ANNALES DE LA FACULTÉ DES SCIENCES DE TOULOUSE

ALBERT FATHI

Regularity of C^1 solutions of the Hamilton-Jacobi equation

Annales de la faculté des sciences de Toulouse 6e série, tome 12, nº 4 (2003), p. 479-516

<http://www.numdam.org/item?id=AFST_2003_6_12_4_479_0>

Annales de la Faculté des Sciences de Toulouse
Vol. XII, n° 4, 2003
pp. 479–516

Regularity of C^1 solutions of the Hamilton-Jacobi equation [(*)]

ALBERT FATHI [(1)]

ABSTRACT. — We study C^1 solutions of the Hamilton-Jacobi equation. We will show that they are necessarily $C^{1,1}$, generalizing some work of Pierre-Louis Lions. We will also prove some compactness theorems, and give some geometrical consequences.

RÉSUMÉ. — Nous étudions les solutions C^1 de l'équation de Hamilton-Jacobi. Nous montrons qu'elles sont nécessairement $C^{1,1}$, généralisant ainsi un résultat de Pierre-Louis Lions. Nous démontrons aussi certains théorèmes de compacité et nous en tirons des conséquences géométriques.

1. Introduction

To state our results we consider a Hamiltonian $H : \mathbb{R} \times \mathbb{R}^k \times \mathbb{R}^k \to \mathbb{R}$, $(t, q, p) \mapsto H(t, q, p)$. We assume the following conditions on H:

(C1) the Hamiltonian H is C^r with $r \geqslant 2$,

(C2) The Hamiltonian H is C^2-strictly convex in the fibers, i.e. for each $t \in \mathbb{R}$ and each $q \in M$, the second derivative of the restriction $H|\{t\} \times \{q\} \times \mathbb{R}^k$ is positive definite as a quadratic form. This can be written as

$$\forall (t, q, p) \in \mathbb{R} \times \mathbb{R}^k \times \mathbb{R}^k, \forall u \in \mathbb{R}^k \setminus \{0\}, \frac{\partial^2 H}{\partial p^2}(t, q, p)(u, u) > 0;$$

(*) Reçu le 20 juin 2003, accepté le 20 octobre 2003
(1) ENS Lyon & CNRS UMR 5669
UMPA, Ecole Normale Supérieure de Lyon, 46, allée d'Italie, 69364 Lyon Cedex 07 France.
Email: afathi@umpa.ens-lyon.fr
The author acknowledges support from the Swiss National Scientific Foundation during his stay at Geneva University in Spring 2000.

(C3) The Hamiltonian H is superlinear, i.e.

$$\forall t \in \mathbb{R}, \forall q \in M, \lim_{\substack{\|v\| \to \infty \\ v \in T_q M}} \frac{H(t,q,p)}{\|p\|} = \infty.$$

We study C^1 Solutions of the Hamilton-Jacobi Equation. Such a solution is a C^1 map $u : U \to \mathbb{R}$, where U is an open subset of $\mathbb{R} \times \mathbb{R}^k$ and

$$\frac{\partial u}{\partial t}(t,q) + H(t,q,\frac{\partial u}{\partial q}(t,q)) = 0.$$

THEOREM 1.1. — *If $u : U \to \mathbb{R}$ is a C^1 solution of the Hamilton-Jacobi equation then u is $C^{1,1}$, i.e. the derivative of u is locally Lipschitzian.*

This generalizes some work of Pierre-Louis Lions, see [Li, Theorem 15.1] and the comments after Theorem 3.3 below.

We will also prove some compactness theorems that do not seem to have been noticed in that full generality. For example the following theorem seems to be new:

THEOREM 1.2. — *Suppose that H is time-independent and that u_n is a sequence of C^1 solutions of the Hamilton-Jacobi equation all defined on the same open subset U of $\mathbb{R} \times \mathbb{R}^k$. If the pointwise limit $u(q)$ of the sequence $u_n(q)$ exists at every point of U, then u itself is a C^1 solution of the Hamilton-Jacobi equation, and u_n converges to u in the (compact open) C^1 topology.*

In fact, under some fairly general conditions a family of C^1 solutions of the Hamilton-Jacobi equation, all defined on the same open subset U, and bounded on compact subsets of U, have derivatives which are equi-Lipschitzian on compact subsets of U, see the precise formulation given in §3 and §4 below.

Some versions of these compactness theorems are known to Differential Geometers or to people working in Lagrangian Dynamics, see the comments below following Corollary 3.3. We give for example an application to the smoothness of Busemann's functions.

Finally we show that the construction of solutions of the Hamilton-Jacobi equation using the method of characteristics can be done when and only when we start with a $C^{1,1}$ function, compare with [Li, Remark 1.1 pages 15–16, and Proposition 15.1 page 265]. These methods can be used to show the following fact, if N is a C^1 codimension 1 submanifold of the Riemannian manifold M, the distance function $q \mapsto d(q, N)$ is C^1 on a set $V \setminus N$, where V is a neighborhood of N in M, if and only if N is $C^{1,1}$. This is to be contrasted with the case of higher differentiability, see the work of Poly and Raby [PR].

It is crucial for the proofs that we can use the dual formalism of the Lagrangian, because action and minimizers will play a fundamental role. The fact that a Hamiltonian satisfying (C1), (C2), and (C3) comes via Legendre-Fenchel duality from a Lagrangian satisfying the same assumptions is well-known. We prefer to state and do things in the invariant framework of manifolds. We recall this framework in §2 below.

A first (awkward) version of this work was presented at the "International Conference in Dynamical Systems, a Tribute to Ricardo Mañé", held from March 27 till April 1, 1995 in Montevideo, Uruguay. The final version was written during a sabbatical semester at the Institut de Mathématiques de l'Université de Genève in Switzerland in Spring 2000.

2. Convex and superlinear Lagrangians

We will consider a fixed smooth manifold M. We will denote a point in its tangent space TM by (q, v) with $q \in M$ and $v \in T_qM$, the space of vectors tangent to M at the point q. The map $\pi : TM \to M$ is the canonical projection $(q, v) \mapsto q$. We will denote a point in the cotangent space T^*M by (q, p) with $p \in T_q^*M$, so p is a linear form on the vector space T_qM. The map $\pi^* : TM^* \to M$ is the canonical projection $(q, p) \mapsto q$.

It is convenient to introduce the Liouville form α on T^*M. This is a differential 1-form. Thus for a given $(q, p) \in T^*M$ we must define a linear map $\alpha_{(x,p)} : T_{(q,p)}(T^*M) \to \mathbb{R}$. To define it let us notice that if $((q, p), (Q, P)) \in T_{(q,p)}T^*M$, then $T\pi^*((q, p), (Q, P)) = (q, Q) \in TM$, and thus the expression $p(Q)$ makes sense, since it is the evaluation of the linear form $p : T_xM \to \mathbb{R}$ on the vector $Q \in T_xM$. We then set

$$\alpha_{(x,p)}(Q, P) = p(Q).$$

It is not difficult to see that this defines a differential 1-form α on the T^*M manifold. If U is an open subset of M and $\theta : U \to \mathbb{R}^k$ is a coordinate chart, we can consider the associated coordinate chart $T^*\theta : T^*U \to \mathbb{R}^k \times \mathbb{R}^k$ on T^*M. If we denote by $(q_1, \cdots, q_k, p_1, \cdots, p_k)$ the canonical coordinates on $\mathbb{R}^k \times \mathbb{R}^k$, we see that

$$\alpha | T^*U = \sum_{i=1}^{k} p_i \, dq_i.$$

The canonical symplectic form on T^*M is $\Omega = -d\alpha$. In the coordinate charts of the type $T^*\theta$ introduced above we have

$$\Omega = \sum_{i=1}^{k} dq_i \wedge dp_i.$$

It will also be convenient to have some fixed Riemannian metric on M. For a tangent vector $(q, v) \in TM$, we will denote the length of the tangent vector for the Riemannian by $\|v\|$.

We will consider a Lagrangian $L : \mathbb{R} \times TM \to \mathbb{R}$ that satisfies the following three conditions:

(C1) the Lagrangian L is C^r with $r \geqslant 2$;

(C2) the Lagrangian L is C^2-strictly convex in the fibers, i.e. for each $t \in \mathbb{R}$ and each $q \in M$, the second derivative of the restriction $L|t \times T_qM$ is positive definite as a quadratic form. This can be written as

$$\forall (t, q, v) \in \mathbb{R} \times TM, \forall u \in T_qM \setminus \{0\}, \frac{\partial^2 L}{\partial v^2}(t, q, v)(u, u) > 0;$$

(C3) the Lagrangian L is superlinear, i.e.

$$\forall t \in \mathbb{R}, \forall q \in M, \lim_{\substack{\|v\| \to \infty \\ v \in T_qM}} \frac{L(t, q, v)}{\|v\|} = \infty.$$

The derivative of $L|t \times T_qM$ at the point (q, v), denoted by $\frac{\partial L}{\partial v}(t, q, v)$, is an element of the dual vector space T_q^*M. We have

$$\forall t \in \mathbb{R}, \forall q \in M, \forall v \in T_qM, \frac{\partial L}{\partial v}(t, q, v)(v) \geqslant L(t, q, v) - L(t, q, 0). \qquad (2.1)$$

As is well-known this follows from the convexity assumption (C2). In fact, the C^2 function $\varphi : \mathbb{R} \to \mathbb{R}$ defined by $\varphi(s) = L(t, q, sv)$ has a non-negative second derivative, hence its first derivative is non-decreasing. By the mean value theorem $L(t, q, v) - L(t, q, 0) = \varphi(1) - \varphi(0) = \varphi'(c)$ for some $c \in]0, 1[$. Since φ' is non-decreasing, we get $\varphi(1) - \varphi(0) \leqslant \varphi'(1) = \frac{\partial L}{\partial v}(t, q, v)(v)$. Using

$$\|\frac{\partial L}{\partial v}(t, q, v)\| \geqslant \frac{\partial L}{\partial v}(t, q, v)(\frac{v}{\|v\|}),$$

from (2.1) and the superlinearity condition (C3), it follows that

$$\forall t \in \mathbb{R}, \forall q \in M, \lim_{\substack{\|v\| \to \infty \\ v \in T_q M}} \|\frac{\partial L}{\partial v}(t, q, v)\| = \infty.$$

This shows that the Legendre transform $\mathcal{L}_{(t,q)} : T_q M \to T_q^* M, v \mapsto \dfrac{\partial L}{\partial v}(t, q, v)$ is a proper map, i.e. the inverse image by $\mathcal{L}_{(t,q)}$ of a compact subset is itself compact. By the convexity condition (C2), the derivative of the Legendre transform $\mathcal{L}_{(t,q)}$ is non-degenerate at every $v \in T_q M$. It follows that $\mathcal{L}_{(t,q)}$ is a C^{r-1} surjective diffeomorphism. Hence the global Legendre transform $\mathcal{L} : \mathbb{R} \times TM \to \mathbb{R} \times T^* M, (t, q, v) \mapsto (t, q, \mathcal{L}_{(t,q)}(v))$ is also a C^{r-1} surjective diffeomorphism. For $t \in \mathbb{R}$, it is convenient to define the diffeomorphism $\mathcal{L}_t : TM \to T^* M, (q, v) \mapsto (q, \mathcal{L}_{(t,q)}(v))$.

