the next section.

In fact, Crandall and Lions [8] developed a notion of weak PDE solution for the Hamilton-Jacobi equation, called viscosity solution. The following global existence theorem was obtained by Lions, Papanicolaou, and Varadhan [18].

**Theorem 1.4.** *Suppose that  $H : \mathbb{T}^k \times \mathbb{R}^k \rightarrow \mathbb{R}$  is continuous and satisfies the superlinearity condition (iii) above. For every  $P$ , there exists a unique constant  $\bar{H}(P)$  such that the Hamilton-Jacobi equation*

$$H(x, P + \nabla u(x)) = \bar{H}(P) \tag{1.2}$$

*admits a global weak (viscosity) solution  $u : \mathbb{T}^k \rightarrow \mathbb{R}$ .*

The solutions obtained in this last theorem are automatically Lipschitz due to the superlinearity of  $H$ . Of course, if we add a constant to a solution of equation (1.2) we still obtain a solution. However, it should be emphasized that there may be a pair of solutions whose difference is not a constant.

Theorem 1.4 above was obtained in 1987, and Mather's work [21] was essentially completed by 1990. John Mather visited the author at the University of Florida in Gainesville in the fall of 1988, and explained that he had obtained some results on existence of Aubry sets for Lagrangians in higher dimension, i.e. beyond twist maps.

In 1996, the author obtained the following result, see [13].

**Theorem 1.5** (Weak Hamilton-Jacobi). *The function  $\alpha$  of Mather, and the function  $\bar{H}$  of Lions-Papanicolaou-Varadhan are equal. Moreover, if  $u : \mathbb{T}^k \rightarrow \mathbb{R}$  is a weak (=viscosity) solution of*

$$H(x, P + \nabla u(x)) = \bar{H}(P) = \alpha(P), \tag{1.3}$$

*then the graph*

$$\text{Graph}(P + \nabla u) = \{(x, \nabla u(x)) \mid \text{for } x \in \mathbb{T}^k, \text{ such that } u \text{ has a derivative at } x\}$$

*satisfies the following properties:*

- 1) *Its closure  $\overline{\text{Graph}(P + \nabla u)}$  is compact, and projects onto the whole of  $\mathbb{T}^k$ .*

2) For every  $t > 0$ , we have  $\phi_{-t}^*(\overline{\text{Graph}(P + \nabla u)}) \subset \text{Graph}(P + \nabla u)$ .

Therefore, the subset  $\tilde{\mathcal{I}}^*(P + u) = \bigcap_{t \geq 0} \phi_{-t}^*(\overline{\text{Graph}(P + \nabla u)})$  is compact non-empty and invariant under  $\phi_t^*$ , for every  $t \in \mathbb{R}$ , and the closure  $\overline{\text{Graph}(P + \nabla u)}$  is contained in the unstable subset  $W^u(P + \tilde{\mathcal{I}}^*(u))$  defined by

$$W^u(\tilde{\mathcal{I}}^*(P + u)) = \{(x, p) \in \mathbb{T}^k \times \mathbb{R}^k \mid \phi_t^*(x, p) \rightarrow \tilde{\mathcal{I}}^*(u), \text{ as } t \rightarrow -\infty\}.$$

Moreover, the Aubry set  $\tilde{\mathcal{A}}^*(P)$  for  $H$  is equal to the intersection of the sets  $\tilde{\mathcal{I}}^*(P + u)$ , where the intersection is taken on all weak (viscosity) solutions of equation (1.3).

In fact, denoting by  $\tilde{\mathcal{A}}_H^*(P)$ , and  $\alpha_H$ , the Aubry sets and the  $\alpha$  function for the Hamiltonian  $H$ , it is not difficult to see that  $\alpha_H(P) = \alpha_{H_P}(0)$ , and also that  $\tilde{\mathcal{A}}_H^*(P)$  can be obtained from  $\tilde{\mathcal{A}}_{H_P}^*(0)$ , where  $H_P$  is the Tonelli Hamiltonian defined by  $H_P(x, p) = H(x, P + p)$ . Therefore we will later on only give the proof of Theorem 1.5 for the case  $P = 0$ .

It is the author’s strong belief that the real discoverer of the above theorem should have been Ricardo Mañé. His untimely death in 1995 prevented him from discovering this theorem as can be attested by his last work [19].

The reader should also be aware that what we are covering is just the beginning of weak KAM theory. It is 18 years old. It does not do justice to the marvelous contributions done by others in this subject since 1996. The author strongly apologizes to all these mathematicians who have carried the theory way beyond the author’s imagination or wildest dream.

## 2. Motivation

Some motivation for Aubry-Mather, and hence for weak KAM theory, came from celestial mechanics, and problems related to more general classical mechanical systems studied by Lagrangian or Hamiltonian methods.

We will give a (very partial) description of this motivation. There are also some historical comments. The reader should not take them seriously. They are here for the sake of a good story. The author does not claim that this historical account is accurate.

Although celestial mechanics is about the motion of several bodies in  $\mathbb{R}^3$  with different masses, we will use a simplified model, and start with the motion of a free particle of mass  $m$  in the Euclidean space  $\mathbb{R}^k$  (if  $k = 3n$ , this is also the motion of  $n$  particles in  $\mathbb{R}^3$ , all with same mass  $m$ ). The trajectory  $\gamma : \mathbb{R} \rightarrow \mathbb{R}^k$  of such a particle satisfies  $\ddot{\gamma}(t) = 0$ , for all  $t \in \mathbb{R}$ . Therefore  $\gamma(t) = x + tv$ , where  $x = \gamma(0)$  is the initial position and  $v$  is the initial speed. The speed of the trajectory is the time derivative  $\dot{\gamma}$ , in particular  $v = \dot{\gamma}(0)$ . It is better to convert the second order ODE given by  $\ddot{\gamma}(t) = 0$  to a first order ODE on the configuration space  $\mathbb{R}^k \times \mathbb{R}^k$  taking into account both position and speed. A point in  $\mathbb{R}^k \times \mathbb{R}^k$  will be denoted by  $(x, v)$ , where  $x \in \mathbb{R}^k$  is the position component and  $v \in \mathbb{R}^k$  is the speed component. The speed curve of  $\gamma$  is  $\Gamma(t) = (\gamma(t), \dot{\gamma}(t))$ . This curve takes values in  $\mathbb{R}^k \times \mathbb{R}^k$  and satisfies the first order ODE

$$\dot{\Gamma}(t) = X_0(\Gamma(t)),$$

where the vector field  $X_0$  on  $\mathbb{R}^k \times \mathbb{R}^k$  is given by  $X_0(x, v) = (v, 0)$ . Conversely, any solution of this ODE is a possible speed curve of a free particle of mass  $m$ . The solutions of the ODE yield a flow  $\phi_t^0$  on  $\mathbb{R}^k \times \mathbb{R}^k$ , defined by

$$\phi_t^0(x, v) = (x + tv, v).$$

Observe that the sets  $\mathbb{R}^k \times \{v\}$ ,  $v \in \mathbb{R}^k$  give a decomposition of  $\mathbb{R}^k \times \mathbb{R}^k$  into subsets which are invariant by the flow  $\phi_t^0$ . We will address the following problem: if we perturb this system a little bit can we still see such a pattern, i.e. a (partial) decomposition, into invariant subsets?

To make things more precise, we add a smooth (at least  $C^2$ ) potential  $V : \mathbb{R}^k \rightarrow \mathbb{R}$  to our mechanical system. To avoid problems caused by non-compactness, we will assume that  $V$  is  $\mathbb{Z}^k$  periodic, i.e. it satisfies  $V(x+z) = V(x)$ , for all  $x \in \mathbb{R}^k$ , and all  $z \in \mathbb{Z}^k$ . Therefore  $V$  is defined on  $\mathbb{T}^k = \mathbb{R}^k / \mathbb{Z}^k$ . The equation of motion is now given by the Newton equation

$$m\ddot{\gamma}(t) = -\nabla V(\gamma(t)).$$

Again this defines a first order ODE on  $\mathbb{T}^k \times \mathbb{R}^k$  using the vector field

$$X(x, v) = \left(v, -\frac{1}{m} \nabla V[\gamma(t)]\right).$$

This ODE has a flow on  $\mathbb{T}^k \times \mathbb{R}^k$  which we will denote by  $\phi_t$ . The orbits of our flow are precisely the speed curves of possible motions of a particle in the potential  $V$ .

Before proceeding further, it is convenient to recall the Lagrangian and Hamiltonian aspects of a classical mechanical system since they will play a major role in the theory. The Lagrangian  $L : \mathbb{T}^k \times \mathbb{R}^k$  is defined by

$$L(x, v) = \frac{1}{2} m \|v\|^2 - V(x),$$

where  $\|\cdot\|$  is the usual Euclidean norm on  $\mathbb{R}^k$ . Using this Lagrangian, the Newton equation becomes

$$\frac{d}{dt} \left[ \frac{\partial L}{\partial v}(\gamma(t), \dot{\gamma}(t)) \right] = \frac{\partial L}{\partial x}(\gamma(t), \dot{\gamma}(t)). \quad (2.1)$$

The equation above is the Euler-Lagrange equation associated to the Lagrangian  $L$ . It shows that the trajectories are extremal curves for the Lagrangian, as we now explain. A Lagrangian like  $L$  is used to define the action  $\mathbb{L}(\gamma)$  of the curve  $\gamma : [a, b] \rightarrow \mathbb{T}^k$  by

$$\mathbb{L}(\gamma) = \int_a^b L(\gamma(s), \dot{\gamma}(s)) ds.$$

A curve  $\gamma : [a, b] \rightarrow \mathbb{T}^k$  is called a minimizer (for  $L$ ) if for every curve  $\delta : [a, b] \rightarrow \mathbb{T}^k$ , with  $\delta(a) = \gamma(a)$ ,  $\delta(b) = \gamma(b)$ , we have  $\mathbb{L}(\delta) \geq \mathbb{L}(\gamma)$ . These curves play a particular role in Aubry-Mather theory. They have to be found among the curves which are critical points for the action functional  $\mathbb{L}$ . These critical points are called extremals. More precisely, a curve  $\gamma : [a, b] \rightarrow \mathbb{T}^k$  is called an extremal for  $L$ , if the functional  $\mathbb{L}$  on the space of curves  $\delta : [a, b] \rightarrow \mathbb{T}^k$ , with  $\delta(a) = \gamma(a)$ ,  $\delta(b) = \gamma(b)$ , has a vanishing derivative  $D_\gamma \mathbb{L}$  at  $\gamma$ . By the classical theory of Calculus of Variations, this is the case if and only if  $\gamma$  satisfies the Euler-Lagrange equation (2.1). Therefore the possible trajectories of our particle for the potential  $V$  are precisely the extremals for  $L$ .

For the Hamiltonian aspects, one has to introduce the dual variable  $p = mv$ . In fact, this dual variable should be understood as an element of the dual space  $(\mathbb{R}^k)^*$ , which means that  $p$  should be considered as the linear form  $\langle p, \cdot \rangle$  on  $\mathbb{R}^k$ . A better way to think of  $p$  is to define it by  $p = \partial L / \partial v(x, v)$ . The Hamiltonian  $H : \mathbb{T}^k \times (\mathbb{R}^k)^*$  is then defined by

$$H(x, p) = \frac{1}{2m} \|p\|^2 + V(x).$$

It is not difficult to see that  $H$  is also given by

$$H(x, p) = \max_{v \in \mathbb{R}^k} p(v) - L(x, v).$$

The Legendre transform  $\mathcal{L} : \mathbb{T}^k \times \mathbb{R}^k \rightarrow \mathbb{T}^k \times (\mathbb{R}^k)^*$  is a diffeomorphism defined by

$$\mathcal{L}(x, v) = \left(x, \frac{\partial L}{\partial v}(x, v)\right).$$

If one uses the Legendre transform to transport the flow  $\phi_t$  to the flow  $\phi_t^* = \mathcal{L}\phi_t\mathcal{L}^{-1}$  on  $\mathbb{T}^k \times (\mathbb{R}^k)^*$ , using the Euler-Lagrange equation and the definition of  $H$ , it is not difficult to see that  $\phi_t^*$  is the flow of the ODE (1.1).

$$\begin{aligned} \dot{x} &= \frac{\partial H}{\partial p}(x, p) \\ \dot{p} &= -\frac{\partial H}{\partial x}(x, p). \end{aligned}$$

This means that  $\phi_t^*$  is the Hamiltonian flow associated to  $H$ .