The Hamiltonian $H : \mathbb{R} \times T^* M \to \mathbb{R}$ is defined by:

$$H(t, q, p) = \sup\{p(v) - L(t, q, v) \mid v \in T_q M\}.$$

This supremum is in fact attained precisely at the unique vector v such that $p = \dfrac{\partial L}{\partial v}(t, q, v)$. The function H is obviously C^{r-1} because it can be also defined as $H(t, q, p) = p(\mathcal{L}^{-1}_{(t,q)}(p)) - L(t, q, \mathcal{L}^{-1}_{(t,q)}(p))$. A closer look at the derivatives shows that H is in fact C^r.

From the definition of H it follows that

$$\forall t \in \mathbb{R}, q \in M, v \in T_q M, p \in T_q^* M, p(v) \leqslant H(t, q, p) + L(t, q, v). \tag{2.2}$$

Moreover

$$p(v) = H(t, q, p) + L(t, q, v) \text{ if and only if } p = \frac{\partial L}{\partial v}(t, q, v). \tag{2.3}$$

It is usual to define the energy $E : \mathbb{R} \times TM \to \mathbb{R}$ associated to the Lagrangian as $E = H \circ \mathcal{L}$, so we have

$$E(t, q, v) = H(t, q, \frac{\partial L}{\partial v}(t, q, v)) = \frac{\partial L}{\partial v}(t, q, v)(v) - L(t, q, v)).$$

The following lemma deserves to be better known, compare with [Be, §4, page 61]:

LEMMA 2.1.— *For each compact subset K of $\mathbb{R} \times M$ and each constant $A \in \mathbb{R}$, there exists a finite constant $C(A, K) < +\infty$ such that*

$$\forall (t, q) \in K, \forall v \in T_{(t,q)} M, \; A\|v\| \leqslant L(t, q, v) + C(A, K),$$

$$\forall (t,q) \in K, \forall p \in T^*_{(t,q)}M,\ A\|p\| \leqslant H(t,q,p) + C(A,K).$$

Proof.— Since for $v \in T_{(t,q)}M$ we have

$$A\|v\| = \sup_{\substack{p \in T^*_{(t,q)}M \\ \|p\|=A}} p(v),$$

we obtain from (2.2)

$$\forall (t,q) \in K, \forall v \in T_{(t,q)},\ A\|v\| \leqslant L(t,q,v) + C_1(A,K),$$

where $C_1(A,K) = \sup\{H(t,q,p) \mid (t,q) \in K, p \in T^*_{(t,q)}M, \|p\| = A\}$ which is finite since the set $\{(t,q,p) \mid (t,q) \in K, p \in T^*_{(t,q)}M, \|p\| = A\}$ is compact.

A dual argument proves the second inequality. □

Let us recall that a C^1 path $\gamma : [a,b] \to M$ is an extremal for the Lagrangian L, if γ is a critical point of the map $\delta \mapsto \int_a^b L(s,\delta(s),\dot\delta(s))\,ds$ defined on the space of continuous piecewise C^1 paths $\delta : [a,b] \to M$ with $\delta(a) = \gamma(a)$ and $\delta(b) = \gamma(b)$. Such an extremal γ is called a minimizer if

$$\int_a^b L(s,\gamma(s),\dot\gamma(s))\,ds \leqslant \int_a^b L(s,\delta(s),\dot\delta(s))\,ds$$

for every continuous piecewise C^1 path $\delta : [a,b] \to M$ with $\delta(a) = \gamma(a)$ and $\delta(b) = \gamma(b)$. We will say that $\gamma : [a,b] \to M$ is a local minimizer if there is an open subset V of M such that $\gamma([a,b]) \subset V$ and

$$\int_a^b L(s,\gamma(s),\dot\gamma(s))\,ds \leqslant \int_a^b L(s,\delta(s),\dot\delta(s))\,ds$$

for every continuous piecewise C^1 path $\delta : [a,b] \to V$ with $\delta(a) = \gamma(a)$ and $\delta(b) = \gamma(b)$.

It is well known that there exists a C^{r-1} time dependent vector field $X_L : \mathbb{R} \times TM \to TTM$, (with $X_L(t,q,v) \in T_{(q,v)}TM$) such that a C^1 path $\gamma : [a,b] \to M$ is an extremal of L if and only if $t \mapsto (t,\gamma(t),\dot\gamma(t)) \in \mathbb{R} \times TM$ is a solution of X_L. This vector field X_L is called the Euler-Lagrange vector field of the Lagrangian L. The (time-dependent) flow of X_L is defined on an open subset $\mathcal{U}_L$ of $\mathbb{R} \times \mathbb{R} \times TM$ containing $\Delta_{\mathbb{R}} \times TM$, with $\Delta_{\mathbb{R}}$ the diagonal of $\mathbb{R} \times \mathbb{R}$. We will denote by $\phi : \mathcal{U}_L \to TM$ this flow, it is of class C^{r-1} and it is called the Euler-Lagrange flow of the Lagrangian. If

$(t, q, v) \in \mathbb{R} \times TM$, then $s \mapsto \phi(s, t, q, v)$ is a maximal solution of X_L with $\phi(t, t, q, v) = (q, v)$. We will also write $\phi_s(t, q, v)$ instead of $\phi(s, t, q, v)$, to emphasize the particular role of the variable s. It is convenient to define $\mathcal{F} : \mathcal{U}_L \to M$ as $\mathcal{F} = \pi \circ \phi$, where π is the projection $TM \to M$. If $(t, q, v) \in \mathbb{R} \times TM$, we have $\mathcal{F}(t, t, q, v) = q$, $\frac{d\mathcal{F}}{ds}(t, t, q, v) = v$ and $s \mapsto \mathcal{F}(s, t, q, v)$ is an extremal of L. Moreover, if $\gamma : [a, b] \to M$ is an extremal of L then $\mathcal{F}(s, t, \gamma(t), \dot\gamma(t))$ is defined for all $s, t \in [a, b]$ and $\mathcal{F}(s, t, \gamma(t), \dot\gamma(t)) = \gamma(s)$ for all $s, t \in [a, b]$.

It is also useful to consider the time-dependent flow ϕ_s^* defined on $\mathbb{R} \times T^*M$ by

$$\phi_s^*(t, q, p) = \mathcal{L}_{t+s} \circ \phi_s \circ \mathcal{L}^{-1}(t, q, p). \tag{2.4}$$

As is well-known this flow is in fact the Hamiltonian flow associated to H, which means that it is the flow of the (time-dependent) vector field X_H^* on $\mathbb{R} \times T^*M$ defined by $d_{(q,v)}H = \Omega(X_H^*, \cdot)$, with $\Omega = -d\alpha$ the canonical symplectic form on T^*M. In a coordinate chart $T^*\theta : T^*U \to \mathbb{R}^k$, where $\theta : U \to \mathbb{R}^k$ is a coordinate chart on M, and $(q_1, \cdots, q_k, p_1, \cdots, p_k)$ are the canonical coordinates on $\mathbb{R}^k \times \mathbb{R}^k$, we have

$$X_H^*(t, q, p) = \sum_{i=1}^{k} \frac{\partial H}{\partial p_i}(t, q, p) \frac{\partial}{\partial q_i} + \sum_{i=1}^{k} -\frac{\partial H}{\partial q_i}(t, q, p) \frac{\partial}{\partial p_i}.$$

If L (or equivalently H) does not depend on time then ϕ_s^* preserves the hamiltonian H and hence ϕ_s preserves the energy E.

3. Solutions of the Hamilton-Jacobi equation

Let U be an open subset of $\mathbb{R} \times TM$. If $u : U \to \mathbb{R}$ is a C^1 function, for a given $(t, q) \in U$, we will denote by $\frac{\partial u}{\partial q}(t, q)$ the partial derivative with respect to the second argument, i.e. the derivative at q of the map $x \mapsto u(t, x)$ which is defined in a neighborhood of q. We think of $\frac{\partial u}{\partial q}(t, q)$ as a linear map $T_qM \to \mathbb{R}$. so $\frac{\partial u}{\partial q}(t, q) \in T_q^*M$. Of course, the derivative $d_{(t,q)}u$ of u at (t, q) is $\frac{\partial u}{\partial t}(t, q)dt + \frac{\partial u}{\partial q}(t, q)$.

We say that such a C^1 function $u : U \to \mathbb{R}$ is $C^{1,1}$ if its derivative $du : U \to T^*(\mathbb{R} \times M)$ is locally Lipschitzian.

Let us recall that the notion of locally Lipschitzian for a map θ from a (finite-dimensional) C^1 manifold N to a (finite-dimensional) C^1 manifold P makes perfect sense. It is simply a map that is locally Lipschitzian when looked at in charts. If M and N are, for example, C^∞ and endowed with smooth Riemannian metrics, then f is locally Lipschitzian if and only if it is locally Lipschitzian with respect to the distances on N and P obtained from the Riemannian metrics.

We will prove the following theorem:

THEOREM 3.1.— *Let U be an open subset of $\mathbb{R} \times M$, and $u : U \to \mathbb{R}$ be a C^1 solution of the Hamilton-Jacobi equation*

$$\frac{\partial u}{\partial t}(t,q) + H(t,q,\frac{\partial u}{\partial q}(t,q)) = 0,$$

then u is $C^{1,1}$.

The version where $M = \mathbb{R}^n$ and $L(t,x,v)$ is a function of v only on $\mathbb{R} \times T\mathbb{R}^n = \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n$ is due to Pierre-Louis Lions, see [Li, Theorem 15.1 page 255]. Lions also observes that it is possible to give a version of his theorem for more general Lagrangians, see [Li, remark 15.2 page 256]. For other particular cases, see the comments following Corollary 3.3.

In Theorem 3.1, we can in fact control the local Lipschitz constant of the derivative du of u. We now explain what this means. We will suppose that $\mathbb{R} \times M$ and $T^*(\mathbb{R} \times M)$ are endowed with distances, both denoted by D, coming from Riemannian metrics and we will set

$$\mathrm{Lip}_{(t,q)}(du) = \lim_{\epsilon\to 0} \sup\{\frac{D[d_{(t_1,q_1)}u, d_{(t_2,q_2)}u]}{D[(t_1,q_1),(t_2,q_2)]} \mid D((t_i,q_i),(t,q)) \leqslant \epsilon, i = 1,2\}.$$

It is clear that the following theorem is stronger than Theorem 3.1.

THEOREM 3.2.— *Let U be an open subset of $\mathbb{R} \times M$. For every pair of compact subsets $K, K' \subset U$, with K contained in the interior of K' and every constant $A \geqslant 0$, we can find a finite constant B such that if $u : U \to \mathbb{R}$ is a C^1 solution of the Hamilton-Jacobi equation*

$$\frac{\partial u}{\partial t}(t,q) + H(t,q,\frac{\partial u}{\partial q}(t,q)) = 0,$$

with $\sup\{\|\frac{\partial u}{\partial q}(t,q)\| \mid (t,q) \in K'\} \leqslant A$, *then*

$$\sup\{\mathrm{Lip}_{(t,q)}(du) \mid (t,q) \in K\} \leqslant B.$$

Using Ascoli's theorem, we obtain from this theorem the following corollary.

COROLLARY 3.3.— *Let U be an open subset of $\mathbb{R} \times M$. Call $\mathcal{S}^1(U,\mathbb{R})$ the set of C^1 maps $u : U \to \mathbb{R}$ which satisfy the Hamilton-Jacobi equation*

$$\frac{\partial u}{\partial t}(t,q) + H(t,q,\frac{\partial u}{\partial q}(t,q)) = 0.$$

Each subset of $\mathcal{S}^1(U,\mathbb{R})$ which is bounded in the compact-open C^1 topology on $\mathcal{C}^1(M,\mathbb{R})$ is a relatively compact subset $\mathcal{S}^1(U,\mathbb{R})$ (endowed with the compact-open C^1 topology).

Theorem 3.2 and its Corollary 3.3 do not seem to be known as such. Particular cases of Theorems 3.1, 3.2, and Corollary 3.3 can be found in Riemannian Geometry (Busemann functions on complete simply connected Riemannian manifolds are $C^{1,1}$, see [Kn]), and in Dynamical Systems (C^0 Lagrangian invariant graphs are Lipschitz, see [He, Théorème 8.14 page 62] and [BP, Theorem 6.1 page 192]). The fact that the Aubry and the Mather sets obtained in Lagrangian Dynamics are Lipschitz graphs [Ma, Theorem 2 page 181] can also be understood as a generalized version of Theorem 3.1, see [Fa, Proposition 3 page 1044].

Before proving Theorem 3.2, we must give a functional formulation of inequality (2.2) and its equality case (2.3). We define $\mathrm{grad}_L u$, the gradient of u with respect to the Lagrangian L, by $\mathrm{grad}_L u(t,q) = \mathcal{L}^{-1}_{(t,q)}(\frac{\partial u}{\partial q}(t,q))$ or $\frac{\partial u}{\partial q}(t,q) = L(t,q,\mathrm{grad}_L u(t,q))$. Then (2.2) and (2.3) give

$$\forall (t,q) \in U, v \in T_qM, d_{(t,q)}u(1,v) \leqslant \frac{\partial u}{\partial t}(t,q) + H(t,q,\frac{\partial u}{\partial q}(t,q)) + L(t,q,v), \tag{3.1}$$

with equality if and only if $v = \mathrm{grad}_L u(t,q)$.

PROPOSITION 3.4.— *Let U be an open subset of $\mathbb{R} \times M$. Let $u : U \to \mathbb{R}$ be a C^1 solution of the Hamilton-Jacobi equation*

$$\frac{\partial u}{\partial t}(t,q) + H(t,q,\frac{\partial u}{\partial q}(t,q)) = 0.$$

Then for each continuous piecewise C^1 path $\gamma : [a,b] \to M, a < b$, whose graph $\{(t,\gamma(t)) \mid t \in [a,b]\})$ is contained in U, we have

$$u(b,\gamma(b)) - u(a,\gamma(a)) \leqslant \int_a^b L(s,\gamma(s),\dot\gamma(s))\,ds,$$

with equality if and only if γ *is an integral curve of the vector field* $\mathrm{grad}_L u$.

Proof.— Since u is a solution of the Hamilton-Jacobi equation, the inequality (3.1) gives

$$L(s,\gamma(s),\dot\gamma(s)) - d_{(s,\gamma(s))}u(1,\dot\gamma(s)) \geqslant 0,$$

with equality if and only if $\dot\gamma(s) = \mathrm{grad}_L u(s,\gamma(s))$.

Since $\int_a^b d_{(s,\gamma(s))}u(1,\dot\gamma(s))\,ds = u(b,\gamma(b)) - u(a,\gamma(a))$, this proves the proposition. □

COROLLARY 3.5.— *Under the same hypothesis as in Theorem 3.1, the integral curves of* $\mathrm{grad}_L u$ *are local minimizers of* L. *The vector field* $\mathrm{grad}_L u$ *is uniquely integrable. Moreover, if* $(t,x) \in U$ *and* $\gamma :]a,b[\to M$ *is the extremal of* L *such that* $\gamma(t) = x$ *and* $\dot\gamma(t) = \mathrm{grad}_L u(t,x)$, *then* γ *is a solution of* $\mathrm{grad}_L u$ *on the interval* $]a',b'[$ *with* $a' = \inf\{s \in]a,t] \mid \forall s' \in [s,t], (s',\gamma(s')) \in U\}$ *and* $b' = \sup\{s \in [t,b[\mid \forall s' \in [t,s], (s',\gamma(s')) \in U\}$

Proof.— By the Cauchy-Peano Theorem, for some $\epsilon > 0$, there exists a C^1 solution $\gamma :]t-\epsilon, t+\epsilon[\to M$ of $\mathrm{grad}_L u$ with $\gamma(t) = x$. By Proposition 3.1, on any compact subinterval $[t_1,t_2] \subset]t-\epsilon,t+\epsilon[$ is a minimizer for L among continuous piecewise C^1 curves $\delta : [t_1,t_2] \to M$ with $(s,\delta(s)) \in U$ for all $s \in [t_1,t_2]$, hence γ is an extremal for L.

Because extremals of L are determined by their position and their speed at one point, we obtain that $\mathrm{grad}_L u$ is uniquely integrable and that its solutions are the extremals of L with speed at one point given by $\mathrm{grad}_L u$. It follows that such an extremal $\gamma :]a,b[\to M$ is a solution of $\mathrm{grad}_L u$ on any interval $[a',b']$ such that $(s,\gamma(s)) \in U$ for all $s \in [a',b']$. □

It is convenient to introduce the following notation

NOTATION 3.6.— For V (resp. U) an open set in M (resp. $\mathbb{R} \times M$) and $[a,b]$ an interval in $\mathbb{R}$, we denote by $\mathcal{PC}^1([a,b],V)$ (resp. $\mathcal{PC}^1_g([a,b],U)$ the set of continuous piecewise C^1 paths $\gamma : [a,b] \to M$ whose image $\gamma([a,b])$ (resp. whose graph $\{(t,\gamma(t)) \mid t \in [a,b]\}$) is contained in V (resp. U).

COROLLARY 3.7.— *Let U be an open subset of $\mathbb{R} \times M$. For every compact subset $K \subset U$ and every constant $A \geqslant 0$, we can find $\epsilon_0 > 0$ such that if $\epsilon \in]0, \epsilon_0]$, and $u : U \to \mathbb{R}$ is a C^1 solution of the Hamilton-Jacobi equation*

$$\frac{\partial u}{\partial t}(t,q) + H(t,q,\frac{\partial u}{\partial q}(t,q)) = 0,$$

with $\sup\{\|\frac{\partial u}{\partial q}(t,q)\| \mid (t,q) \in K\} \leqslant A$, *then for every* $(t,q) \in K$:

$$\begin{aligned} u(t,q) &= \inf_{\gamma \in \mathcal{PC}^1_g([t-\epsilon,t],U),\gamma(t)=q} u(t-\epsilon,\gamma(t-\epsilon)) + \int_{t-\epsilon}^{t} L(s,\gamma(s),\dot\gamma(s))\,ds; \\ &= \sup_{\gamma \in \mathcal{PC}^1_g([t,t+\epsilon],U),\gamma(t)=q} u(t+\epsilon,\gamma(t-\epsilon))) - \int_{t}^{t+\epsilon} L(s,\gamma(s),\dot\gamma(s))\,ds. \end{aligned}$$

Moreover, the infimum and the supremum are attained by solutions of $\mathrm{grad}_L u$ *(which are extremals for L).*

Proof.— Since the Legendre transform is a global diffeomorphism $\mathcal{L} : \mathbb{R} \times TM \to \mathbb{R} \times T^*M$, there is a constant B such that $\sup\{\|\frac{\partial u}{\partial q}(t,q)\| \mid (t,q) \in K\} \leqslant A$ implies $\sup\{\|\mathrm{grad}_L u(t,q)\| \mid (t,q) \in K\} \leqslant B$. The set $\{(t,q,v) \in \mathbb{R} \times TM \mid (t,q) \in K, \|v\| \leqslant B\}$ is compact, hence, there exists ϵ_0 such that $\mathcal{F}(s,t,q,v)$ is defined on $\{(s,t,q,v) \in \mathbb{R} \times TM \mid (t,q) \in K, \|v\| \leqslant B, |s-t| \leqslant \epsilon_0\}$ and $(s,\mathcal{F}(s,t,q,v)) \in U$ for all such (s,t,q,v). In particular, the solution of $\mathrm{grad}_L u$ at the point $(t,q) \in K$ will be defined on the interval $[t-\epsilon_0, t+\epsilon_0]$. The corollary now follows from Proposition 3.3. □

Proof of Theorem 3.2.— It suffices to consider the case where $M = \mathbb{R}^n$ and the Riemannian metric is the usual Euclidean metric. We fix some $\epsilon_0 > 0$ such that we can apply 3.7 to K', and moreover for every $(t,q) \in K'$ and every $v \in \mathbb{R}^n$, with $\|\mathcal{L}_{(t,q)}(v)\| \leqslant A$, the extremal $\gamma_{(t,q,v)}$ of L, such that $\gamma_{(t,q,v)}(t) = q$, and $\dot\gamma_{(t,q,v)}(t) = v$, is defined on $[t-\epsilon_0, t+\epsilon_0]$ and $(s,\gamma_{(t,q,v)}(s)) \in U$ for each $s \in [t-\epsilon_0, t+\epsilon_0]$. Since the set $Z = \{(t,q,v) \mid (t,q) \in K', \|\mathcal{L}_{(t,q)}(v)\| \leqslant A\}$ is compact and $(s,t,q,v) \mapsto \gamma_{(t,q,v)}(s)$ is C^1 wherever it is defined, the set $Y = \{(s,\gamma_{(t,q,v)}(s),\dot\gamma_{(t,q,v)}(s)) \mid (t,q,v) \in Z, s \in [t-\epsilon_0, t+\epsilon_0]\}$ is a compact subset of $U \times \mathbb{R}^n$. We choose some $\alpha > 0$ such that $\bar V_\alpha(Y)$, the closed neighborhood of points in $\mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n$ at distance $\leqslant \alpha$ from Y, is compact, and contained in $U \times \mathbb{R}^n$ (the distance used on $\mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n$ is the Euclidean distance).