Since we are now interested in perturbing the motion of the free particle, we will denote by  $\phi_t^V, L_V, H_V, \dots$  the objects associated to the potential  $V$ . Of course, for  $V = 0$ , we get back the flow  $\phi_t^0$ , or rather the induced flow on the quotient  $\mathbb{T}^k \times \mathbb{R}^k$ . In that case  $L_0(x, v) = \|v\|^2/2$ , and  $H_0(x, p) = \|p\|^2/2$ , the flow  $\phi_t^0$  is the geodesic flow of the flat canonical metric on  $\mathbb{T}^k$ , and  $\phi_t^{*0}$  is the geodesic flow on the cotangent bundle. The decomposition into invariant sets for the flow  $\phi_t^{*0}$  is given by  $\{(x, p) \mid p = P\}, P \in \mathbb{R}^k$ . Notice that this is the graph of the solution  $u = 0$  of the Hamilton-Jacobi equation

$$H_0(x, P + d_x u) = \frac{1}{2}\|P\|^2.$$

One could try to understand the persistence or non-persistence of the invariant sets by trying to solve for  $V$  small the Hamilton-Jacobi equation

$$H_V(x, P + d_x u) = c(P).$$

Unfortunately, it is almost impossible to find a  $C^1$  solution of such an equation for a given  $V$ , and all  $P$ . In fact, as we now see in the simple example of a pendulum, there must be some condition on  $P$  to be able to do that.

**Example 2.1.** We consider the function  $V_\epsilon(t) = \epsilon \cos 2\pi t$  on the 1-dimensional torus  $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ . The Hamiltonian  $H_\epsilon(x, p) = 1/2p^2 + \epsilon \cos 2\pi t$  has levels which are 1-dimensional. The Hamiltonian flow  $\phi_t^\epsilon$  is the flow of the ODE

$$\begin{aligned} \dot{x} &= p \\ \dot{p} &= -2\pi\epsilon \sin(2\pi x). \end{aligned}$$

Therefore the flow  $\phi_t^\epsilon$  has exactly two fixed points  $(0, 0)$  and  $(1/2, 0)$ . There are also two orbits homoclinic to the fixed point  $(0, 0)$  (i.e. converging to the fixed point when  $t \rightarrow \pm\infty$ ). The union of the fixed point  $(0, 0)$  and its two homoclinic orbits is the level  $H_\epsilon = \epsilon$ , see the figure below. The other orbits of the Hamiltonian flow are periodic. A level set  $H_\epsilon = c$  is

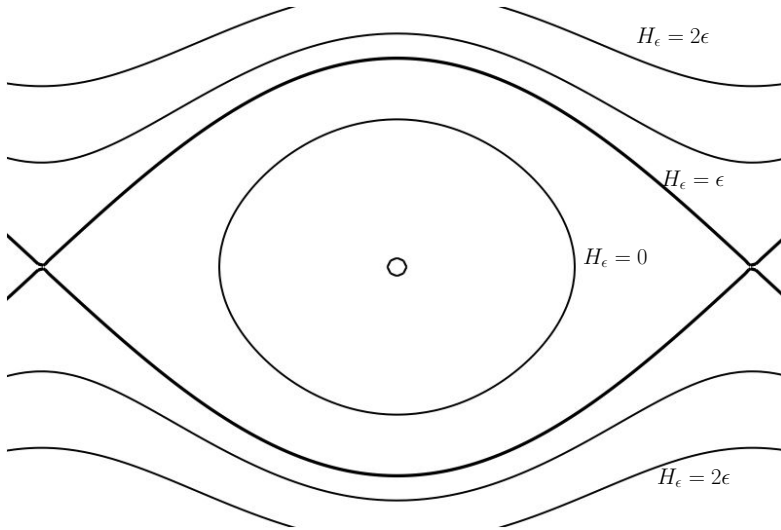


Figure 1. The pendulum.

just one orbit if  $c < \epsilon$ , and a pair of orbits if  $c > \epsilon$ . The level set for  $c = \epsilon$  is given by the equation

$$\frac{1}{2}p^2 + \epsilon \cos 2\pi x = \epsilon.$$

Hence the region  $\{(x, p) \mid H_\epsilon(x, p) \leq \epsilon\}$  is enclosed between the two graphs  $p = \pm\sqrt{\epsilon}(2 - 2 \cos 2\pi x)^{1/2}$ . The area  $A_\epsilon$  of  $\{(x, p) \mid H_\epsilon(x, p) \leq \epsilon\}$  rescales as  $A_\epsilon = \sqrt{\epsilon}A_1$ , where  $A_1 > 0$  is the area of  $\{(x, p) \mid p^2 + 2 \cos 2\pi x \leq 2\}$ . Suppose that for a given  $P \in \mathbb{R}$ , we can find, for some  $c \in \mathbb{R}$ , a  $C^1$  solution  $u : \mathbb{T} \rightarrow \mathbb{R}$  of

$$H_\epsilon(x, P + u'(x)) = c. \tag{2.2}$$

This implies that the level set  $H_\epsilon = c$ , which is a subset of  $\mathbb{T} \times \mathbb{R}$ , projects onto the whole of  $\mathbb{T}$ . Therefore  $c \geq \epsilon$ , and the area between the curve  $p = P + u'(x)$ , and the curve  $p = 0$  is an absolute value larger than  $A_\epsilon/2$ , i.e.  $|\int_{\mathbb{T}} P + u'(x) dx| \geq A_\epsilon/2$ . But  $\int_{\mathbb{T}} u'(x) dx = 0$ , since  $u'$  is the derivative of a  $C^1$  function on  $\mathbb{T}$ . It follows that  $|P| \geq \sqrt{\epsilon}A_1/2$ . In particular, the set  $p = 0$  does not deform to an invariant set for  $\epsilon$  as small as we want.

Note that  $\phi_t^{*V_\epsilon}$  still remembers part of the set  $p = 0$ . In fact, the points in  $p = 0$  are fixed points for  $\phi_t^{*0}$ , and the flow  $\phi_t^{*V_\epsilon}$  must also have fixed points, because the fixed points of  $\phi_t^{*V_\epsilon}$  are precisely the critical points of  $H_{V_\epsilon}$ , and by superlinearity the function  $H_{V_\epsilon}$  must have critical points (at least a minimum) on  $\mathbb{T} \times \mathbb{R}$ .

Another fact that can be readily seen on Figure 1, is that for  $|P| \geq \epsilon$ , the Hamilton-Jacobi equation (2.2) has a solution. This solution is  $C^\infty$  for  $|P| > \epsilon$ . But it is only  $C^1$  for  $|P| = \epsilon$ , in which case the derivative  $u'$  is only piecewise  $C^1$  because its graph has a corner at  $x = 0$ .

In fact, there are always problems with resonances. This goes back to the work of Henri Poincaré [24] on the three body problem. To explain this in our case, we come back to the flow  $\phi_t^0$  defined on the tangent bundle of  $\mathbb{T}^k$ . The invariant sets are given by  $T_v = \{(x, v) \mid x \in \mathbb{T}^k\}$ ,  $v \in \mathbb{R}^k$ . If the coordinates of  $v$  are all rational, then the motion on  $T_v$

is periodic. It can be shown that perturbing the system destroys most of the  $T_v$ 's. However necessarily some periodic orbits must still exist. In fact, the periodic orbits on  $T_v$  are all in the same homotopy class, and they minimize, in that homotopy class, the action for the Lagrangian  $L_0(x, v) = \|v\|^2/2$ . If we perturb the Lagrangian  $L_0$  to a Tonelli Lagrangian  $L$ , by the direct method in the Calculus of Variations, there are minimizers of the  $L$ -action in this homotopy class. For a long time, it was believed that most of the  $T_v$ 's would disappear under a general perturbation except maybe for some periodic orbits. It came as a surprise, when A.N. Kolmogorov [17] announced the stability property for  $T_v$ , for  $v$  far away from the rational vectors, at least for analytic perturbations. This was extended by V.I. Arnold [1], and also by J. Moser [23] to cover differentiable perturbations in the  $C^r$  topology. This is the now famous KAM theory. In fact, not only do the  $T_v$  persist for some  $v$ 's, but they persist for more and more  $v$ 's as the size of the perturbation becomes smaller and smaller, for example in the  $C^\infty$  topology. The set of  $v$ 's for which this is possible tends to a set of full Lebesgue measure as the perturbation vanishes.

It turns out that the KAM method proves more than what we just said. Fix a  $v_0 \in \mathbb{R}^k$  to which the KAM theorem applies. The invariant set  $T_{v_0}$  for  $\phi_t^0$  is a torus and on that torus  $\phi_t^0$  is the linear flow  $(t, x) \mapsto (x + tv_0)$ . For  $V$  small enough, KAM theory finds a smooth imbedding map  $i_{V,v_0} : T_{v_0} \rightarrow \mathbb{T}^k \times \mathbb{R}^k$  such that:

- 1) the image  $i_{V,v_0}(T_{v_0})$  is invariant under  $\phi_t^V$ ;
- 2) the imbedding  $i_{V,v_0}$  is a conjugation between the linear flow  $\phi_t^0|_{T_{v_0}}$  and the restriction of  $\phi_t^V$  to the image  $i_{V,v_0}(T_{v_0})$ .

So not only does the set persist (with a deformation) but the dynamics remain the same. The imbedding  $i_{V,v_0}$  is the identity for  $V = 0$ , and it depends continuously on  $V$ . Moreover, we have:

- 3) the image  $\mathcal{L}_V(i_{V,v_0}(T_{v_0}))$  of  $i_{V,v_0}(T_{v_0})$  by the Legendre transform, which is invariant under  $\phi_t^{*V}$ , is a Lagrangian graph on the base. This means that we can find  $P_{V,v_0} \in \mathbb{R}^k$ , and a smooth function  $u_{V,v_0} : \mathbb{T}^k \rightarrow \mathbb{R}$  such that  $\mathcal{L}_V(i_{V,v_0}(T_{v_0})) = \text{Graph}(P_{V,v_0} + du_{V,v_0}) = \{(x, P_{V,v_0} + d_x u_{V,v_0}) \mid x \in \mathbb{T}^k\}$ .

Since this graph  $\text{Graph}(P_{V,v_0} + du_{V,v_0})$  is invariant by  $\phi_t^{*V}$ , by the Hamilton-Jacobi theorem the function  $u_{V,v_0}$  solves the equation  $H_V(x, P_{V,v_0} + d_x u_{V,v_0}) = c_{V,v_0}$ , for some constant  $c_{V,v_0}$ . Hence, although the KAM theory is rooted in Dynamical Systems and tries to find a part of the dynamics that is conjugated to a simple linear dynamic on the torus, it nevertheless produces smooth solutions to the Hamilton-Jacobi equation.

Of course, it remained to understand what happens to the invariant tori when they disappear. In 1982, independently, Aubry [2] and Mather [20] studied twist maps on the annulus (they can be thought as giving examples of a discretization of Tonelli Hamiltonians on  $\mathbb{T}^2$ ). They showed that the invariant circles of the standard twist diffeomorphism  $(x, r) \mapsto (x + r, r)$  of  $\mathbb{T} \times [0, 1]$  never completely disappear. In fact, periodic orbits persist for  $r$  rational, and for  $r$  irrational there exists either an invariant Cantor subset or an invariant curve. In all cases, the invariant sets are Lipschitz graphs on (part of) the base  $\mathbb{T}$ . It is important to note that these results are not only perturbative, but that they also hold for all area preserving twist maps of  $\mathbb{T} \times [0, 1]$ .

Around 1989, John Mather extended the existence of these sets to Tonelli Hamiltonians in higher dimension [21, 22].



### 3. The general setting

We will consider the more general setting of a Hamiltonian  $H : T^*M \rightarrow \mathbb{R}$  defined on the cotangent space  $T^*M$  of the compact connected manifold without boundary  $M$ . We will denote by  $(x, p)$  a point in  $T^*M$ , where  $x \in M$ , and  $p \in T_x^*M$ .