Let us now fix $(t,q) \in K'$ and let $\gamma : [t-\epsilon_0, t] \to \mathbb{R}^n$ be the extremal such that $\gamma(t) = q$ and $\dot\gamma(t) = \mathrm{grad}_L(t,q)$. By Corollary 3.7, we have

$$u(t,q) = u(t-\epsilon_0, \gamma(t-\epsilon_0)) + \int_{t-\epsilon_0}^{t} L(s,\gamma(s),\dot\gamma(s))\,ds.$$

In the sequel, it will be convenient to introduce the affine bijection $\theta : [t-\epsilon_0, t] \to [0,1]$. We have

$$\theta(t-\epsilon_0) = 0, \theta(t) = 1, 0 \leqslant \theta(s) \leqslant 1, \dot\theta(s) = \epsilon_0^{-1}.$$

If $\eta \in \mathbb{R}$ is such that $|\eta| \leqslant \epsilon_0/2$ we introduce the affine bijection $\psi_\eta : [t-\epsilon_0, t] \to [t-\epsilon_0, t+\eta]$. We have

$$\forall s \in [t-\epsilon_0, t], \psi_\eta(s) - s = \eta\epsilon_0^{-1}(s-t+\epsilon_0),$$

hence

$$\forall s \in [t-\epsilon_0, t], |\psi_\eta(s) - s| \leqslant |\eta|.$$

We will also need to know that

$$\forall s \in [t-\epsilon_0, t], \dot\psi_\eta(s) = (\eta+\epsilon_0)\epsilon_0^{-1},$$

and

$$\forall s \in [t-\epsilon_0, t+\eta], \dot{\psi}_\eta^{-1}(s) = \epsilon_0(\eta+\epsilon_0)^{-1}.$$

If $\eta \in \mathbb{R}$ is such that $|\eta| \leqslant \epsilon_0/2$, and $h \in \mathbb{R}^n$, we define a path $\tilde\gamma_{\eta,h} : [t-\epsilon_0, t+\eta] \to \mathbb{R}^n$ by

$$\tilde\gamma_{\eta,h}(s) = \gamma(\psi_\eta^{-1}(s)) + \theta(\psi_\eta^{-1}(s))h,$$

in particular

$$\|\tilde\gamma_{\eta,h}(s) - \gamma(\psi_\eta^{-1}(s))\| \leqslant \|h\|.$$

Moreover, we have

$$\dot{\tilde\gamma}_{\eta,h}(s) = \epsilon_0(\eta+\epsilon_0)^{-1}\dot\gamma(\psi_\eta^{-1}(s)) + (\eta+\epsilon_0)^{-1}h,$$

in particular

$$\dot{\tilde\gamma}_{\eta,h}(s) - \dot\gamma(\psi_\eta^{-1}(s)) = -\eta(\eta+\epsilon_0)^{-1}\dot\gamma(\psi_\eta^{-1}(s)) + (\eta+\epsilon_0)^{-1}h,$$

If we set $C_1 = \sup\{\|v\| \mid \exists(t,q), (t,q,v) \in Y\}$, which is finite since Y is compact, using $(s,\gamma(s),\dot\gamma(s)) \in Y$ for $s \in [t-\epsilon_0, t]$ and $|(\eta+\epsilon_0)^{-1}| \leqslant 2\epsilon_0^{-1}$ for $|\eta| \leqslant \epsilon_0/2$, we obtain

$$\|\dot{\tilde\gamma}_{\eta,h}(s) - \dot\gamma(\psi_\eta^{-1}(s))\| \leqslant 2\epsilon_0^{-1}(C_1|\eta| + \|h\|).$$

Remark that C_1 depends only on K', L and ϵ_0.

From the estimates given above, we have

$$\forall s \in [t-\epsilon_0, t+\eta],$$

$$\|(s, \tilde{\gamma}_{\eta,h}(s), \dot{\tilde{\gamma}}_{\eta,h}(s)) - (\psi_\eta^{-1}(s), \gamma(\psi_\eta^{-1}(s)), \dot{\gamma}(\psi_\eta^{-1}(s)))\| \leqslant C_2\|(\eta,h)\|,$$

where C_2 depends only on K', L and ϵ_0. Since $(s,\gamma(s),\dot{\gamma}(s)) \in Y$, for $s \in [t-\epsilon_0, t]$, it follows that, for $\|(\eta,h)\| \leqslant \min(\epsilon_0/2, \alpha/C_2)$, the whole segment joining the point $(s, \tilde{\gamma}_{\eta,h}(s), \dot{\tilde{\gamma}}_{\eta,h}(s))$ to to the point $(\psi_\eta^{-1}(s), \gamma(\psi_\eta^{-1}(s)),$ $\dot{\gamma}(\psi_\eta^{-1}(s)))$ in $\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^n$ is contained in the compact subset $\bar{V}_\alpha(Y) \subset U\times\mathbb{R}^n$. In particular $(s, \tilde{\gamma}_{\eta,h}(s)) \in U$ for all $s \in [t\epsilon_0, t+\eta]$. From the infimum part of Corollary 3.6, we get

$$u(t+\eta, q+h) \leqslant u(t-\epsilon_0, \gamma(t-\epsilon_0)) + \int_{t-\epsilon_0}^{t+\eta} L(r, \tilde{\gamma}_{\eta,h}(r), \dot{\tilde{\gamma}}_{\eta,h}(r))\, dr,$$

hence using the change of variable $r = \psi_\eta(s)$

$$u(t+\eta, q+h) \leqslant u(t-\epsilon_0, \gamma(t-\epsilon_0))$$

$$+ \int_{t-\epsilon_0}^{t} L(\psi_\eta(s), \tilde{\gamma}_{\eta,h}(\psi_\eta(s)), \dot{\tilde{\gamma}}_{\eta,h}(\psi_\eta(s)))(\eta+\epsilon_0)\epsilon_0^{-1}\, ds,$$

subtracting the equality case $u(t,q) = u(t-\epsilon_0, \gamma(t-\epsilon_0)) + \int_{t-\epsilon_0}^{t} L(s,\gamma(s),\dot{\gamma}(s))\, ds$ gives

$$u(t+\eta, q+h) - u(t,q) \leqslant$$

$$\int_{t-\epsilon_0}^{t} [(\eta+\epsilon_0)\epsilon_0^{-1} L(\psi_\eta(s), \tilde{\gamma}_{\eta,h}(\psi_\eta(s)), \dot{\tilde{\gamma}}_{\eta,h}(\psi_\eta(s))) - L(s,\gamma(s),\dot{\gamma}(s))]\, ds.$$

Since the segment joining $(\psi_\eta(s), \tilde{\gamma}_{\eta,h}(\psi_\eta(s)), \dot{\tilde{\gamma}}_{\eta,h}(\psi_\eta(s)))$ to $(s,\gamma(s),\dot{\gamma}(s))$ in $\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^n$ is contained in the compact subset $\bar{V}_\alpha(Y)$ of $U\times\mathbb{R}^n$, taking into account that

$$(\psi_\eta(s), \tilde{\gamma}_{\eta,h}(\psi_\eta(s)), \dot{\tilde{\gamma}}_{\eta,h}(\psi_\eta(s))) - (s,\gamma(s),\dot{\gamma}(s)) =$$

$$(\eta\epsilon_0^{-1}(s-t+\epsilon_0), \theta(s)h, -(\eta+\epsilon_0)^{-1}\eta\dot{\gamma}(\psi_\eta^{-1}(s)) + (\eta+\epsilon_0)^{-1}h)$$

and

$$\|(\psi_\eta(s), \tilde{\gamma}_{\eta,h}\psi_\eta((s), \dot{\tilde{\gamma}}_{\eta,h}(\psi_\eta(s))) - (s,\gamma(s),\dot{\gamma}(s))\| \leqslant C_2\|(\eta,h)\|,$$

by Taylor formula applied to L, we have

$$L(\psi_\eta(s),\tilde\gamma_{\eta,h}(\psi_\eta(s)),\dot{\tilde\gamma}_{\eta,h}(\psi_\eta(s))) - L(s,\gamma(s),\dot\gamma(s)) \leqslant$$

$$DL(s,\gamma(s),\dot\gamma(s))(\eta\epsilon_0^{-1}(s-t+\epsilon_0),\theta(s)h,-(\eta+\epsilon_0)^{-1}\eta\dot\gamma((s))+(\eta+\epsilon_0)^{-1}h)$$
$$+C_3C_2^2\|(\eta,h)\|^2,$$

where the constant C_3 is given by

$$C_3 = \sup_{(t,q,v)\in\bar V_\alpha(Y)} \|D^2L(t,q,v)\|/2.$$

Since we can write

$$(\eta+\epsilon_0)\epsilon_0^{-1}L(\psi_\eta(s),\tilde\gamma_{\eta,h}(\psi_\eta(s)),\dot{\tilde\gamma}_{\eta,h}(\psi_\eta(s))) - L(s,\gamma(s),\dot\gamma(s))$$

in the following equivalent way

$$(\eta+\epsilon_0)\epsilon_0^{-1}[L(\psi_\eta(s),\tilde\gamma_{\eta,h}(\psi_\eta(s)),\dot{\tilde\gamma}_{\eta,h}(\psi_\eta(s))) - L(s,\gamma(s),\dot\gamma(s))]$$
$$+\frac{\eta}{\epsilon_0}L(s,\gamma(s),\dot\gamma(s)),$$

we obtain

$$(\eta+\epsilon_0)\epsilon_0^{-1}L(\psi_\eta(s),\tilde\gamma_{\eta,h}(\psi_\eta(s)),\dot{\tilde\gamma}_{\eta,h}(\psi_\eta(s))) - L(s,\gamma(s),\dot\gamma(s)) \leqslant$$

$$(\eta+\epsilon_0)\epsilon_0^{-1}\frac{\partial L}{\partial t}(s,\gamma(s),\dot\gamma(s))(\eta\epsilon_0^{-1}(s-t+\epsilon_0))+(\eta+\epsilon_0)\epsilon_0^{-1}\frac{\partial L}{\partial q}(s,\gamma(s),\dot\gamma(s))(\theta(s)h)$$
$$+(\eta+\epsilon_0)\epsilon_0^{-1}\frac{\partial L}{\partial v}(s,\gamma(s),\dot\gamma(s))(-(\eta+\epsilon_0)^{-1}\eta\dot\gamma(s)+(\eta+\epsilon_0)^{-1}h)$$
$$+(\eta+\epsilon_0)\epsilon_0^{-1}C_3C_2\|(\eta,h)\|^2+\eta\epsilon_0^{-1}L(s,\gamma(s),\dot\gamma(s)),$$

We set

$$C_4 = \sup\{\max(|\frac{\partial L}{\partial t}(t,q,v)|,\|\frac{\partial L}{\partial q}(t,q,v)\|) \mid (t,q,v)\in Y\},$$

which is finite by compactness of Y. Using that $(\eta+\epsilon_0)\epsilon_0^{-1} = 1+\eta\epsilon_0^{-1} \leqslant 3/2$, for $|\eta|\leqslant\epsilon_0/2$, and that $s\in[t-\epsilon_0,t]$, we obtain

$$(\eta+\epsilon_0)\epsilon_0^{-1}L(\psi_\eta(s),\tilde\gamma_{\eta,h}(\psi_\eta(s)),\dot{\tilde\gamma}_{\eta,h}(\psi_\eta(s))) - L(s,\gamma(s),\dot\gamma(s)) \leqslant$$

$$\frac{\partial L}{\partial t}(s,\gamma(s),\dot\gamma(s))(\eta\epsilon_0^{-1}(s-t+\epsilon_0))+\epsilon_0^{-1}\frac{\partial L}{\partial q}(s,\gamma(s),\dot\gamma(s))(\theta(s)h)$$

$$+\epsilon_0^{-1}\frac{\partial L}{\partial v}(s,\gamma(s),\dot\gamma(s)(-\eta\dot\gamma(s))+h)+\eta\epsilon_0^{-1}L(s,\gamma(s),\dot\gamma(s))$$
$$+C_4\epsilon_0^{-1}(|\eta|^2+|\eta|\|h\|)+3C_3C_2\|(\eta,h)\|^2/2$$

Setting $C=2C_4\epsilon_0^{-1}+3C_3C_2/2$, by integration we obtain

$$u(t+\eta,q+h)-u(t,q)\leqslant$$
$$\int_{t-\epsilon_0}^{t}[\frac{\partial L}{\partial t}(s,\gamma(s),\dot\gamma(s))(\eta\epsilon_0^{-1}(s-t+\epsilon_0))+\epsilon_0^{-1}\frac{\partial L}{\partial q}(s,\gamma(s),\dot\gamma(s))(\theta(s)h)$$
$$+\epsilon_0^{-1}derLv(s,\gamma(s),\dot\gamma(s)(-\eta\dot\gamma(s))+h)+\eta\epsilon_0^{-1}L(s,\gamma(s),\dot\gamma(s))]\,ds$$
$$+C\|(\eta,h)\|^2.$$

Since u is C^1 and

$$\int_{t-\epsilon_0}^{t}[\frac{\partial L}{\partial t}(s,\gamma(s),\dot\gamma(s))(\eta\epsilon_0^{-1}(s-t+\epsilon_0))+\epsilon_0^{-1}\frac{\partial L}{\partial q}(s,\gamma(s),\dot\gamma(s))(\theta(s)h)$$
$$+\epsilon_0^{-1}\frac{\partial L}{\partial v}(s,\gamma(s),\dot\gamma(s)(-\eta\dot\gamma(s))+h)+\eta\epsilon_0^{-1}L(s,\gamma(s),\dot\gamma(s))]\,ds$$

is a linear map of (η,h), we must have

$$du(t,q)(\eta,h)=$$
$$\int_{t-\epsilon_0}^{t}[\frac{\partial L}{\partial t}(s,\gamma(s),\dot\gamma(s))(\eta\epsilon_0^{-1}(s-t+\epsilon_0))+\epsilon_0^{-1}\frac{\partial L}{\partial q}(s,\gamma(s),\dot\gamma(s))(\theta(s)h)$$
$$+\epsilon_0^{-1}\frac{\partial L}{\partial v}(s,\gamma(s),\dot\gamma(s)(-\eta\dot\gamma(s))+h)+\eta\epsilon_0^{-1}L(s,\gamma(s),\dot\gamma(s))]\,ds.$$

Finally we have obtained that for $\|(\eta,h)\|\leqslant\min(\epsilon_0/2,\alpha/C_2)$ and $(t,q)\in K'$

$$u(t+\eta,q+h)-u(t,q)-du(t,q)(\eta,h)\leqslant C\|(\eta,h)\|^2,$$

where C does not depend on u provided that $\|\mathrm{grad}_L u(t,q)\|\leqslant A$, for all $(t,q)\in K'$.

Using the supremum part of Corollary 3.6, in the same way, we obtain for $\|(\eta,h)\|\leqslant\min(\epsilon_0/2,\alpha/C_2)$ and $(t,q)\in K'$

$$u(t+\eta,q+h)-u(t,q)-du(t,q)(\eta,h)\geqslant -C\|(\eta,h)\|^2.$$

□

The following lemma implies that u is $\mathrm{C}^{1,1}$ on the interior of K' with local Lipschitz constant $\leqslant 6C'$ at every point. This lemma can be found in different forms either implicitly or explicitly in the literature, see [CC, Proposition 1.2 page 8], [HE, Proof of 8.14, pages 63–65], [Kn], [Li, Proof of Theorem 15.1, pages 258-259], and also [Ki] for far reaching generalizations. The simple proof given below evolved from discussions with Bruno Sevennec.

LEMMA 3.8.— *Let E be a normed space. Let $\varphi : B(a,r) \to \mathbb{R}$ be a map defined on the open ball of center a and radius r in E. If there is a finite constant $C \geqslant 0$, and for each $x \in B(a,r)$ there is a linear continuous map $\varphi_x : E \to \mathbb{R}$ such that*

$$\forall x, y \in B(a,r), |\varphi(y) - \varphi(x) - \varphi_x(y-x)| \leqslant C\|(y-x)\|^2,$$

then φ is $\mathrm{C}^{1,1}$ with $D\varphi(x) = \varphi_x$, and the Lipschitz constant of the derivative is $\leqslant 6C$.

Proof.— Obviously φ is differentiable at every point $x \in B(a,r)$ with $D\varphi(x) = \varphi_x$.

Let $x \in B(a,r)$ and $\delta = (r - \|x\|)/2$. If $h, k \in E$ both have norms $< \delta$ then $x+h, x+k, x+h+k$ are all in $B(a,r)$. We do have

$$\begin{aligned} |\varphi(x+h+k) - \varphi(x) - \varphi_x(h+k)| &\leqslant C\|h+k\|^2 \\ |\varphi(x+h+k) - \varphi(x+h) - \varphi_{x+h}(k)| &\leqslant C\|k\|^2 \\ |\varphi(x+h) - \varphi(x) - \varphi_x(h)| &\leqslant C\|h\|^2 \end{aligned}$$

Changing the sign inside the absolute value of the first inequality and adding the three lines gives

$$|\varphi_x(k) - \varphi_{x+h}(k)| \leqslant C[\|h+k\|^2 + \|k\|^2 + \|h\|^2] \leqslant 3C[\|k\|^2 + \|h\|^2].$$

If we fix $h \neq 0$ and remark that the norm of a continuous linear map $\psi : E \to \mathbb{R}$ is also given by

$$\|\psi\| = \|h\|^{-1} \sup\{|\psi(k)| \mid k \in E, \|k\| = \|h\|\},$$

we see that

$$\forall h \in E, \|h\| < \delta \Rightarrow \|\varphi_{x+h} - \varphi_x\| \leqslant 6C\|h\|.$$

this shows that $x \mapsto \varphi_x$ is locally Lipschitz with local Lipschitz constant $\leqslant 6C$. Since $B(a,r)$ is convex, the map $x \mapsto \varphi_x$ is in fact Lipschitz on $B(a,r)$ with Lipschitz constant $\leqslant 6C$. □

Remark 3.9.— By inspection of the proof, it is possible to obtain some compactness theorems for families of Lagrangians. For example, if $\mathcal{K}$ is a compact space, and we have a family of Lagrangians $L_k, k \in \mathcal{K}$, all defined on $\mathbb{R} \times TM$, satisfying assumptions (C1), (C2), and (C3), and such that the map $\mathcal{K} \to C^2(TM,\mathbb{R}), k \mapsto L_k$, is continuous, where $C^2(TM,\mathbb{R})$ is provided with the (compact-open) C^2 topology, then the constant B_k given by Theorem 3.2 applied to L_k can be chosen to be independent of $k \in \mathcal{K}$.

It might be worthwhile to consider the case of a time-independent Lagrangian $L : TM \to \mathbb{R}$ and "time-independent" solutions of the Hamilton-Jacobi equation. Of course, the Hamiltonian $H : T^*M \to \mathbb{R}$ is also time-independent. In fact, if $u : V \to \mathbb{R}$ is some C^1 function defined on the open subset V of M and such that $H(q, d_q u) = c$ is a constant then the function $\hat{u} : \mathbb{R} \times V \to \mathbb{R}, (t, q) \mapsto -tc + u(q)$ is a solution of the Hamilton-Jacobi equation. We will say that such a function u is a time-independent solution of the Hamilton-Jacobi equation, and we will call the constant $c = H(q, d_q u)$ the Hamiltonian constant of u, we will denote this constant by $\mathbf{H}(u)$.

In that setting, Theorems 3.1 and 3.2 can be stated as:

THEOREM 3.10.— *Suppose that the Lagrangian $L : TM \to \mathbb{R}$ is time-independent and satisfies assumptions (C1), (C2), and (C3). Denote by $H : T^*M \to R$ its associated Hamiltonian. Any C^1 time-independent solution $u : V \to \mathbb{R}$ of the Hamilton-Jacobi equation is $\mathrm{C}^{1,1}$.*

Moreover, if we call $\mathcal{SI}^1_c(V)$ the set of such C^1 time-independent solutions $u : V \to \mathbb{R}$ of the Hamilton-Jacobi equation, defined all on the same open subset V of M, and satisfying $\mathbf{H}(u) \leqslant c$, then, the family of derivatives of functions in $\mathcal{SI}^1_c(V)$ is equi-Lipschitzian on any compact subset of V.

It is worth to state the particular case of a Riemannian metric, and to show how one can obtain easily from it the smoothness of the Busemann functions for complete simply connected Riemannian manifolds without conjugate points (compare with [Es] and [Kn], and notice the strong similarity of our proof of Theorem 3.1 and their proofs for Busemann functions).

THEOREM 3.11.— *Suppose that M is a Riemannian manifold. If V is an open subset of M, denote by $\mathcal{G}_1(V)$ the set of C^1 functions $f : V \to \mathbb{R}$, such that the derivative $d_x f$ is of norm 1 at each point $x \in V$. Any function in $\mathcal{G}_1(V)$ is $\mathrm{C}^{1,1}$.*

Moreover, the set of derivatives of functions in $\mathcal{G}_1(V)$ is equi-Lipschitzian on any compact subset of V.

In particular, if C is a finite constant and x_0 is some fixed point, the set $\{f \in \mathcal{G}_1(V) \mid |f(x_0)| \leqslant C\}$ is compact for the (compact-open) C^1-topology.

It is surprising that such a basic result has not appeared before in this form in the literature.

Notice that there is no assumption about the completeness of the Riemannian manifold in Theorem 3.11.

We now obtain the smoothness of Busemann functions alluded to above. Let us recall the definition of Busemann functions. A ray in a Riemannian manifold M is a curve $\gamma : [0, +\infty[\to M$ such that $d(\gamma(s), \gamma(s')) = |s - s'|$, for each $s, s' \in [0, +\infty[$, where d is the Riemannian metric. A ray is necessarily a geodesic parametrized by arc-length. For each $t \geqslant 0$, the function $b_t : M \to \mathbb{R}$, defined by $b_t(x) = t - d(x, \gamma(t))$, is Lipschitzian with Lipschitz constant 1. From the triangle inequality, we have $t \leqslant d(\gamma(0), x) + d(x, \gamma(t))$, and $d(x, \gamma(t')) \leqslant d(x, \gamma(t)) + t' - t$, for $t' \geqslant t$. This gives $b_t(x) \leqslant d(\gamma(0), x)$, and $b_t(x) \leqslant b_{t'}(x)$, for $t' \geqslant t$. It follows that $B_\gamma(x) = \lim_{t\to+\infty} b_t(x)$ exists. The function $B_\gamma : M \to \mathbb{R}$ is called the Busemann function of the ray γ. It is Lipschitzian with Lipschitz constant 1, since all b_t have Lipschitz constant 1.