The Hamiltonian  $H$  is said to be Tonelli, if it satisfies conditions (i), (ii), and (iii) of the Introduction. Only condition (iii) needs an explanation. We replace the Euclidean norm on  $\mathbb{R}^k$ , by any family of norms  $\|\cdot\|_x, x \in M$ , on the fibers of  $TM \rightarrow M$ , coming from a Riemannian metric on  $M$ . Note that, by the compactness of  $M$ , any two such families are uniformly equivalent, i.e. their ratio is uniformly bounded away from 0 and from  $+\infty$ . Therefore condition (iii) is

(iii) (Superlinearity)  $H(x, p)/\|p\|_x \rightarrow +\infty$ , uniformly in  $x \in M$ , as  $\|p\|_x \rightarrow +\infty$ .

The Hamiltonian flow  $\phi_t^*$  is still defined on  $T^*M$ . In local coordinates in  $M$  it is still the flow of the ODE (1.1). The flow is complete because  $H$  is constant on orbits, and has compact level sets by superlinearity.

We introduce the Lagrangian  $L : TM \rightarrow \mathbb{R}$ , defined on the tangent bundle  $TM$  of  $M$ , by

$$L(x, v) = \sup_{p \in T_x^*M} p(v) - H(x, p). \quad (3.1)$$

Since  $H$  is superlinear, this sup is always attained. Moreover, since the function  $p \mapsto p(v) - H(x, p)$  is  $C^1$  and strictly convex, this sup is achieved at the only  $p$  at which its derivative vanishes, namely the only  $p$ , where  $v = \partial H / \partial p(x, p)$ .

The Lagrangian  $L$  is as differentiable as  $H$  is, and it satisfies the Tonelli properties (i), (ii), (iii) of the introduction. The Legendre transform  $\mathcal{L} : TM \rightarrow T^*M$  is defined by

$$\mathcal{L}(x, v) = \left(x, \frac{\partial L}{\partial v}(x, v)\right). \quad (3.2)$$

Using the Tonelli properties, it can be shown that  $\mathcal{L} : TM \rightarrow T^*M$  is a global  $C^{r-1}$  diffeomorphism. Moreover, its inverse is given by

$$\mathcal{L}^{-1}(x, p) = \left(x, \frac{\partial H}{\partial p}(x, p)\right). \quad (3.3)$$

Definition (3.1) of the Lagrangian yields the Fenchel inequality

$$p(v) \leq L(x, v) + H(x, p). \quad (3.4)$$

Furthermore, there is equality in the Fenchel inequality if and only if  $(x, p) = \mathcal{L}(x, v)$ , which is equivalent to  $p = \partial L / \partial v(x, v)$ , and also to  $v = \partial H / \partial p(x, p)$ .

Since for any given  $p \in T_x^*M$ , we can find a  $v \in T_xM$ , for which the Fenchel inequality is an equality, we obtain

$$H(x, p) = \sup_{v \in T_xM} p(v) - L(x, v). \quad (3.5)$$

The Lagrangian  $L$  is used to define the action  $\mathbb{L}(\gamma)$  of the curve  $\gamma : [a, b] \rightarrow M$  by

$$\mathbb{L}(\gamma) = \int_a^b L(\gamma(s), \dot{\gamma}(s)) ds.$$

The notion of minimizer and extremal for  $L$  are the same as in §2 above. By the classical theory of Calculus of Variations, the curve  $\gamma : [a, b] \rightarrow M$  is an extremal if and only if it satisfies, in local coordinates on  $M$ , the Euler-Lagrange equation

$$\frac{d}{dt} \left[ \frac{\partial L}{\partial v}(\gamma(t), \dot{\gamma}(t)) \right] = \frac{\partial L}{\partial x}(\gamma(t), \dot{\gamma}(t)). \tag{3.6}$$

If we carry out the derivation with respect to  $t$  in this last equation, we get

$$\frac{\partial^2 L}{\partial v^2}(\gamma(t), \dot{\gamma}(t))(\ddot{\gamma}(t), \cdot) = \frac{\partial L}{\partial x}(\gamma(t), \dot{\gamma}(t))(\cdot) - \frac{\partial^2 L}{\partial x \partial v}(\gamma(t), \dot{\gamma}(t))(\dot{\gamma}(t), \cdot),$$

where the dot  $\cdot$  means that we consider maps on the linear space  $T_x M$ . Since  $L$  is Tonelli, the bilinear form  $\partial^2 L / \partial v^2(\gamma(t), \dot{\gamma}(t))$  is invertible. Therefore we can solve for  $\ddot{\gamma}(t)$ , and obtain  $\ddot{\gamma}(t) = X(\gamma(t), \dot{\gamma}(t))$ , where  $X$  is a vector field  $U \rightarrow \mathbb{R}^k$ , with  $U$  a coordinate patch in  $M$ , and  $k = \dim(M)$ . The solutions of this second order ODE are exactly the extremals, a concept which does not depend on the choice of a coordinate system. Hence, these local second order ODE's define a global second order ODE on  $M$ . Taking into account not only position, but also speed, it becomes a first order ODE on  $TM$ , which is called the Euler-Lagrange ODE, and its flow  $\phi_t$  is called the Euler Lagrange flow. We give in the next proposition the well-known properties of the Euler Lagrange flow.

**Proposition 3.1.** *If  $\gamma : [a, b] \rightarrow M$  is an extremal for the Lagrangian  $L$ , then its speed curve  $t \mapsto (\gamma(t), \dot{\gamma}(t))$  is a piece of an orbit of the Euler-Lagrange flow  $\phi_t$ , i.e., we have  $(\gamma(t), \dot{\gamma}(t)) = \phi_{t-t_0}(\gamma(t_0), \dot{\gamma}(t_0))$ , for all  $t_0, t \in [a, b]$ .*

*Conversely, denoting by  $\pi : TM \rightarrow M$  is the canonical projection, for every  $(x, v) \in TM$ , the curve  $\gamma_{(x,v)}(t) = \pi \phi_t(x, v)$  is an extremal, whose speed curve satisfies  $(\gamma_{(x,v)}(t), \dot{\gamma}_{(x,v)}(t)) = \phi_t(x, v)$ .*

We now come to the relation between the Euler-Lagrange flow and the Hamiltonian flow.

**Proposition 3.2.** *The Legendre transform  $\mathcal{L} : TM \rightarrow T^*M$  is a conjugacy between the Euler-Lagrange flow  $\phi_t$  and the Hamiltonian flow  $\phi_t^*$ . This means that  $\phi_t^* = \mathcal{L} \phi_t \mathcal{L}^{-1}$ . In particular, the flow  $\phi_t$  is complete, since this is the case for  $\phi_t^*$ .*

An important property of Tonelli Lagrangians is existence and regularity of minimizers.

**Theorem 3.3** (Tonelli [5, 7, 14, 21]). *Suppose that  $L$  is a  $C^r$  Tonelli Lagrangian on the compact manifold  $M$ . For every  $x, y \in M$ , for every  $a, b \in \mathbb{R}$ , with  $a < b$ , we can find a minimizer  $\gamma : [a, b] \rightarrow M$ , with  $\gamma(a) = x$ , and  $\gamma(b) = y$ . Moreover, any minimizer is automatically a  $C^r$  extremal. In particular, its speed curve is a piece of an orbit of the Euler-Lagrange flow.*

We now define  $h_t(x, y)$  as the minimal action of a curve from  $x$  to  $y$  in the time  $t > 0$ .

$$h_t(x, y) = \inf_{\gamma} \int_a^b L(\gamma(s), \dot{\gamma}(s)) ds, \tag{3.7}$$

where the infimum is taken over all piecewise  $C^1$  curves  $\gamma : [a, b] \rightarrow M$ , with  $b - a = t$ ,  $\gamma(a) = x$ , and  $\gamma(b) = y$ . Since  $L$ , in our setting, does not depend on time, the action of a curve  $\gamma : [a, b] \rightarrow M$  is the same as the action of any of its shifted in time curves

$\gamma_\sigma : [a + \sigma, b + \sigma] \rightarrow M, \gamma_\sigma(s) = \gamma(s - \sigma)$ , with  $\sigma \in \mathbb{R}$ . Therefore, if  $a_0, b_0$  are fixed with  $b_0 - a_0 = t$ , to define  $h_t(x, y)$  we could have taken the infimum over all curves  $\gamma : [a_0, b_0] \rightarrow M$ , with  $\gamma(a_0) = x$ , and  $\gamma(b_0) = y$ . Moreover, by Tonelli's theorem the infimum is always achieved on a  $C^r$  curve, if  $L$  is  $C^r, r \geq 2$ . Hence we could have restricted the curves to obtain the infimum to  $C^r$  curves (even to  $C^\infty$  by density, although the minimizer may not be  $C^\infty$  if the Lagrangian  $L$  is not  $C^\infty$ ).

Here are the elementary properties of  $h_t$

**Lemma 3.4.** *If  $L$  is a  $C^r$  Tonelli Lagrangian, and  $d$  is a distance on  $M$  obtained from a Riemannian metric, we have:*

- 1) *For every  $x, y$ , and every  $a, b \in \mathbb{R}$  with  $b - a = t$ , there exists a  $C^r$  minimizer  $\gamma : [a, b] \rightarrow M$ , with  $\gamma(a) = x, \gamma(b) = y$  such that  $h_t(x, y) = \int_a^b L(\gamma(s), \dot{\gamma}(s)) ds$ .*
- 2) *There exists a finite constant  $A$  such that  $h_t(x, x) \leq At$ , for every  $x \in M$ .*
- 3) *There exists a finite constant  $B$  such that  $h_{d(x,y)}(x, y) \leq Bd(x, y)$ , for every  $x, y \in M$ , with  $x \neq y$ .*
- 4) *For every  $x, y \in M$ , and every  $t, t' > 0$ , we have*

$$h_{t+t'}(x, y) = \inf_{z \in M} h_t(x, z) + h_{t'}(z, y).$$

- 5) *For every  $K \geq 0$ , we can find a finite constant  $C(K)$  such that*

$$h_t(x, y) \geq Kd(x, y) + C(K)t, \text{ for all } x, y \in M.$$

*Proof.* Part 1) is a consequence of Tonelli's theorem 3.3 above.

To prove Part 2), we use the constant path  $s \mapsto x$ , to obtain

$$h_t(x, x) \leq tL(x, 0) \leq At, \text{ with } A = \max_{x \in M} L(x, 0).$$

We now prove part 3). By the compactness of  $M$ , we can find a geodesic  $\gamma : [0, d(x, y)] \rightarrow M$  parametrized by arc-length (i.e.  $\|\dot{\gamma}(s)\|_{\gamma(s)} = 1$  everywhere), with  $\gamma(0) = x$ , and  $\gamma(d(x, y)) = y$ . If we set

$$B = \sup\{L(x, v) \mid x \in M, v \in T_x M, \|v\|_x \leq 1\},$$

we see that  $h_{d(x,y)}(x, y) \leq \mathbb{L}(\gamma) \leq Bd(x, y)$ .

Part 4) follows from the fact that to go from  $x$  to  $y$  in time  $t + t'$ , we have first to go in time  $t$  to some point  $z \in M$  then we go from  $z$  to  $y$  in time  $t'$ , and, moreover, the action for the concatenated path is the sum of the action of the path from  $x$  to  $z$  and of the action of the path from  $z$  to  $y$ .

For part 5), fix  $K \geq 0$ . We first prove that there exists a finite constant  $C(K)$  such that

$$L(x, v) \geq K\|v\|_x + C(K), \tag{3.8}$$

By the superlinearity of  $L$ , we know that  $(L(x, v) - K\|v\|_x)/\|v\|_x$  tends uniformly to  $+\infty$  as  $\|v\|_x \rightarrow +\infty$ . By the compactness of  $M$ , it follows that the constant  $C(K) = \inf_{TM} L(x, v) - K\|v\|_x$  is finite. Therefore, the inequality (3.8) holds with this  $C(K)$ .

Given a curve  $\gamma : [a, b] \rightarrow M$ , if we apply (3.8) with  $x = \gamma(s)$ , and  $v = \dot{\gamma}(s)$ , and integrate on  $[a, b]$ , we obtain

$$\mathbb{L}(\gamma) \geq K \int_a^b \|\dot{\gamma}(s)\|_{\gamma(s)} ds + C(K)(b - a).$$

But the Riemannian length  $\int_a^b \|\dot{\gamma}(s)\|_{\gamma(s)} ds$  of  $\gamma$  is  $\geq d(\gamma(a), \gamma(b))$ . Hence

$$\mathbb{L}(\gamma) \geq Kd(\gamma(a), \gamma(b)) + C(K)(b - a).$$

Taking the infimum on all paths  $\gamma : [0, t] \rightarrow M$ , with  $\gamma(0) = x, \gamma(t) = y$  finishes the proof of part 5). □

We now come to the most important property of  $h_t$ . This is what started weak KAM theory. This property has been discovered independently by many people. When the author himself discovered it back in 1996, he explained it to Michel Herman in his office in Paris. After hearing it, Michel Herman opened the drawer of his desk, got out a copy of the paper of W.H. Fleming [16] published in 1969, which contained an equivalent form of this statement. This is the oldest instance that the author knows of the following lemma.