THEOREM 3.12 (Eschenburg-Knieper).— *Suppose that M is a complete simply connected Riemannian manifold without conjugate points. If $\gamma : [0, +\infty[\to M$ is a ray, its Busemann function B_γ is $C^{1,1}$.*

Proof.— The main point here is that for each $y \in M$, the distance function $d_y : M \to \mathbb{R}, x \mapsto d(y, x)$ is C^∞ on $M \setminus \{y\}$. In fact, the exponential map $\exp_y : T_yM \to M$ is a surjective C^∞ diffeomorphism, and $d(x, y) = \|\exp_y^{-1}(x)\|_y$. It is not difficult to check that the function $x \mapsto d(x, y)$ has a derivative of norm 1 at each point of $M \setminus \{y\}$. If V is an open and relatively compact subset of M, for t large enough, the functions b_t are all C^∞ and have derivatives of norm 1 at each point of V. It follows from Theorem 3.11 that $B_\gamma|V$ is $C^{1,1}$. □

4. The case of an almost complete Lagrangian

DEFINITION 4.1 (Almost complete Lagrangian).— *We say that a Lagrangian $L : \mathbb{R} \times TM \to \mathbb{R}$ is almost complete, if for every curve $\gamma :]a, b[\to M$ for L, which is an extreamal for L, and whose graph $\{(t, \gamma(t)) \mid t \in]a, b[\}$ is relatively compact in $\mathbb{R} \times M$, the norm of its speed $\|\dot\gamma(t))\|$ (for some or any Riemannian metric on M) remains uniformly bounded for $t \in]a, b[$ (or equivalently the graph of the speed curve $\{(t, \gamma(t), \dot\gamma(t)) \mid t \in]a, b[\}$ is contained in a compact subset of $\mathbb{R} \times TM$).*

Examples 4.2.— 1) Let us recall that a vector field (or its flow) is said to be complete if its flow is globally defined for all time (or equivalently maximal solutions are defined on $\mathbb{R}$). If the Euler-Lagrange vector field X_L of the Lagrangian L is complete then it is almost complete. In fact, if $\gamma :]a,b[\to M$ is an extremal curve for such an L with its graph $\{(t,\gamma(t)) \mid t \in]a,b[\}$ relatively compact in $\mathbb{R} \times M$ then a and b are finite and the extremal γ can be extended to an extremal defined on $\mathbb{R}$.

2) If L is an almost complete Lagrangian defined on $\mathbb{R} \times TM$, then for every open subset $U \subset M$ the restricted Lagrangian $L|\mathbb{R} \times TU$ is also almost complete.

3) A time-independent Lagrangian is always almost complete. This follows from Lemma 2.1 and the fact that the flow ϕ_t^* preserves the Hamiltonian.

Here is the property of almost complete Lagrangians that will be used in the sequel

PROPOSITION 4.3.— *Let $L : \mathbb{R} \times TM \to \mathbb{R}$ be a C^2 almost complete Lagrangian.*

1) If $\gamma :]a,b[\to M$ is a maximal extremal curve for L and $b < +\infty$ (resp. $a > -\infty$) then for each compact subset K of M there exists a sequence $t_i \to b$ (resp. $t_i \to a$) in $]a,b[$ such that $\gamma(t_i) \notin K$.

2) If K is a compact subset of $\mathbb{R} \times M$ and ϵ is > 0. The set E of points $(t,q,v) \in \mathbb{R} \times TM$ for which there exists an extremal curve $\gamma : [t-\epsilon, t+\epsilon] \to M$ with $(\gamma(t), \dot\gamma(t)) = (q,v)$ and $(s,\gamma(s)) \in K$, for each $s \in [t-\epsilon, t+\epsilon]$ is a closed subset of $\mathbb{R} \times TM$

Proof.— We prove 1). Suppose $b < +\infty$ and $\epsilon \in]0, b-a[$. If the extremal curve $\gamma|[b-\epsilon, b[$ were entirely contained in the compact set K, its graph would be contained in the compact set $[b-\epsilon, b] \times K$, hence the graph of its speed curve $\{(t,\gamma(t),\dot\gamma(t)) \mid t \in [b-\epsilon, b[\}$ would be contained in a compact subset of $\mathbb{R} \times TM$, we could then extend this solution of the Euler-lagrange vector field on M beyond b. This would contradict the maximality of $\gamma :]a,b[\to M$.

To prove 2), we consider a sequence $(t_i, q_i, v_i) \in E$ tending to $(t_\infty, q_\infty, v_\infty)$. Let us prove, for example that the maximal extremal $\gamma :]c,d[\to M$ with $\gamma(t_\infty) = q_\infty$ and $\dot\gamma(t_\infty) = v_\infty$ verifies $d > t_\infty + \epsilon$. If we call $\gamma_i : [t_i - \epsilon, t_i + \epsilon] \to M$ with $\gamma_i(t_i) = q_i$ and $\dot\gamma_i(t_i) = v_i$. We set $\eta = \inf(\epsilon, d - t_\infty)$. The continuity of the Euler-Lagrange flow wherever it is defined shows that for $s \in [0, \eta[$ we have $\gamma(t_\infty) = \lim_{i\to\infty} \gamma_i(t_i)$. Since the graph $\{(s,\gamma_i(s)) \mid s \in [t_i-\epsilon, t_i+\epsilon]\}$ is contained in the compact subset K, this is also the case for the graph $\{(s,\gamma(s)) \mid s \in [t_\infty, t_\infty+\eta[\}$. The local completeness of L implies that there exists some $\delta > 0$ such that $t_\infty + \eta + \delta \leqslant d$. Hence $\eta = \inf(\epsilon, d - t_\infty) = \epsilon$, and $d \geqslant t_\infty + \epsilon + \delta > t_\infty + \epsilon$. □

PROPOSITION 4.4.— *Suppose the Lagrangian L almost complete. Let U be an open subset of $\mathbb{R}\times M$. For every pair of compact subsets $K, K' \subset U$, with K contained in the interior of K' and every constant $A \geqslant 0$, we can find a constant B such that if $u : U \to \mathbb{R}$ is a C^1 solution of the Hamilton-Jacobi equation*

$$\frac{\partial u}{\partial t}(t,q) + H(t,q,\frac{\partial u}{\partial q}(t,q)) = 0,$$

with $\sup_{(t,q)\in K'} u(t,q) - \inf_{(t,q)\in K'} u(t,q) \leqslant A$,

then $\sup\{\|\frac{\partial u}{\partial q}(t,q)\| \mid (t,q) \in K'\} \leqslant B$.

Proof.— Our first objective is to show that there is $\epsilon > 0$ such that for each u as above the solution γ of $\mathrm{grad}_L u$ starting at $(t,q) \in K$ is such that $(t+s, \gamma(t+s))$ is defined and remains in K' for all $s \in [0,\epsilon]$. We choose some Riemannian metric on M. If $D \geqslant 0$, the superlinearity of L gives us a C_D such that

$$\forall (t,q) \in K', \forall v \in T_qM, D\|v\| \leqslant L(t,q,v) + C_D.$$

Suppose that $(t+s, \gamma(t+s)) \in K'$ for $s \in [0,\alpha]$ and $(t+\alpha, \gamma(t+\alpha))$ is in $\partial K'$ the boundary of K' and $(t,\gamma(t)) \in K$. If we denote by l_γ the length of $\gamma|[t, t+\alpha]$, we have

$$Dl_\gamma \leqslant \int_t^{t+\alpha} L(s,\gamma(s),\dot\gamma(s))\,ds + \alpha C_D. \qquad (4.1)$$

Since γ is an integral curve of $\mathrm{grad}_L u$, we have

$$\int_t^{t+\alpha} L(s,\gamma(s),\dot\gamma(s))\,ds = u(t+\alpha, \gamma(t+\alpha)) - u(t,q) \leqslant A,$$

hence

$$l_\gamma \leqslant AD^{-1} + \alpha C_D D^{-1}.$$

We set $\delta = d(K, \partial K')$, where the distance d on $\mathbb{R}\times TM$ is given by $d((t,m),(t',m')) = [(t-t')^2 + d(m,m')^2]^{1/2}$ with $d(m,m')$ the Riemannian distance in M between m and m'. Since K and $\partial K'$ are disjoint compact sets, the constant δ is > 0. By the definition of δ and the definition of the Riemannian distance, we have

$$\delta^2 \leqslant \alpha^2 + l_\gamma^2.$$

Combining inequalities, we get

$$\delta^2 \leqslant \alpha^2 + (AD^{-1} + \alpha C_D D^{-1})^2.$$

We now fix D such that AD^{-1} is small compared to δ. This implies that there exists $\epsilon > 0$ depending only on δ, A, D, C_D such that $\alpha > \epsilon$.

Using a symmetric argument, we see that we can find $\epsilon > 0$ such that for each u as in the statement of the proposition the solution γ of $\mathrm{grad}_L u$ starting at $(t, q) \in K$ is such that $(t + s, \gamma(t + s))$ is defined and remains in K' for all $s \in [-\epsilon, \epsilon]$. We fix now such an ϵ.

Since by (4.1) with $D = 1$

$$\int_0^{\epsilon/2} \|\dot\gamma(t+s)\|\, ds \leqslant \int_0^{\epsilon/2} L(t+s, \gamma(t+s), \dot\gamma(t+s))\, ds + \epsilon C_1/2,$$

and by Proposition 3.4

$$\int_0^{\epsilon/2} L(t+s, \gamma(t+s), \dot\gamma(t+s))\, ds = u(t+\epsilon/2, \gamma(t+\epsilon/2)) - u(t,q) \leqslant A,$$

we obtain

$$\int_0^{\epsilon/2} \|\dot\gamma(t+s)\|\, ds \leqslant A + \epsilon C_1/2$$

This implies that there is some $s_0 \in [t, t+\epsilon/2]$ such that

$$\|\dot\gamma(s_0)\| \leqslant 2A\epsilon^{-1} + C_1/2.$$

Hence the point $(s_0, \gamma(s_0), \dot\gamma(s_0))$ belongs to the set E of points $(s, q, v) \in TM$ such that $|v| \leqslant 2A\epsilon^{-1} + C_1/2$, and $\sigma \to \mathcal{F}(\sigma, s, q, v)$ is defined on $[s - \epsilon/2, s + \epsilon/2]$, with its graph $\{(\sigma + s, \mathcal{F}_\sigma(s, q, v)) \mid \sigma \in [-\epsilon/2, \epsilon/2]\}$ contained in K'. By Proposition 4.3, the set E is closed, therefore compact, because it is contained in the compact subset $\{(s, q, v) \mid (s, q) \in K', |v| \leqslant 2A\epsilon^{-1} + C_1/2\}$. The continuity of the Euler-Lagrange flow where it is defined then implies that there exists a constant B such that

$$\forall (s,q,v) \in E, \forall \sigma \in [s-\epsilon/2, s+\epsilon/2], \|\frac{\partial \mathcal{F}}{\partial \sigma}(\sigma, s, q, v)\| \leqslant B \;.$$

Returning to the situation considered above we have $\mathrm{grad}_L u(t,q) = \dot\gamma(t)$ and there exists s_0 with $|t - s_0| \leqslant \epsilon/2$ such that $(s_0, \gamma(s_0), \dot\gamma(s_0)) \in E$. Since γ

is an extremal, we obtain $\gamma(t) = \mathcal{F}(t, s_0, \gamma(s_0), \dot\gamma(s_0))$, hence $\mathrm{grad}_L u(t,q) = \frac{\partial \mathcal{F}}{\partial t}(t, s_0, \gamma(s_0), \dot\gamma(s_0))$. It follows that

$$\|\mathrm{grad}_L u(t,q)\| = \|\dot\gamma(t)\| \leqslant B_1.$$

Since the Legendre transform is a diffeomorphism sending $\mathrm{grad}_L u(t,q)$ to $\frac{\partial u}{\partial q}(t,q)$ and $(t,q) \in K$, with K compact, it follows that there exists a constant B_2 depending only on L, K and B_1 such that $\|\frac{\partial u}{\partial q}(t,q)\| \leqslant B_2$, for all $(t,q) \in K$. Since u satisfies the Hamilton-Jacobi equation the is a constant B_3 depending only on H, K and B_2 such that $|\frac{\partial t}{\partial u}(t,q)| = |H(t,q,\frac{\partial u}{\partial q}(t,q))| \leqslant B_3$. □

Using Theorem 3.2 and the proposition above we obtain the following corollary which is reminiscent of Montel's theorem for holomorphic functions:

COROLLARY 4.5.— *Suppose the Lagrangian L almost complete. Let U be an open subset of* $\mathbb{R} \times M$. *Call* $\mathcal{S}^1(U,\mathbb{R})$ *the set of maps* $u : U \to \mathbb{R}$ *of class* C^1 *which satisfy the Hamilton-Jacobi equation*

$$\frac{\partial u}{\partial t}(t,q) + H(t,q,\frac{\partial u}{\partial q}(t,q)) = 0.$$

Any subset of $\mathcal{S}^1(U,\mathbb{R})$ *which is bounded in the compact-open* C^0 *topology on* $\mathcal{C}^0(U,\mathbb{R})$ *is a relatively compact subset* $\mathcal{S}^1(U,\mathbb{R})$ *for the compact-open* C^1 *topology on* $\mathcal{C}^1(U,\mathbb{R})$.

THEOREM 4.6.— *Suppose the Lagrangian L almost complete. Let U be an open subset of* $\mathbb{R} \times M$. *Let* $(u_n)_{n\in\mathbb{N}}$ *be a sequence of* C^1 *solutions of the Hamilton-Jacobi equation, all of them defined on U. Suppose that for each* $x \in U$ *the limit* $u(t,q) = \lim_{n\to\infty} u_n(t,q)$ *exists, then u is* C^1 *on U and* $u_n \to u$ *in the compact-open* C^1 *topology. In particular, the map* $u : U \to \mathbb{R}$ *is also a solution of the Hamilton-Jacobi equation.*

Proof.— As usual we suppose that $M = \mathbb{R}^n$. By the previous corollary, it suffices to show that for each $(t_0, q_0) \in U$ there is a neighborhood $V_0 \subset U$ of (t_0, q_0) and a constant C such that

$$\forall n \in \mathbb{N}, \forall (t,q) \in V_0,\ |u_n(t,q)| \leqslant C.$$

We choose $\epsilon > 0$ and $r > 0$ such that $[t_0 - 2\epsilon, t_0 + 2\epsilon] \times \bar{B}(q_0, r) \subset U$.

For $(t,q) \in [t_0-\epsilon, t_0+\epsilon]\times\bar{B}(q_0,r)$, we can define a path $\gamma_{t,q} : [t_0-2\epsilon, t] \to \bar{B}(q_0,r)$ by $\gamma_{t,q}(s) = \frac{s-(t_0-2\epsilon)}{t-(t_0-2\epsilon)}(q-q_0)+q_0$.

Since $t \geqslant t_0-\epsilon$ and $q \in \bar{B}(q_0,r)$, we have $\|\dot{\gamma}_{t,q}(s)\| \leqslant \epsilon^{-1}r$. It follows that there is a constant C_1 depending only on ϵ, r and L such that for all $(t,q) \in [t_0-\epsilon, t_0+\epsilon]\times\bar{B}(q_0,r)$ we have, for $s \in [t_0-2\epsilon, t]$, the inequality $L(s,\gamma_{t,q}(s),\dot{\gamma}_{t,q}(s)) \leqslant C_1$. Since by Proposition 3.4

$$u_n(t,q) - u_n(t_0-2\epsilon, q_0) \leqslant \int_{t_0-2\epsilon}^{t} L(s,\gamma_{t,q}(s),\dot{\gamma}_{t,q}(s)),$$

and $u_n(t_0-2\epsilon,q_0) \to u(t_0-2\epsilon,q_0)$, we can find a constant C_2 such that

$$\forall n \in \mathbb{N}, \forall (t,q) \in [t_0-\epsilon, t_0+\epsilon]\times\bar{B}(q_0,r),\ u_n(t,q) \leqslant C_2 + 3\epsilon C_1.$$

In the same way, using a path from (t,q) to $(t_0+2\epsilon, q_0)$, we can find constants C_1', C_2' such that

$$\forall n, \in \mathbb{N}, \forall (t,q) \in [t_0-\epsilon, t_0+\epsilon]\times\bar{B}(q_0,r),\ u_n(t,q) \geqslant C_2' - 3\epsilon C_1'. \qquad \square$$

We would like to prove a version of Corollary 4.5 for a time-independent Lagrangian and a time-independent solution.

THEOREM 4.7.— *Let $L : TM \to \mathbb{R}$ be a time-independent Lagrangian satisfying the conditions (C1),(C2), and (C3). If V is an open subset of M, we call $\mathcal{SI}^1(V,\mathbb{R})$ the set of C^1 maps $u : V \to \mathbb{R}$ which are time-independent solutions of the Hamilton-Jacobi equation, i.e. the set of C^1 maps $u : V \to \mathbb{R}$ for which there exists a constant $\mathbf{H}(u)$ such that*

$$\forall q \in V, H(q, \frac{\partial u}{\partial q}(q)) = \mathbf{H}(u).$$

Any subset of $\mathcal{SI}^1(V,\mathbb{R})$ which is bounded in the compact-open C^0 topology on $C^0(V,\mathbb{R})$ is a relatively compact subset $\mathcal{SI}^1(V,\mathbb{R})$ for the compact-open C^1 topology on $C^1(V,\mathbb{R})$.

We first notice that this is stronger than to apply Corollary 4.5 to the functions $\hat{u} : \mathbb{R}\times V \to \mathbb{R}$ defined, for $u \in \mathcal{SI}^1(V,\mathbb{R})$, by $\hat{u}(t,q) = u(q) - \mathbf{H}(u)t$. In fact, if we assume that $|\hat{u}|$ is bounded by A on some compact set of the form $[-\epsilon,\epsilon]\times K$, with $K \subset V$ compact and non-empty, and $\epsilon > 0$, we conclude that $|\mathbf{H}(u)| \leqslant 2A\epsilon^{-1}$, but there is no such assumption in our statement above. The point of Theorem 4.7 is precisely to obtain a bound for $\mathbf{H}(u)$ from a bound on u on a compact subset of V.

Proof of Theorem 4.7.— By the uniform superlinearity of H above compact subsets of M, see Lemma 2.1, it suffices to show that $\mathbf{H}(u)$ is bounded above if we assume that $u \in \mathcal{SI}^1(V,\mathbb{R})$ is bounded in absolute value on some open non-empty ball contained in V. In fact, this will show that any bounded subset of $\mathcal{SI}^1(V,\mathbb{R})$ for the compact-open C^0 topology on $\mathcal{C}^0(V,\mathbb{R})$ is bounded for the compact-open C^1 topology on $\mathcal{C}^1(V,\mathbb{R})$ and hence, by Theorem 3.10, relatively compact for the compact-open C^1 topology.

As usual, we suppose that V is open in $M = \mathbb{R}^k$, and $\bar{B}(q,r) \subset V$ with $r > 0$. By the uniform superlinearity of L above compact sets, see Lemma 2.1, for each $A \geqslant 0$ there exists $C(A) \in \mathbb{R}$ such that

$$\forall q \in \bar{B}(q,r), \forall v \in \mathbb{R}^k, L(x,v) \geqslant A|v| + C(A).$$

The vector field $\mathrm{grad}_L u$ is time-independent, it is uniquely integrable (either by the same reasoning as in Corollary 3.5, or simply because we now know that $\mathrm{grad}_L u$ is locally Lipschitzian by Theorem 3.10), and for each integral curve $\gamma : [a,b] \to V$ of $\mathrm{grad}_L u$ we have

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

We suppose that $|u|$ is uniformly bounded by B on $\bar{B}(q,r)$ and we choose A such that $2B \leqslant Ar$, this is possible because $r > 0$. Suppose that $\gamma : [0,b[\to V$ is a maximal integral curve of $\mathrm{grad}_L u$ with $\gamma(0) = q$. If $\gamma([0,b[) \subset \bar{B}(q,r)$ then, by compactness of $\bar{B}(q,r)$, we must have $b = \infty$. In particular, we obtain

$$u(\gamma(1)) - u(q) = \mathbf{H}(u) + \int_0^1 L(\gamma(s),\dot\gamma(s))\,ds.$$

Since $|u|$ is bounded by B on $\bar{B}(q,r)$, and L is bounded below by $C(0)$ on $\bar{B}(q,r) \times \mathbb{R}^k$, we must have $\mathbf{H}(u) \leqslant -C(0) + 2B$. If on the other hand $\gamma([0,b[)$ is not entirely contained in $\bar{B}(q,r)$, we can find $t > 0$ such that $\gamma([0,t]) \subset \bar{B}(q,r)$ and $\gamma(t) \in \partial B(q,r)$. We obtain in that case

$$\mathbf{H}(u)t + \int_0^t L(\gamma(s),\dot\gamma(s))\,ds = u(\gamma(t)) - u(q) \leqslant 2B.$$

Moreover, by the superlinearity of L, and using $\int_0^t \|\dot\gamma(s)\|\,ds \geqslant d(\gamma(t),q) = r$, we know that $\int_0^t L(\gamma(s),\dot\gamma(s))\,ds \geqslant Ar + C(A)t$. Finally, we see that

$$\mathbf{H}(u)t + Ar + C(A)t \leqslant 2B.$$

Since $t > 0$, and $Ar \geqslant 2B$, we must have $\mathbf{H}(u) \leqslant -C(A)$ □

5. The method of characteristics

THEOREM 5.1.— *Let V be some open set in M and $u_0 : V \to \mathbb{R}$ be a function. There is a* C^1 *solution u of the Hamilton-Jacobi defined on the open neighborhood U of $\{0\} \times V$ in $\mathbb{R} \times M$ with $u(0,q) = u_0(q)$, for each $q \in V$, if and only if u_0 is* $\mathrm{C}^{1,1}$.

From Theorem 3.1 we already know that if u is C^1 then it has to be $\mathrm{C}^{1,1}$, hence its restriction u_0 to $\{0\} \times V$ is also $\mathrm{C}^{1,1}$.

It remains to show the existence of u on an open neighborhood U of $\{0\} \times V$, when u_0 is $\mathrm{C}^{1,1}$.

We will, of course use the so-called Method of Characteristics that usually people use when u_0 is C^2, see [Be, §8 page 23]. We will see that this method of constructing solutions of the Hamilton-Jacobi equation works for $\mathrm{C}^{1,1}$ initial data. This was already observed in [Li, Remark 1.1 page 15], at least for the case where $M = \mathbb{R}^k$ and $L(t,q,v)$ depends only on the vertical variable v (we use here the canonical identification $T\mathbb{R}^k = \mathbb{R}^k \times \mathbb{R}^k$).

There are several ways to show that the method of characteristics works for $\mathrm{C}^{1,1}$ initial data, even if none of them seems to our knowledge to be present in the literature, except for the reference to the work of Pierre-Louis Lions given above. The case of C^2 initial data is well documented and appears in almost every classical treatise on the Calculus of Variations or PDE, see [Be].