**Lemma 3.5** (Fleming, [16]). *For every  $t_0 > 0$ , the family of functions  $h_t : M \times M \rightarrow \mathbb{R}$ ,  $t \geq t_0$  is equi-Lipschitzian.*

For the proof of Fleming’s lemma see §8 below. In fact, more is true: the family  $h_t : M \times M \rightarrow \mathbb{R}$ ,  $t \geq t_0$  is equi-semiconcave. Again this fact has been well-known for sometime now in the theory of viscosity solutions [6]. For a proof, in our setting, of this extension of Fleming’s lemma see the appendices of [15].

#### 4. The Lax-Oleinik semi-group, and its fixed points

Rather than introducing the theory of viscosity solutions, we are going to give its evolution semi-group, i.e. the semi-group obtained by solving (in the viscosity sense) the equation

$$\frac{\partial U}{\partial t}(t, x) + H(x, \frac{\partial U}{\partial x}(t, x)) = 0,$$

on  $[0, +\infty[ \times M$  with given initial condition  $u : M \rightarrow \mathbb{R}$ , for  $t = 0$ . This semi-group  $T_t^-$ , called the Lax-Oleinik semi-group, acts on the space  $\mathcal{C}^0(M, \mathbb{R})$  of real-valued continuous functions on  $M$ . It can be expressed directly in that case using the functions  $h_t, t > 0$ , by

$$T_t^- u(x) = \inf_{y \in M} u(y) + h_t(y, x). \tag{4.1}$$

This definition is valid for  $t > 0$ , of course  $T_0^-$  is the identity. Notice that by Fleming’s lemma 3.5, not only is  $T_t^- u$  continuous for  $t > 0$ , but for every  $t_0 > 0$  the whole family  $T_t^- u$ , for  $t \geq t_0, u \in \mathcal{C}^0(M, \mathbb{R})$  is equi-Lipschitzian. By the Ascoli-Arzelá theorem, this suggests that the image of  $T_t^-$  is “almost” relatively compact. In fact, to be able to apply the Ascoli-Arzelá theorem, we would also need uniform boundedness. This is not the case but as we will see, we can easily overcome this small difficulty.

We first give the properties of  $T_t^-$ .

**Proposition 4.1.** *The Lax-Oleinik satisfies the following properties:*

- 1) For every  $u \in \mathcal{C}^0(M, \mathbb{R})$ , and every  $c \in \mathbb{R}$ , we have  $T_t^-(c + u) = c + T_t^-(u)$ , for all  $t \in [0, +\infty[$ .
- 2) For every  $u, v \in \mathcal{C}^0(M, \mathbb{R})$ , with  $u \leq v$ , we have  $T_t^- u \leq T_t^- v$ , for all  $t \in [0, +\infty[$ .
- 3) For every  $u, v \in \mathcal{C}^0(M, \mathbb{R})$ , we have  $\|T_t^- u - T_t^- v\|_0 \leq \|u - v\|_0$ , for all  $t \in [0, +\infty[$ , where  $\|\cdot\|_0$  is the sup (or  $\mathcal{C}^0$ ) norm on  $\mathcal{C}^0(M, \mathbb{R})$ .
- 4) The family  $T_t^-$ ,  $t \geq 0$  is a semi-group, i.e. for every  $u \in \mathcal{C}^0(M, \mathbb{R})$ , we have  $T_{t+t'}^- u = T_t^- [T_{t'}^- (u)]$ .
- 5) For every given  $u \in \mathcal{C}^0(M, \mathbb{R})$ , the curve  $t \mapsto T_t^- u$  is continuous for the sup norm topology on  $\mathcal{C}^0(M, \mathbb{R})$ .
- 6) For every  $t_0 > 0$ , the family  $T_t^- u$ ,  $t \geq t_0$ ,  $u \in \mathcal{C}^0(M, \mathbb{R})$  is equi-Lipschitz.

*Proof.* Parts 1) and 2) are obvious from the definition of the semi-group  $T_t^-$ .

To show part 3), we observe that  $-\|u - v\|_0 + v \leq u \leq v + \|u - v\|_0$ . Therefore using 2) and 1), we obtain  $-\|u - v\|_0 + T_t^- v \leq T_t^- u \leq T_t^- v + \|u - v\|_0$ , which implies part 3). Part 4) is a consequence of part 4) of Lemma 3.4.

It remains to prove part 5). We first consider the case where  $u : M \rightarrow \mathbb{R}$  is Lipschitz. We prove that  $\|T_t^- u - u\|_0 \rightarrow 0$ , when  $t \rightarrow 0$ . By part 2) of Lemma 3.4 and the definition of  $T_t^-$ , we obtain

$$T_t^- u(x) \leq u(x) + h_t(x, x) \leq u(x) + At.$$

Hence  $T_t^- u - u \leq At$ . If we denote by  $K$  a Lipschitz constant for  $u$ , we have  $u(y) + Kd(y, x) \geq u(x)$ , combining with part 5) of Lemma 3.4, we get

$$u(y) + h_t(y, x) \geq u(y) + Kd(y, x) + C(K)t \geq u(x) + C(K)t.$$

Taking the infimum over  $y \in M$ , we conclude that  $T_t^- u(x) \geq u(x) + C(K)t$ . Hence  $u - T_t^- u \leq -C(K)t$ . Combining the two inequalities yields

$$\|T_t^- u - u\|_0 \leq t \max(A, -C(k)).$$

Therefore  $\|T_t^- u - u\|_0 \rightarrow 0$ , when  $t \rightarrow 0$ .

If  $u \in \mathcal{C}^0(M, \mathbb{R})$ , we can find a sequence of  $\mathcal{C}^1$  functions  $u_n : M \rightarrow \mathbb{R}$  such that  $\|u_n - u\|_0 \rightarrow 0$ , as  $n \rightarrow +\infty$ . Since a  $\mathcal{C}^1$  function on the compact manifold  $M$  is Lipschitz, we have  $\|T_t^- u_n - u_n\|_0 \rightarrow 0$ , as  $t \rightarrow 0$ , for every  $n$ . Since  $\|T_t^- u_n - T_t^- u\|_0 \leq \|u_n - u\|_0$ , for every  $t > 0$ , it is not difficult to conclude that  $\|T_t^- u - u\|_0 \rightarrow 0$ , when  $t \rightarrow 0$ . To show the continuity of  $t \rightarrow T_t^- u$  on  $[0, +\infty[$ , we use the semi-group property 4), and 2), to obtain that for  $t' \geq t$ , we have  $\|T_{t'}^- u - T_t^- u\|_0 = \|T_t^- (T_{t'-t}^- u) - T_t^- u\|_0 \leq \|T_{t'-t}^- u - u\|_0$ . This can be rewritten as  $\|T_{t'}^- u - T_t^- u\|_0 \leq \|T_{|t'-t|}^- u - u\|_0$ , which is valid also in the case  $t \geq t'$ . Therefore the continuity of  $t \rightarrow T_t^- u$  at 0 implies the continuity on  $[0, +\infty[$ .  $\square$

We are now in a position to prove the existence of global weak solutions of the Hamilton-Jacobi equation.

**Theorem 4.2 (Weak KAM Solution).** *We can find  $c \in \mathbb{R}$ , and a function  $u \in \mathcal{C}^0(M, \mathbb{R})$  such that  $u = T_t^- u + ct$ , for every  $t > 0$ . Necessarily  $u$  is Lipschitz, and  $c = -\lim_{t \rightarrow +\infty} T_t^- v/t$ , for every  $v \in \mathcal{C}^0(M, \mathbb{R})$ .*

*Proof.* We define  $\text{Lip}_K(M, \mathbb{R})$  as the subset of Lipschitz functions in  $\mathcal{C}^0(M, \mathbb{R})$  with Lipschitz constant  $\leq K$ . This subset is closed and convex in  $\mathcal{C}^0(M, \mathbb{R})$ . Moreover, if we fix a base point  $x_0 \in M$ , by the Arzelà-Ascoli theorem, the closed convex subset  $\text{Lip}_K^{x_0}(M, \mathbb{R}) = \{u \in \text{Lip}_K(M, \mathbb{R}) \mid u(x_0) = 0\}$  is compact. Fix  $t_0 > 0$ , by part 6) of Proposition 4.1, there exists a constant  $K(t_0)$  such that for every  $t \geq t_0$ , the image of  $T_t^-$  is contained in  $\text{Lip}_{K(t_0)}(M, \mathbb{R})$ . Therefore, for  $t \geq t_0$ , we can define the continuous non-linear operator  $\hat{T}_t^- : \mathcal{C}^0(M, \mathbb{R}) \rightarrow \text{Lip}_{K(t_0)}^{x_0}(M, \mathbb{R})$  by  $u \mapsto T_t^- u - T_t^- u(x_0)$ . Since  $\hat{T}_t^-$  sends the compact convex subset  $\text{Lip}_{K(t_0)}^{x_0}(M, \mathbb{R})$  to itself, by the Schauder-Tykhonov theorem [10, Theorem 2.2, pages 414-415], the map  $\hat{T}_t^-$  has a fixed point. We now show that we can find a common fixed point for the family  $\hat{T}_t^-, t > 0$ . We first note that  $\hat{T}_t^-, t > 0$ , is a semi-group. In fact  $\hat{T}_{t'}^-(\hat{T}_t^- u) = T_{t'}^- \hat{T}_t^- u - T_{t'}^- \hat{T}_t^- u(x_0)$ . Since

$$\begin{aligned} T_{t'}^- \hat{T}_t^- u &= T_{t'}^- (T_t^- u - T_t^- u(x_0)) = T_{t'}^- (T_t^- u) - T_t^- u(x_0) \\ &= T_{t'+t}^- u - T_t^- u(x_0), \end{aligned}$$

we obtain

$$\begin{aligned} \hat{T}_{t'}^- \hat{T}_t^- u &= T_{t'+t}^- u - T_t^- u(x_0) - [T_{t'+t}^- u(x_0) - T_t^- u(x_0)] = T_{t'+t}^- u - T_{t'+t}^- u(x_0) \\ &= \hat{T}_{t'+t}^- u. \end{aligned}$$

This semi-group property implies that  $\text{Fix}(\hat{T}_{1/2^{n+1}}^-) \subset \text{Fix}(\hat{T}_{1/2^n}^-)$ , for every integer  $n \geq 1$ , where  $\text{Fix}(\hat{T}_t^-)$  is the set of fixed points of  $\hat{T}_t^-$  in  $\mathcal{C}^0(M, \mathbb{R})$ . Since, for  $t > 0$ , the non-empty set  $\text{Fix}(\hat{T}_t^-)$  is closed and contained in  $\text{Lip}_{K(t)}^{x_0}(M, \mathbb{R})$ , it is compact. Therefore the non-increasing sequence  $\text{Fix}(\hat{T}_{1/2^n}^-), n \geq 1$ , has a non-empty intersection. If  $u$  is in this intersection, it is fixed by every  $\hat{T}_{1/2^n}^-$ . By the semi-group property we obtain  $u = \hat{T}_t^- u$ , for every  $t$  in the dense set of rational numbers of the form  $p/2^n, p \in \mathbb{N}, n \geq 1$ . But  $t \mapsto \hat{T}_t^- u$  is continuous by part 5) of Proposition 4.1. Hence  $u = \hat{T}_t^- u$ , for every  $t \geq 0$ . Therefore, we obtained a  $u \in \mathcal{C}^0(M, \mathbb{R})$  such that  $u = T_t^- u + c_t$ , for every  $t \geq 0$ , where  $c_t = -T_t^- u(x_0) \in \mathbb{R}$ . Since

$$T_{t'}^- u = T_{t'}^- [T_t^- u + c_t] = T_{t'}^- T_t^- u + c_t = T_{t'+t}^- u + c_t,$$

we infer

$$u = T_{t'}^- u + c_{t'} = T_{t'+t}^- u + c_t + c_{t'}.$$

This implies that  $c_{t'+t} = c_{t'} + c_t$ . The continuity of  $t \mapsto c_t = u - T_t^- u$  implies  $c_t = tc$ , where  $c = c_1$ . This finishes the proof of the existence of  $u$  and  $c$ .

Note that  $u$  is necessarily Lipschitz, since  $T_t^- u$  is Lipschitz for  $t > 0$ . To prove the last claim of the theorem on  $c$ , we first observe that  $T_t^- u/t = u/t - c$ . Since the function  $u$  is bounded on the compact set  $M$ , we do get  $\lim_{t \rightarrow +\infty} T_t^- u/t = -c$ . By part 3) of Proposition 4.1, for any  $v \in \mathcal{C}^0(M, \mathbb{R})$ , we have  $\|T_t^- v/t - T_t^- u/t\|_0 \leq \|v - u\|_0/t \rightarrow 0$ , as  $t \rightarrow +\infty$ .  $\square$

**Definition 4.3** (Critical value). We will denote by  $c(H)$ , or  $c(L)$  the only constant  $c$  for which we can find a weak KAM solution, i.e. the only constant  $c$  for which we can find a function  $u : M \rightarrow \mathbb{R}$ , with  $u = T_t^- u + c$ , for every  $t > 0$ . This constant is called the Mañé critical value.

## 5. Domination and calibration

The proof of the following proposition is straightforward from the definitions.

**Proposition 5.1** (Characterization of subsolutions). *Let  $u : M \rightarrow \mathbb{R}$  be a function, and  $c \in \mathbb{R}$ . The following are equivalent*

- 1) for every  $t > 0$ , we have  $u \leq T_t u + ct$ ;
- 2) for every  $t > 0$ , and every  $x, y \in M$ , we have  $u(y) - u(x) \leq h_t(x, y) + ct$ ;
- 3) for every continuous, piecewise  $C^1$  curve  $\gamma : [a, b] \rightarrow M$ , we have

$$u(\gamma(b)) - u(\gamma(a)) \leq \int_a^b L(\gamma(s), \dot{\gamma}(s)) ds + c(b - a). \quad (5.1)$$

It is convenient to introduce the following definition.

**Definition 5.2** (Domination). If  $u : M \rightarrow \mathbb{R}$  is a function, and  $c \in \mathbb{R}$ , we say that  $u$  is dominated by  $L + c$ , which we denote by  $u \prec L + c$ , if it satisfies inequality (5.1) above, for every piecewise  $C^1$  curve  $\gamma : [a, b] \rightarrow M$ .

**Lemma 5.3.** *There is a constant  $B$  such that any function  $u : M \rightarrow \mathbb{R}$ , dominated by  $L + c$ , is Lipschitz with Lipschitz constant  $\leq B + c$ .*

*Proof.* By the domination condition  $u(y) - u(x) \leq h_{d(x,y)}(x, y) + cd(x, y)$ . Therefore, we obtain  $u(y) - u(x) \leq (B + c)d(x, y)$ , with  $B$  given by part 3) of Lemma 3.4.  $\square$

Recall that by Rademacher's theorem, Lipschitz functions are differentiable a.e.

**Lemma 5.4.** *Let  $u : M \rightarrow \mathbb{R}$  be dominated by  $L + c$ . If the derivative  $d_x u$  exists at some given  $x \in M$ , then  $H(x, d_x u) \leq c$ . In particular, the function  $u$  is an almost everywhere subsolution of the Hamilton-Jacobi equation  $H(x, d_x u) = c$ .*

*Proof.* Suppose  $d_x u$  exists at  $x \in M$ . For a given  $v \in T_x M$ , let  $\gamma : [0, 1] \rightarrow M$  be a  $C^1$  curve with  $\gamma(0) = x, \dot{\gamma}(0) = v$ . Applying (5.1) to the curve  $\gamma|_{[0, t]}$ , for every  $t \in [0, 1]$ , we obtain

$$u(\gamma(t)) - u(\gamma(0)) \leq \int_0^t L(\gamma(s), \dot{\gamma}(s)) ds + ct.$$

Dividing by  $t > 0$  and letting  $t \rightarrow 0$ , we get  $d_{\gamma(0)} u(\dot{\gamma}(0)) \leq L(\gamma(0), \dot{\gamma}(0)) + c$ . By the choice of  $\gamma$ , we conclude that  $d_x u(v) - L(x, v) \leq c$ . But  $H(x, d_x u) = \sup_{v \in T_x M} d_x u(v) - L(x, v)$ . Hence  $H(x, d_x u) \leq c$ .  $\square$

It should not come as a surprise that curves satisfying the equality in (5.1) enjoy special properties. It is convenient to give them a name.

**Definition 5.5** (Calibrated curve). Suppose that  $u : M \rightarrow \mathbb{R}$  is dominated by  $L + c$ . A curve  $\gamma : [a, b] \rightarrow M$  is said to be  $(u, L, c)$ -calibrated if

$$u(\gamma(b)) - u(\gamma(a)) = \int_a^b L(\gamma(s), \dot{\gamma}(s)) ds + c(b - a).$$

Recall that for  $t \in \mathbb{R}$ , and  $\gamma : [a, b] \rightarrow M$ , the curve  $\gamma_t : [a + t, b + t] \rightarrow M$  is defined by  $\gamma_t(s) = \gamma(s - t)$ . Since  $L$  does not depend on time, we have  $\mathbb{L}(\gamma_t) = \mathbb{L}(\gamma)$ .

**Proposition 5.6.** *If  $u \prec L + c$ , then any  $(u, L, c)$ -calibrated curve  $\gamma : [a, b] \rightarrow M$  is a minimizer. In particular, it is as smooth as  $L$ . Moreover, for every  $[a', b'] \subset [a, b]$ , the restriction  $\gamma|_{[a', b']}$  is also  $(u, L, c)$ -calibrated, and so is the curve  $\gamma_t$  for all  $t \in \mathbb{R}$ .*

*Proof.* If  $\delta : [a, b] \rightarrow M$  is a curve with  $\delta(a) = \gamma(a), \delta(b) = \gamma(b)$ , we have

$$\mathbb{L}(\gamma) + c(b - a) = u(\gamma(b)) - u(\gamma(a)) = u(\delta(b)) - u(\delta(a)) \leq \mathbb{L}(\delta) + c(b - a).$$

Therefore  $\mathbb{L}(\gamma) \leq \mathbb{L}(\delta)$ , and  $\gamma$  is a minimizer. The regularity of  $\gamma$  is given by Tonelli's theorem.

We next use the domination  $u \prec L + c$  to obtain

$$\begin{aligned} u(\gamma(a')) - u(\gamma(a)) &\leq \int_a^{a'} L(\gamma(s), \dot{\gamma}(s)) ds + c(a' - a) \\ u(\gamma(b')) - u(\gamma(a')) &\leq \int_{a'}^{b'} L(\gamma(s), \dot{\gamma}(s)) ds + c(b' - a') \\ u(\gamma(b)) - u(\gamma(b')) &\leq \int_{b'}^b L(\gamma(s), \dot{\gamma}(s)) ds + c(b - b'). \end{aligned} \tag{5.2}$$

If we add these three inequalities, we obtain

$$u(\gamma(b)) - u(\gamma(a)) \leq \int_a^b L(\gamma(s), \dot{\gamma}(s)) ds + c(b - a),$$

which is an equality. Therefore the three inequalities in (5.2) are equalities. The middle equality means that  $\gamma|_{[a', b']}$  is  $(u, L, c)$ -calibrated. The last part follows from  $\gamma_t(a + t) = \gamma(a), \gamma_t(b + t) = \gamma(b)$ , and  $\mathbb{L}(\gamma_t) = \mathbb{L}(\gamma)$ .  $\square$

We now extend the notion of calibration to non-compact curves. For a curve  $\gamma : I \rightarrow M$  defined on the not-necessarily compact interval  $I \subset \mathbb{R}$ , we say that  $\gamma$  is  $(u, L, c)$ -calibrated if the restriction  $\gamma|_{[a, b]}$  is  $(u, L, c)$ -calibrated for every compact subinterval  $[a, b] \subset \mathbb{R}$ . By Proposition 5.6 above this definition coincides with Definition 5.5 when  $I$  is compact.

Although a dominated function is differentiable almost everywhere, it might not be obvious to explicitly find a point where the derivative exists. The following lemma provides such points.

**Lemma 5.7.** *Assume that  $u : M \rightarrow \mathbb{R}$  is  $L + c$  dominated, and let  $\gamma : [a, b] \rightarrow M$  be  $(u, L, c)$ -calibrated. We have:*

1) *If  $d_{\gamma(t)}u$  exists at some  $t \in [a, b]$ , then*

$$H(\gamma(t), d_{\gamma(t)}u) = c, \text{ and } d_{\gamma(t)}u = \partial L / \partial v(\gamma(t), \dot{\gamma}(t)).$$

2) *If  $t \in ]a, b[$ , then the derivative  $d_{\gamma(t)}u$  does indeed exist.*



*Proof.* We prove 1) for  $t \in [a, b]$ . The argument can be slightly modified to obtain the proof for  $t \in ]a, b]$ . By Proposition 5.6, for  $t + \epsilon \leq b$ , we have

$$u(\gamma(t + \epsilon)) - u(\gamma(t)) = \int_t^{t+\epsilon} L(\gamma(s), \dot{\gamma}(s)) ds + c\epsilon.$$

Dividing by  $\epsilon > 0$  and letting  $\epsilon \rightarrow 0$ , we obtain  $d_{\gamma(t)}u(\dot{\gamma}(t)) = L(\gamma(t), \dot{\gamma}(t)) + c$ . By (3.5), this implies  $H(\gamma(t), d_{\gamma(t)}u) \geq d_{\gamma(t)}u(\dot{\gamma}(t)) - L(\gamma(t), \dot{\gamma}(t)) = c$ . But by Lemma 5.4, we also know that  $H(\gamma(t), d_{\gamma(t)}u) \leq c$ . Therefore, we get  $c = H(\gamma(t), d_{\gamma(t)}u) = d_{\gamma(t)}u(\dot{\gamma}(t)) - L(\gamma(t), \dot{\gamma}(t))$ . This proves the first part of 1), but also the second one because the last equality shows that we have equality in the Fenchel inequality  $H(\gamma(t), d_{\gamma(t)}u) + L(\gamma(t), \dot{\gamma}(t)) \geq d_{\gamma(t)}u(\dot{\gamma}(t))$ .

To prove part 2), we will construct two  $C^1$  functions  $\psi, \theta : V \rightarrow \mathbb{R}$ , defined on the neighborhood  $V$  of  $x = \gamma(t)$ , and such that

$$\psi(y) \leq u(y) - u(x) \leq \theta(y),$$

on  $V$ , with equality at  $x$ . We leave it to the reader to show that  $d_x\theta = d_x\psi$ , and that this common derivative is also the derivative of  $u$  at  $x$ . We will construct  $\theta$ , since the argument for  $\psi$  is analogous. Let us first choose a domain  $U$  of a smooth chart  $\varphi : U \rightarrow \mathbb{R}^k$  of the manifold  $M$ , with  $x = \gamma(t) \in U$ , we can find  $a' < t < b'$  such that  $\gamma([a', b']) \subset U$ . To simplify notations we use  $\varphi$  to identify  $U$  with its image in  $\mathbb{R}^k$ . For  $y$  close enough to  $x$  the path  $\gamma_y : [a', t] \rightarrow \mathbb{R}^k$  defined by

$$\gamma_y(s) = \gamma(s) + \frac{s - a'}{t - a'}(y - x),$$

will have an image contained in  $U$ , and therefore can be considered as a path in  $M$ . Note that  $\gamma_y$  starts at  $\gamma(a')$ , and ends at  $y$ . Hence by  $u \prec L + c$ , we obtain

$$u(y) - u(\gamma(a')) \leq \int_{a'}^t L(\gamma_y(s), \dot{\gamma}_y(s)) ds + c(t - a')$$

Moreover for  $y = x$ , we have  $\gamma_x = \gamma$ , and the inequality above is an equality. Therefore subtracting the equality at  $x$  from the inequality at  $y$ , we get

$$u(y) - u(x) \leq \int_{a'}^t L(\gamma_y(s), \dot{\gamma}_y(s)) - L(\gamma(s), \dot{\gamma}(s)) ds.$$

We can now define  $\theta(y)$  for  $y$  close to  $x$  by

$$\begin{aligned} \theta(y) &= \int_{a'}^t L(\gamma_y(s), \dot{\gamma}_y(s)) - L(\gamma(s), \dot{\gamma}(s)) ds \\ &= \int_{a'}^t L\left(\gamma(s) + \frac{s - a'}{t - a'}(y - x), \dot{\gamma}(s) + \frac{1}{t - a'}(y - x)\right) - L(\gamma(s), \dot{\gamma}(s)) ds. \end{aligned}$$

From the last expression, it is clear that  $\theta$  is as smooth as  $L$ . Moreover, we have  $u(y) - u(x) \leq \theta(y)$ , and  $\theta(x) = 0$  as required.  $\square$

For  $(x, v) \in TM$ , let us recall that  $\gamma_{(x,v)}$  is the curve defined by  $\gamma_{(x,v)}(t) = \pi\phi_t(x, v)$ , see Proposition 3.1. It satisfies  $(\gamma_{(x,v)}(t), \dot{\gamma}_{(x,v)}(t)) = \phi_t(x, v)$ .

If  $u : M \rightarrow \mathbb{R}$  is dominated by  $L + c$ , for  $a, b \in \mathbb{R}$ , with  $a < b$ , we define the sets  $\tilde{\mathcal{G}}_{a,b}(u)$ , and  $\tilde{\mathcal{G}}_b(u)$  by

$$\begin{aligned} \tilde{\mathcal{G}}_{a,b}(u) &= \{(x, v) \in TM \mid \gamma_{(x,v)} \text{ is } (u, L, c)\text{-calibrated on } [a, b]\}, \\ \tilde{\mathcal{G}}_b(u) &= \{(x, v) \in TM \mid \gamma_{(x,v)} \text{ is } (u, L, c)\text{-calibrated on } ] - \infty, b]\}. \end{aligned} \tag{5.3}$$

Of course, the sets  $\tilde{\mathcal{G}}_{a,b}(u)$  depend not only on  $u$ , but also on  $c$ , and a better notation would be  $\tilde{\mathcal{G}}_{a,b}(u, c)$ . However, we will only use them later with  $u$  a weak KAM solution, and  $c = c(H)$ .

**Proposition 5.8.** *Suppose  $u : M \rightarrow \mathbb{R}$  is dominated by  $L + c$ . Given  $a < b$ , the set  $\tilde{\mathcal{G}}_{a,b}(u)$  is compact. Moreover, any  $(u, L, c)$ -calibrated curve  $\gamma : [a, b] \rightarrow M$  is of the form  $\gamma_{(x,v)}|_{[a, b]}$ , for some  $(x, v) \in \tilde{\mathcal{G}}_{a,b}(u)$ .*

*Proof.* We first observe that  $\tilde{\mathcal{G}}_{a,b}(u)$  is closed in  $TM$ . We have  $(x, v) \in \tilde{\mathcal{G}}_{a,b}(u)$  if and only if

$$u \circ \pi(\phi_b(x, v)) - u \circ \pi(\phi_a(x, v)) = \int_a^b L\phi_s(x, v) ds + c(b - a).$$

It follows that  $\tilde{\mathcal{G}}_{a,b}(u)$  is closed in  $TM$ , since both sides of the equality above are continuous as functions of  $(x, v) \in TM$ .

We now prove the compactness of  $\tilde{\mathcal{G}}_{a,b}(u)$ . By part 1) of Proposition 5.6, we know that  $\gamma_{(x,v)}|_{[a, b]}$  is a minimizer. Therefore by Lemma 8.1, we can find a finite constant  $\kappa_{b-a}$  such that  $\|\dot{\gamma}_{(x,v)}(a)\|_{\gamma_{(x,v)}(a)} \leq \kappa_{b-a}$ , for every  $(x, v) \in \tilde{\mathcal{G}}_{a,b}(u)$ . Since  $(x, v) = \phi_{-a}(\gamma_{(x,v)}(a), \dot{\gamma}_{(x,v)}(a))$ , the compactness of  $\tilde{\mathcal{G}}_{a,b}(u)$  follows.

If  $\gamma : [a, b] \rightarrow M$  is  $(u, L, c)$ -calibrated it is a minimizer. Hence its speed curve satisfies  $(\gamma(t), \dot{\gamma}(t)) = \phi_{t-a}(\gamma(a), \dot{\gamma}(a))$ . Therefore  $\gamma = \gamma_{(x,v)}|_{[a, b]}$ , where  $(x, v) = \phi_{-a}(\gamma(a), \dot{\gamma}(a))$ . □

**Proposition 5.9.** *Suppose  $u \prec L + c$ , for  $a' \leq a \leq b \leq b'$ , and  $t \in \mathbb{R}$ , we have*

- 1)  $\tilde{\mathcal{G}}_b(u) = \cap_{a < b} \tilde{\mathcal{G}}_{a,b}(u)$ ,  $\tilde{\mathcal{G}}_{a',b'}(u) \subset \tilde{\mathcal{G}}_{a,b}(u)$ , and  $\tilde{\mathcal{G}}_{b'}(u) \subset \tilde{\mathcal{G}}_b(u)$ .
- 2)  $\phi_{-t}\tilde{\mathcal{G}}_{a,b}(u) = \tilde{\mathcal{G}}_{a+t,b+t}(u)$ , and  $\phi_{-t}\tilde{\mathcal{G}}_b(u) = \tilde{\mathcal{G}}_{b+t}(u)$ .
- 3)  $\mathcal{L}[\tilde{\mathcal{G}}_b(u)] \subset \text{Graph}(du)$ , for  $b > 0$ .
- 4)  $H \circ \mathcal{L}[\tilde{\mathcal{G}}_0(u)] = c$ , and  $\mathcal{L}[\tilde{\mathcal{G}}_0(u)] \subset \overline{\text{Graph}(du)}$ , where

$$\text{Graph}(du) = \{(x, d_x u) \mid \text{for } x \in M \text{ at which } d_x u \text{ exists}\}. \tag{5.4}$$

*Proof.* A curve  $\gamma : ] - \infty, b] \rightarrow M$  is  $(u, L, c)$ -calibrated if and only if its restriction to any compact interval  $[a, b]$ ,  $a < b$  is  $(u, L, c)$ -calibrated. This proves the first equality. The inclusions follow from Proposition 5.6. We prove the equality  $\phi_{-t}\tilde{\mathcal{G}}_{a,b}(u) = \tilde{\mathcal{G}}_{a+t,b+t}(u)$ . The equality  $\phi_{-t}\tilde{\mathcal{G}}_b(u) = \tilde{\mathcal{G}}_{b+t}(u)$  follows from this last one by taking intersections over  $a < b$ . We have  $(x, v) \in \phi_{-t}\tilde{\mathcal{G}}_{a,b}(u)$  if and only if  $\phi_t(x, v) \in \tilde{\mathcal{G}}_{a,b}(u)$ . This is equivalent to  $\gamma_{\phi_t(x,v)}$  is  $(u, L, c)$ -calibrated on  $[a, b]$ . This last condition is equivalent to  $\gamma_{(x,v)}$  is  $(u, L, c)$ -calibrated on  $[a + t, b + t]$ , since  $\gamma_{\phi_t(x,v)}(s) = \gamma_{(x,v)}(s + t)$ . Hence  $(x, v) \in \phi_{-t}\tilde{\mathcal{G}}_{a,b}(u)$  if and only if  $(x, v) \in \tilde{\mathcal{G}}_{a+t,b+t}(u)$ .

We now prove parts 3) and 4). If  $(x, v) \in \tilde{\mathcal{G}}_b(u)$ ,  $b > 0$ , by Lemma 5.7, since  $(x, v) = (\gamma_{(x,v)}(0), \dot{\gamma}_{(x,v)}(0))$ , the function  $u$  has a derivative at  $\gamma_{(x,v)}(0) = x$ , which satisfies

$$H(x, d_x u) = c, \text{ and } d_x u = \partial L / \partial v(x, v).$$

Hence  $\mathcal{L}(x, v) = (x, d_x u) \in \text{Graph}(du)$ , and  $H \circ \mathcal{L}(x, v) = c$ . To finish the proof, we note that  $\phi_{-b} \tilde{\mathcal{G}}_0(u) = \tilde{\mathcal{G}}_b(u)$ , for  $b > 0$ . Therefore  $\mathcal{L}[\phi_{-b} \tilde{\mathcal{G}}_0] \subset \text{Graph}(du)$ , and  $H\mathcal{L}[\phi_{-b} \tilde{\mathcal{G}}_0] = c$ . If we let  $b \rightarrow 0$ , we obtain  $\mathcal{L}[\tilde{\mathcal{G}}_0(u)] \subset \overline{\text{Graph}(du)}$ .  $\square$

**Proposition 5.10.** *Let  $u : M \rightarrow \mathbb{R}$  be a weak KAM solution, then*

$$\pi(\tilde{\mathcal{G}}_0(u)) = M, \text{ and } \mathcal{L}[\tilde{\mathcal{G}}_0(u)] = \overline{\text{Graph}(du)}.$$

Moreover  $H(x, d_x u) = c(H)$  at every point  $x \in M$  where  $d_x u$  exists.

*Proof.* We first prove that  $\pi^{-1}(x) \cap \tilde{\mathcal{G}}_0(u)$  is not empty, for every  $x \in M$ . Since  $\tilde{\mathcal{G}}_0(u)$  is the decreasing intersection of the compact sets  $\tilde{\mathcal{G}}_{[-t,0]}(u)$ ,  $t > 0$ , see Proposition 5.9, it suffices to show that for a given  $x \in M$ , and a given  $t > 0$ , we have  $\pi^{-1}(x) \cap \tilde{\mathcal{G}}_{[-t,0]}(u) \neq \emptyset$ . Since, the function  $u$  is a weak KAM solution, we have  $u(x) = \inf_{y \in M} u(y) + h_t(y, x) + c(H)t$ . By the compactness of  $M$  and the continuity of both  $u$  and  $h_t$ , we can find  $y \in M$  such that  $u(x) = u(y) + h_t(y, x) + c(H)t$ . By part 1) of Lemma 3.4, we can find  $\gamma : [-t, 0] \rightarrow M$ , with  $\gamma(-t) = y$ ,  $\gamma(0) = x$ , and

$$h_t(y, x) = \int_{-t}^0 L(\gamma(s), \dot{\gamma}(s)) ds.$$

Hence

$$u(\gamma(0)) - u(\gamma(-t)) = \int_{-t}^0 L(\gamma(s), \dot{\gamma}(s)) ds + c(H)t,$$

and  $\gamma$  is  $(u, L, c(H))$ -calibrated. This implies  $(x, \dot{\gamma}(0)) = (\gamma(0), \dot{\gamma}(0)) \in \tilde{\mathcal{G}}_{[-t,0]}(u)$ . Therefore  $\pi^{-1}(x) \cap \tilde{\mathcal{G}}_{[-t,0]}(u) \neq \emptyset$ , as was to be shown.

From the previous Proposition 5.9, we already know that the compact set  $\mathcal{L}[\tilde{\mathcal{G}}_0(u)]$  is contained in the closure  $\overline{\text{Graph}(du)}$ . To finish the proof, it suffices to show that  $(x, d_x u) \in \mathcal{L}[\tilde{\mathcal{G}}_0(u)]$ , for every  $x$  at which  $d_x u$  exists. Fix such an  $x$ . By the first part of the proposition, we can find  $v \in T_x M$  with  $(x, v) \in \tilde{\mathcal{G}}_0(u)$ . Therefore, the curve  $\gamma_{(x,v)}$  is  $(u, L, c)$ -calibrated on  $]-\infty, 0]$ , with  $(\gamma_{(x,v)}(0), \dot{\gamma}_{(x,v)}(0)) = (x, v)$ . By Lemma 5.7, we have  $(x, d_x u) = \mathcal{L}(x, v)$ .  $\square$

**Corollary 5.11.** *A  $C^1$  weak KAM solution is a solution of the Hamilton-Jacobi equation  $H(x, d_x u) = c(H)$ .*

We will prove the converse of Corollary 5.11 in §9.

## 6. The weak Hamilton-Jacobi theorem

**Theorem 6.1** (Weak Hamilton-Jacobi theorem). *If  $u : M \rightarrow \mathbb{R}$  is a weak KAM solution, then*

$$\phi_{-t}^*(\overline{\text{Graph}(du)}) \subset \text{Graph}(du),$$

for every  $t > 0$ . Therefore the intersection

$$\tilde{\mathcal{I}}^*(u) = \bigcap_{t \geq 0} \phi_{-t}^*[\text{Graph}(du)] = \bigcap_{t \geq 0} \phi_{-t}^*[\overline{\text{Graph}(du)}] \tag{6.1}$$

is a non-empty compact  $\phi_t^*$ -invariant set, contained in  $\text{Graph}(du)$ . This implies that  $\tilde{\mathcal{I}}^*(u)$  is a (partial) graph on the base  $M$ .

If we set  $\tilde{\mathcal{I}}(u) = \mathcal{L}^{-1}[\tilde{\mathcal{I}}^*(u)]$ , then this last set is non-empty, compact,  $\phi_t$ -invariant, and is also a graph on the base  $M$ . Moreover, we have

$$\tilde{\mathcal{I}}(u) = \{(x, v) \in TM \mid \gamma_{(x,v)} \text{ is } (u, L, c(H))\text{-calibrated on } ]-\infty, +\infty[ \}. \tag{6.2}$$

Both sets  $\tilde{\mathcal{I}}(u), \tilde{\mathcal{I}}^*(u)$  are called the Aubry set of the weak KAM solution  $u$ .

*Proof.* By Propositions 5.9 and 5.10, for  $t > 0$ , we know that  $\phi_{-t}\tilde{\mathcal{G}}_0 = \tilde{\mathcal{G}}_t$  is decreasing,  $\mathcal{L}[\tilde{\mathcal{G}}_0(u)] = \overline{\text{Graph}(du)}$ , and  $\tilde{\mathcal{G}}_t \subset \overline{\text{Graph}(du)}$ . Since the diffeomorphism  $\mathcal{L}$  conjugates  $\phi_t$  and  $\phi_t^*$ , we obtain  $\phi_{-t}^*(\overline{\text{Graph}(du)}) \subset \text{Graph}(du)$ , for every  $t > 0$ . This implies (6.1). The non-emptiness follows from the fact that one of the intersections in (6.1) is a decreasing intersection of compact sets. The graph property follows from the inclusion  $\tilde{\mathcal{I}}^*(u) \subset \text{Graph}(du)$ .

To prove (6.2), using again the conjugacy property of  $\mathcal{L}$ , we obtain that  $\tilde{\mathcal{I}}(u)$  is the decreasing intersection of  $\phi_{-t}(\mathcal{L}^{-1}[\overline{\text{Graph}(du)}]) = \phi_{-t}\tilde{\mathcal{G}}_0 = \tilde{\mathcal{G}}_t, t > 0$ . Hence a point  $(x, v)$  is in  $\tilde{\mathcal{I}}(u)$  if and only if it is in  $\tilde{\mathcal{G}}_t(u)$ , for every  $t > 0$ . By definition of  $\tilde{\mathcal{G}}_t(u)$ , this means that  $\gamma_{(x,v)}$  is  $(u, L, c(H))$ -calibrated on  $]-\infty, t]$ , for every  $t > 0$ , or equivalently  $\gamma_{(x,v)}$  is  $(u, L, c(H))$ -calibrated on  $]-\infty, +\infty[$ .  $\square$

### 7. Mather measures, Aubry and Mather sets

Let  $\tilde{\mu}$  be a Borel probability measure on  $TM$ . Since  $L$  is bounded below the integral  $\int_{TM} L d\tilde{\mu} \in \mathbb{R} \cup \{+\infty\}$  always makes sense. Moreover, if  $u : M \rightarrow \mathbb{R}$  is a continuous function then  $u \circ \pi$  is continuous bounded on  $TM$ , therefore  $u \circ \pi$  is  $\tilde{\mu}$ -integrable.

**Theorem 7.1.** *Suppose that  $\tilde{\mu}$  is a Borel probability measure on  $TM$  which is invariant under the Euler-Lagrange flow  $\phi_t$ , then*

$$\int_{TM} L d\tilde{\mu} \geq -c(H).$$

Moreover, there are such invariant measures  $\tilde{\mu}$  which realize the equality.

In fact, if  $u : M \rightarrow \mathbb{R}$  is a weak KAM solution then an invariant measure  $\tilde{\mu}$  satisfies  $\int_{TM} L d\tilde{\mu} = -c(H)$  if and only if the support  $\text{supp}(\tilde{\mu})$  of  $\tilde{\mu}$  is contained in the Aubry set  $\tilde{\mathcal{I}}(u)$  of  $u$ .

*Proof.* If  $L$  is not  $\tilde{\mu}$  integrable then  $\int_{TM} L d\tilde{\mu} = +\infty$ , and there is nothing to prove. Therefore we can assume that  $L$  is integrable for  $\tilde{\mu}$ .

Fix a weak KAM solution  $u$ . For  $(x, v) \in TM$ , expressing the domination condition  $u \prec L + c(H)$  along the curve  $\gamma_{(x,v)}(s) = \pi\phi_s(x, v)$  yields

$$u \circ \pi(\phi_{t'}(x, v)) - u \circ \pi(\phi_t(x, v)) \leq \int_t^{t'} L(\phi_s(x, v)) ds + c(H)(t' - t), \tag{7.1}$$

for all  $t, t' \in \mathbb{R}$ , with  $t \leq t'$ , and all  $(x, v) \in TM$ . If we integrate this inequality with respect to the measure  $\tilde{\mu}$ , we obtain

$$\int_{TM} u\pi\phi_{t'} d\tilde{\mu} - \int_{TM} u\pi\phi_t d\tilde{\mu} \leq \int_{TM} \int_t^{t'} L\phi_s ds d\tilde{\mu} + c(H)(t' - t).$$

By the  $\phi_s$ -invariance of  $\tilde{\mu}$ , the left hand side above is 0. Moreover, using Fubini theorem together with the  $\phi_s$ -invariance on the right hand side, we find that the inequality above is

$$0 \leq (t' - t) \int_{TM} L d\tilde{\mu} + c(H)(t' - t). \quad (7.2)$$

This of course implies  $\int_{TM} L d\tilde{\mu} \geq -c(H)$ .

We have  $\int_{TM} L d\tilde{\mu} = -c(H)$ , if and only if (7.2) is an equality. But this last inequality was obtained by integration of (7.1), therefore (7.2) is an equality if and only if (7.1) is an equality for  $\tilde{\mu}$ -almost every  $(x, v) \in TM$ . Since both sides of (7.1) are continuous in  $(x, v)$ , we conclude that  $\int_{TM} L d\tilde{\mu} = -c(H)$  if and only if (7.1) is an equality on the support on  $\text{supp}(\tilde{\mu})$ . By (6.2) this last condition is equivalent to  $\text{supp}(\tilde{\mu}) \subset \tilde{\mathcal{I}}(u)$ .

Since the compact set  $\tilde{\mathcal{I}}(u)$  is non-empty and invariant by the flow, we can find an invariant measure  $\tilde{\mu}$  with  $\text{supp}(\tilde{\mu}) \subset \tilde{\mathcal{I}}(u)$ . Therefore  $\int_{TM} L d\tilde{\mu} = -c(H)$ .  $\square$

**Definition 7.2** (Mather measures, Mather set). A Mather measure (for the Lagrangian  $L$ ) is a Borel probability  $\phi_s$ -invariant measure  $\tilde{\mu}$  satisfying  $\int_{TM} L d\tilde{\mu} = -c(H)$ . The Mather set  $\tilde{\mathcal{M}}$  (of the Lagrangian  $L$ ) is the closure of  $\bigcup_{\tilde{\mu}} \text{supp} \tilde{\mu}$ , where the union is taken over all Mather measures  $\tilde{\mu}$ .

By Theorem 7.1, the Mather set is not empty. The Aubry set  $\tilde{\mathcal{A}}(u)$  depends on the choice of the weak KAM solution. The way to make it independent of choice is the following definition.

**Definition 7.3** (Aubry set). The Aubry set  $\tilde{\mathcal{A}}$  of the Lagrangian  $L$  (resp.  $\tilde{\mathcal{A}}^*$  of the Hamiltonian  $H$ ) is  $\bigcap_u \tilde{\mathcal{I}}(u)$  (resp.  $\bigcap_u \tilde{\mathcal{I}}^*(u)$ ), where the intersection is taken over all weak KAM solutions  $u : M \rightarrow \mathbb{R}$ .

Note that we use here the notation  $\tilde{\mathcal{A}}^*$  instead of the notation  $\tilde{\mathcal{A}}^*(0)$  used in the Introduction §1.

**Corollary 7.4.** *The Aubry sets  $\tilde{\mathcal{A}}$  and  $\tilde{\mathcal{A}}^*$  are not empty. In fact, we have  $\tilde{\mathcal{M}} \subset \tilde{\mathcal{A}}$ , and  $\mathcal{L}(\tilde{\mathcal{A}}) = \tilde{\mathcal{A}}^*$ . Both the Mather set and the Aubry sets are graphs on the base  $M$ , since  $\tilde{\mathcal{A}}^* \subset \text{Graph}(du)$ , for any weak KAM solution  $u : M \rightarrow \mathbb{R}$ .*

The results obtained in this section finishes the proof of Theorem 1.5 for the case  $P = 0$ . As explained in the introduction, the case for a general  $P \in \mathbb{R}^k$  follows from this one.

## 8. Proof of Fleming's lemma

It will be helpful to consider the energy  $E : TM \rightarrow \mathbb{R}$ , defined by  $E = H \circ \mathcal{L}$ . Since  $H$  is superlinear, and  $\mathcal{L}$  is a homeomorphism, for every  $K \in \mathbb{R}$ , the set  $\{(x, v) \mid E(x, v) \leq K\}$  is compact. Moreover, since  $\mathcal{L}$  conjugates the Lagrangian flow  $\phi_t$  to the Hamiltonian  $\phi_t^*$ , the energy is constant along speed curves of extremals.

**Lemma 8.1.** *Given  $t_0 > 0$ , there exists a finite constant  $\kappa_{t_0}$ , such that every minimizer  $\gamma : [a, b] \rightarrow M$ , with  $b - a \geq t_0$ , satisfies  $\|\dot{\gamma}(s)\|_{\gamma(s)} \leq \kappa_{t_0}$ , for every  $s \in [a, b]$ .*

*Proof.* Call  $\delta : [a, b] \rightarrow M$  a geodesic, parametrized proportionally to arc-length, with  $\delta(a) = \gamma(a)$ ,  $\delta(b) = \gamma(b)$ , and whose length is  $d(\gamma(a), \gamma(b))$ . The speed  $\|\dot{\delta}(s)\|_{\delta(s)}$  of the geodesic is constant for  $s \in [a, b]$ . The length of  $\gamma$  is therefore  $(b - a)\|\dot{\delta}(s)\|_{\delta(s)}$ , for any  $s \in [a, b]$ . This implies

$$(b - a)\|\dot{\delta}(s)\|_{\delta(s)} = d(\gamma(a), \gamma(b)) \leq \text{diam}(M).$$

Hence  $\|\dot{\delta}(s)\|_{\delta(s)} \leq \text{diam}(M)/t_0$ . If we set  $C_{t_0}^1 = \sup\{L(x, v) \mid \|v\|_x \leq \text{diam}(M)/t_0\}$ , we see that the action of  $\delta$  is bounded by  $(b - a)C_{t_0}^1$ . Since  $\gamma$  is a minimizer, we obtain  $\int_a^b L(\gamma(s), \dot{\gamma}(s)) ds \leq (b - a)C_{t_0}^1$ . This implies that there exists  $s_0 \in [a, b]$  such that  $L(\gamma(s_0), \dot{\gamma}(s_0)) \leq C_{t_0}^1$ . By the superlinearity of  $L$ , the constant

$$C_{t_0}^2 = \sup\{\|v\|_x \mid L(x, v) \leq C_{t_0}^1\}$$

is finite. Since the energy is constant along the speed curve of an extremal, we get

$$E(\gamma(s), \dot{\gamma}(s)) = E(\gamma(s_0), \dot{\gamma}(s_0)) \leq C_{t_0}^3,$$

where  $C_{t_0}^3 = \sup\{E(x, v) \mid \|v\|_x \leq C_{t_0}^2\}$ . Hence, for every  $s \in [a, b]$ , we have

$$\|\dot{\gamma}(s_0)\|_{\gamma(s_0)} \leq \kappa_{t_0},$$

where  $\kappa_{t_0} = \sup\{\|v\|_x \mid E(x, v) \leq C_{t_0}^3\}$ . □

**Lemma 8.2.** *If  $t_0 > 0$ , and a finite  $\beta \geq 1$  are given, we can find a constant  $K_{t_0, \beta}$  such that*

$$|h_t(x, y) - h_{t'}(x, y)| \leq K_{t_0, \beta}|t - t'|,$$

for every  $x, y \in M$ , and  $t, t' \geq t_0$ , with  $\max(t/t', t'/t) \leq \beta$ .

*Proof.* By Tonelli's theorem, we can find a minimizer  $\gamma : [0, t'] \rightarrow M$ , with  $\gamma(0) = x$ , and  $\gamma(t') = y$ . Since  $\gamma$  is a minimizer

$$h_{t'}(x, y) = \int_0^{t'} L(\gamma(s), \dot{\gamma}(s)) ds.$$

Note that by Lemma 8.1 above we have  $\|\dot{\gamma}(s)\|_{\gamma(s)} \leq \kappa_{t_0}$ , for every  $s \in [0, t']$ . If we define  $\tilde{\gamma} : [0, t] \rightarrow M$  by  $\tilde{\gamma}(s) = \gamma(t't^{-1}s)$ . Since  $\tilde{\gamma}(0) = x$ , and  $\tilde{\gamma}(t) = y$ , we get

$$\begin{aligned} h_t(x, y) &\leq \int_0^t L(\tilde{\gamma}(s), \dot{\tilde{\gamma}}(s)) ds \\ &\leq \int_0^t L(\gamma(t't^{-1}s), t't^{-1}\dot{\gamma}(t't^{-1}s)) ds \\ &= \int_0^{t'} L(\gamma(s'), t't^{-1}\dot{\gamma}(s')) t't^{-1} ds'. \end{aligned}$$

where the last line was obtained by the change of variable  $s' = t't^{-1}s$ . Therefore, we have

$$h_t(x, y) - h_{t'}(x, y) \leq \int_0^{t'} L(\gamma(s), t't^{-1}\dot{\gamma}(s))tt'^{-1} - L(\gamma(s), \dot{\gamma}(s)) ds. \quad (8.1)$$

Since  $L$  is at least  $C^1$ , we can find a Lipschitz constant  $K_{t_0, \beta}$  of the map  $(x, v, \alpha) \mapsto L(x, \alpha^{-1}v)\alpha$ , on the compact set  $\{(x, v, \alpha) \mid (x, v) \in TM, \|v\|_x \leq \kappa_{t_0}, \beta^{-1} \leq \alpha \leq \beta\}$ . This fact together with inequality (8.1) yield

$$h_t(x, y) - h_{t'}(x, y) \leq \int_0^{t'} K_{t_0, \beta} |tt'^{-1} - 1| ds = K_{t_0, \beta} |t - t'|.$$

By symmetry this finishes the proof.  $\square$

*Proof of Fleming's lemma 3.5.* Assume  $t \geq t_0$ . By part 3) and 4) of Lemma 3.4, we have

$$\begin{aligned} h_{t+d(x, x') + d(y, y')}(x', y') &\leq h_{d(x', x)}(x', x) + h_t(x, y) + h_{d(y, y')}(y, y') \\ &\leq h_t(x, y) + B(d(x', x) + d(y, y')). \end{aligned} \quad (8.2)$$

If we set  $t' = t + d(x, x') + d(y, y')$ , we have  $t, t' \geq t_0$ ,  $t/t' \leq 1$ , and  $t'/t = 1 + (d(x', x) + d(y, y'))/t \leq 1 + 2 \operatorname{diam}(M)t_0^{-1}$ . By Lemma 8.2 with  $\beta = 1 + 2 \operatorname{diam}(M)t_0^{-1}$ , we get

$$h_t(x', y') \leq h_{t+d(x, x') + d(y, y')}(x', y') + K_{t_0, \beta}(d(x', x) + d(y, y')).$$

Combining with the inequality (8.2), we obtain

$$h_t(x', y') - h_t(x, y) \leq (B + K_{t_0, \beta})(d(x', x) + d(y, y')).$$

By symmetry this finishes the proof of Fleming's Lemma.  $\square$

## 9. A $C^1$ solution of the Hamilton-Jacobi equation is a weak KAM solution

In this section, we assume that  $u : M \rightarrow \mathbb{R}$  is a  $C^1$  solution of the Hamilton-Jacobi equation  $H(x, d_x u) = c$  (for every  $x \in M$ ). We first prove that  $u \prec L + c$ . Assume  $\gamma : [a, b] \rightarrow M$  is a  $C^1$  curve, together with the Hamilton-Jacobi equation, the Fenchel inequality gives  $d_{\gamma(s)} u(\dot{\gamma}(s)) \leq L(\gamma(s), \dot{\gamma}(s)) + H(\gamma(s), d_{\gamma(s)} u) = L(\gamma(s), \dot{\gamma}(s)) + c$ . Integrating on  $[a, b]$  yields  $u(\gamma(b)) - u(\gamma(a)) \leq \mathbb{L}(\gamma) + c(b - a)$ . Therefore, by Proposition 5.1, we have  $u \leq T_t^- u + ct$ , for every  $t \geq 0$ . To show the opposite inequality, we prove the following lemma.

**Lemma 9.1.** *For every  $x \in M$ , there exists a curve  $\gamma_x : ]-\infty, +\infty[ \rightarrow M$ , which is  $(u, L, c)$ -calibrated.*

*Proof.* Since the Legendre transform  $\mathcal{L}$  is a homeomorphism, we can define a continuous vector field  $X_u$  on  $M$  by  $X_u(x) = \partial H / \partial p(x, d_x u)$ . By the equality case in Fenchel equality, and the fact that  $H(x, d_x u) = c$ , we have

$$d_x u(X_u(x)) = L(x, X_u(x)) + c, \text{ for every } x \in M. \quad (9.1)$$

Since  $X_u$  is continuous, we can apply the Cauchy-Peano theorem to find a solution  $\gamma_x$  of  $X_u$  with  $\gamma_x(0) = X$ . Moreover, by compactness of  $M$ , we can assume that the solution  $\gamma_x$  of  $X_u$  is defined on the whole of  $\mathbb{R}$ . Using (9.1) along  $\gamma_x$ , we obtain

$$d_{\gamma_x(s)}u(\dot{\gamma}_x(s)) = L(\gamma_x(s), \dot{\gamma}_x(s)) + c, \text{ for every } x \in M.$$

It remains now to integrate this equality on an arbitrary compact interval  $[a, b]$  to see that  $\gamma_x$  is  $(u, L, c)$ -calibrated on  $\mathbb{R}$ .  $\square$

We now show that  $u \geq T_t^- u + ct$ , for every  $t \geq 0$ . Fix  $x \in M$ , and pick  $\gamma_x$  given by Lemma 9.1. For  $t > 0$ , we have

$$u(x) - u(\gamma_x(-t)) = \int_{-t}^0 L(\gamma_x(s), \dot{\gamma}_x(s)) ds + ct \geq h_t(\gamma_x(-t), x) + ct.$$

Therefore  $u(x) \geq u(\gamma_x(-t)) + h_t(\gamma_x(-t), x) + ct \geq T_t^- u(x) + ct$ .

**Acknowledgements.** The author benefited from support from Institut Universitaire de France, and from ANR ANR-07-BLAN-0361-02 KAM & ANR-12-BS01-0020 WKBHJ. He also enjoyed the hospitality of both the Korean IBS Center For Geometry and Physics in Pohang, and CIMAT in Guanajuato, while writing this paper.

## References

- [1] Arnold, V.I., *Proof of a theorem of A. N. Kolmogorov on the preservation of conditionally periodic motions under a small perturbation of the Hamiltonian* (Russian), *Uspehi Mat. Nauk* **18** (1963), 13–40.
- [2] Aubry, S., *The devil's staircase transformation in incommensurate lattices*, in *The Riemann problem, complete integrability and arithmetic applications*, Bures-sur-Yvette/New York, 1979/1980. *Lecture Notes in Math.* 925, Springer, Berlin-New York, 1982. 221–245.
- [3] Bardi, M. and Cappuzzo-Dolceta, I., *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations*, *Systems & Control: Foundations and Applications*, Birkhäuser Boston Inc., Boston, MA, 1997.
- [4] Barles, G., *Solutions de viscosité des équations de Hamilton-Jacobi*, *Mathématiques et Applications* 17, Springer-Verlag, Paris, 1994.
- [5] Buttazzo, G., Giaquinta, M., and Hildebrandt, S., *One-dimensional variational problems*, *Oxford Lecture Series in Mathematics and its Applications* 15, Oxford Univ. Press, New York, 1998.
- [6] Cannarsa, P. and Sinestrari, C., *Semiconcave Functions, Hamilton-Jacobi Equations, and Optimal Control*, *Progress in Nonlinear Differential Equations and Their Applications* 58, Birkhäuser, Boston, 2004.
- [7] Clarke, F.H., *Methods of dynamic and nonsmooth optimization*, *CBMS-NSF Reg. Conf. Series in App. Math.* 57, SIAM, Philadelphia, PA, 1989.



- [8] Crandall, M.G. and Lions, P.L., *Viscosity solutions of Hamilton-Jacobi equations*, Trans. Amer. Math. Soc. **277**(1983), 1–42.
- [9] Dias Carneiro, M.J., *On minimizing measures of the action of autonomous Lagrangians*, Nonlinearity **8** (1995), 1077–1085.
- [10] Dugundji, J., *Topology*, Allyn and Bacon Inc., Boston, 1970.
- [11] E. Weinan, *Aubry-Mather theory and periodic solutions of the forced Burgers equation*, Comm. Pure Appl. Math. **52** (1999), 811–828.
- [12] Evans, L. C. and Gomes, D. *Effective Hamiltonians and averaging for Hamiltonian dynamics. I-II*, Arch. Ration. Mech. Anal. **157** (2001), 1–33 & **161** (2002), 271–305.
- [13] Fathi, A., *Théorème KAM faible et théorie de Mather sur les systèmes lagrangiens*, C. R. Acad. Sci. Paris Sér. I Math. **324**(9) (1997), 1043–1046.
- [14] ———, *Weak KAM Theorem and Lagrangian Dynamics*, Preprint, <http://tinyurl.com/AFnotesWeakKam>, <http://tinyurl.com/oh4qjv3>.
- [15] Fathi, A. and A. Figalli, *Optimal transportation on non-compact manifolds*, Israel J. Math. **175** (2010), 1–59.
- [16] Fleming, W.H., *The Cauchy problem for a nonlinear first order partial differential equation*, J. Differential Equations **5** (1969), 515–530.
- [17] Kolmogorov, A.N., *Théorie générale des systèmes dynamiques et mécanique classique*, in Proceedings of the International Congress of Mathematicians, Amsterdam, 1954, Vol. 1 315–333. Erven P. Noordhoff N.V., Groningen; North-Holland Publishing Co 1957.
- [18] Lions, P.-L., Papanicolau, G., and Varadhan, S.R.S., *Homogenization of Hamilton-Jacobi equation*, Unpublished preprint, (1987).
- [19] Mañé, R., *Lagrangian flows: the dynamics of globally minimizing orbits*, Bol. Soc. Brasil. Mat. (N.S.) **28** (1997), 141–153.
- [20] Mather, J.N., *Existence of quasiperiodic orbits for twist homeomorphisms of the annulus*, Topology **21** (1982), 457–467.
- [21] ———, *Action minimizing measures for positive definite Lagrangian systems*, Math. Z. **207** (1991), 169–207.
- [22] ———, *Variational construction of connecting orbits*, Ann. Inst. Fourier **43** (1993), 1349–1386.
- [23] Moser, J., *A new technique for the construction of solutions of nonlinear differential equations*, Proc. Nat. Acad. Sci. U.S.A. **47** (1961), 1824–1831.
- [24] Poincaré, H., *Les Méthodes nouvelles de la Mécanique Céleste*, 3 tomes, Paris, Gauthier-Villars, 1892–1899.

46 allée d'Italie, 69364 Lyon, Cedex 07, France

E-mail: albert.fathi@ens-lyon.fr