Since we will be mostly using the geometric framework, we will recall the geometric setting of the method of characteristics. This will fix the notations and (we hope!) will help the more analytically minded (or trained) reader to follow our arguments which are clearly well-known to the geometrically minded reader (who may skip most of it).

We introduce the (time-independent) flow Φ_s on $\mathbb{R} \times T^*M$ defined by

$$\Phi_s(t,q,p) = (t+s, \phi_s^*(t,q,p)).$$

The (time-independent) vector $\hat{X}_H^*$ generating Φ_s is

$$\hat{X}_H^* = \frac{\partial}{\partial t} + X_H^*,$$

where X_H^* is the (time-dependent) Hamiltonian vector field on $\mathbb{R} \times T^*M$ associated with H, and generating ϕ_s^*.

We show that this flow Φ_s preserves the closed 2-form $\Omega_H = -d\alpha_H$ on $\mathbb{R} \times T^*M$, where the differential 1-form α_H on $\mathbb{R} \times M$ is defined by

$$\alpha_H = \alpha - Hdt,$$

and α is the Liouville form on T^*M. More precisely, we should write

$$\alpha_H = p_2^*\alpha - Hdt,$$

where $p_2 : \mathbb{R} \times T^*M \to T^*M$ is the projection on the second factor, and dt is the differential on $T^*M \times \mathbb{R}$ of the projection $T^*M \times \mathbb{R} \to \mathbb{R}$ on the first factor. We have

$$\Omega_H = \Omega + dH \wedge dt,$$

where $\Omega = -d\alpha$ is the canonical symplectic form on T^*M. In fact, it is easy to check that $\Omega_H(\hat{X}_H^*, \cdot)$ is identically 0. Since Ω_H is closed, by Cartan's formula $L_X = i_X d + di_X$, this of course implies that Φ_s preserves Ω_H.

If we start with $v_0 : V \to \mathbb{R}$ a C^1 function defined on the open subset V of M, we can consider the graph of the derivative dv_0 in T^*M

$$\text{Graph}(dv_0) = \{(q, d_q v_0) \mid q \in V\}.$$

We now "propagate" Graph(dv_0) into a C^0 submanifold of dimension $k+1$ of $\mathbb{R}\times T^*M$ using the flow Φ_s (here k is the dimension of the base manifold M). More precisely, call Ξ the open subset $\mathbb{R} \times \mathbb{R} \times T^*M$ consisting of the points (s, t, q, p) where $\Phi_s(t, q, p)$ is defined. This set Ξ contains $\{0\} \times \mathbb{R} \times T^*M$. We set

$$\mathcal{V}_0 = \{\Phi_s(0, q, d_q v_0) \mid q \in V, (s, 0, q, d_q v_0) \in \Xi\},$$

and we call $\mathcal{V}_0$ the propagated graph of dv_0. It is easy to check that $\mathcal{V}_0 \supset \{0\} \times \text{Graph}(dv_0)$, and that it is a C^0 submanifold of dimension $k + 1$ of $\mathbb{R} \times T^*M$. In fact, it is the image of an open subset of the C^0 submanifold $\mathbb{R} \times \text{Graph}(dv_0)$ of $\mathbb{R} \times T^*M$ by the C^2 diffeomorphism $(s, q, p) \mapsto \Phi_s(0, q, p) = (s, \phi_s(0, q, p))$ which is defined on the open subset $\{(s, q, p) \in \mathbb{R} \times T^*M \mid (s, 0, q, p) \in \Xi\}$.

To simplify a little bit our exposition and take advantage of the geometry of the graph of a derivative, we will explain the classical method of characteristics for C^3 initial data, although most of what we say can be adapted to the case where u_0 is merely C^2.

From now on, we assume that $v_0 : V \to \mathbb{R}$ is a C^3 function defined on the open subset V of M. In that case, both Graph(dv_0) and $\mathcal{V}_0$ are C^2 submanifolds. The submanifold $\mathcal{V}_0$ is a graph over the projection $\text{Id}_{\mathbb{R}} \times \pi^* : \mathbb{R} \times T^*M \to \mathbb{R} \times M$, $(t, q, p) \mapsto (t, q)$ in a neighborhood of $\{0\} \times \text{Graph}(dv_0)$.

In fact, at a point $(0, q, d_q v_0) \in \{0\} \times \mathrm{Graph}(dv_0)$ the derivative of $\mathrm{Id}_{\mathbb{R}} \times \pi^*$ restricted to $\mathcal{V}_0$ is surjective because its image contains $\{0\} \times T_q M = T_{(0,q,d_q v_0)}[\mathrm{Id}_{\mathbb{R}} \times \pi^*](T_{(0,q,d_q v_0)}\mathrm{Graph}(dv_0))$ and the vector $T_{(q,d_q v_0)}\pi^*(X_H) + \frac{\partial}{\partial t}$. By dimension argument, this derivative of $\mathrm{Id}_{\mathbb{R}} \times \pi^* | \mathcal{V}_0$ is an isomorphism at each point of $\{0\} \times \mathrm{Graph}(dv_0)$. It follows that this restriction is a local diffeomorphism at each point in a neighborhood of $\{0\} \times \mathrm{Graph}(dv_0)$. Since $\mathrm{Id}_{\mathbb{R}} \times \pi^*$ is a homeomorphism on $\{0\} \times \mathrm{Graph}(dv_0)$, an adaptation of a well-known topological argument, see [La, pages 109–110], shows that $\mathrm{Id}_{\mathbb{R}} \times \pi^*$ is a diffeomorphism from an open neighborhood of $\{0\} \times \mathrm{Graph}(dv_0)$ onto its image.

As is well known the Liouville form α restricted to the submanifold $\mathrm{Graph}(dv_0)$ is exact. In fact, if we pullback the Liouville form α by the C^2 derivative section $q \mapsto (q, d_q v_0)$ we obtain the differential 1-form dv_0 on V which is exact (this can be easily checked in coordinates). Since the C^2 derivative section $q \mapsto (q, d_q v_0)$ is a C^2 diffeomorphism of V onto $\mathrm{Graph}(dv_0)$, we do obtain that α is exact. More precisely, we can write $\alpha | \mathrm{Graph}(dv_0) = d[v_0 \circ \pi^*]$.

The important observation is that the restriction of the 1-form α_H to $\mathcal{V}_0$ is exact. To prove this, we first show that $\Omega_H = -d\alpha_H$ is identically 0 as a 2-form on $\mathcal{V}_0$. In fact, the tangent space of $\mathcal{V}_0$ at some point $\Phi_{s_0}(0, q_0, d_{q^0} v_0)$ is the sum of the space generated by $\hat{X}_H^* = X_H^* + \frac{\partial}{\partial t}$ and the image of the space tangent to $\mathrm{Graph}(dv_0)$ at $(0, q_0, d_{q^0} v_0)$ by the derivative of the diffeomorphism Φ_{s_0}. But the vector $\hat{X}_H^*$ is such that $\Omega_H(\hat{X}_H^*, \cdot) = 0$, moreover, the diffeomorphism Φ_{s_0} preserves Ω_H and the 2-form Ω_H restricted to $\{0\} \times \mathrm{Graph}(dv_0)$ is nothing but Ω restricted to $\mathrm{Graph}(dv_0)$ which is identically 0. This implies that α_H is closed when restricted to $\mathcal{V}_0$.

To show that it is exact, we remark that $\mathcal{V}_0$ can be retracted by C^2 deformation to its subset $\{0\} \times \mathrm{Graph}(dv_0)$, since, for $(q, p) \in T^*M$ fixed, the set of s such that $\Phi_s(0, q, p)$ is defined is an open interval of $\mathbb{R}$. Moreover, the restriction of α_H to $\{0\} \times \mathrm{Graph}(dv_0)$ is nothing but the restriction of α to $\mathrm{Graph}(dv_0)$ and $\alpha | \mathrm{Graph}(dv_0) = d[v_0 \circ \pi^*]$. By Poincaré's lemma, we can find a C^2 function $\hat{S}_{v_0}$ defined on the C^2 submanifold $\mathcal{V}_0$ such that $\alpha_h \mid \mathcal{V}_0 = d\hat{S}_{v_0}$ and $\hat{S}_{v_0} | \{0\} \times \mathrm{Graph}(dv_0)$ is nothing but $v_0 \circ \pi^*$.

If $\hat{U}$ is an open neighborhood of $\{0\} \times \mathrm{Graph}(dv_0)$ in the propagated graph $\mathcal{V}_0$ such that $\mathrm{Id}_{\mathbb{R}} \times \pi^*$ induces a diffeomorphism from $\hat{U}$ onto its image which we call U, then we can construct a solution $v : U \to \mathbb{R}$ of the Hamilton-Jacobi equation such that $v(0, q) = v_0(q)$ for each $q \in V$. In fact, if we call $\sigma : U \to \hat{U}$ the inverse of the restriction of $\mathrm{Id}_{\mathbb{R}} \times \pi^*$ to $\hat{U}$, we can write

$\sigma^*\alpha_H = dv$, with $v = \hat{S}_{v_0} \circ \sigma$. We do have $v(0,q) = v_0(q)$ for each $q \in V$. To check that v satisfies the Hamilton-Jacobi equation, we use local coordinates $(q_1, \dots, q_k$ on M, and the associated coordinates $(q_1, \dots, q_k, p_1, \dots, p_k)$ on T^*M. In these coordinates $\alpha_H = \sum_{i=1}^k p_i dq_i - H(t, q_1, \dots, q_k, p_1, \dots, p_k)dt$. Using obvious notations, the equation $\sigma^*\alpha_H = dv$ gives $\dfrac{\partial v}{\partial q_i} = p_i \circ \sigma$ and $\dfrac{\partial v}{\partial t} = -H(t, q_1, \dots, q_k, p_1 \circ \sigma, \dots, p_k \circ \sigma)$. By substitution this does yield

$$\frac{\partial v}{\partial t} + H(t, q_1, \dots, q_k, \frac{\partial v}{\partial q_1}, \dots, \frac{\partial v}{\partial q_k}) = 0.$$

We also remark that the previous argument proves that $\sigma(t,q) = (t, q, \dfrac{\partial v}{\partial q})$, for each $(t,q) \in U$. It follows that v is C^3, since both σ and H are C^2.

We summarize what we obtained in the following lemma:

LEMMA 5.2.— *Let $v_0 : V \to \mathbb{R}$ be a C^3 function defined on the open subset V of M. If we define the propagated graph $\mathcal{V}_0$ of dv_0 as the set of $\Phi_s(0, q, d_q v_0)$ such that $q \in V$ and $\Phi_s(0, q, d_q v_0)$ is defined, then the propagated graph $\mathcal{V}_0$ is a submanifold of $\mathbb{R} \times T^*M$ of dimension $\dim M + 1$, and class C^2, which contains $\{0\} \times \mathrm{Graph}(dv_0)$, where $\mathrm{Graph}(dv_0) = \{(q, d_q v_0) \mid q \in V\}$ is the graph of the derivative dv_0. Moreover, there exists an open neighborhood of $\{0\} \times \mathrm{Graph}(dv_0)$ in $\mathcal{V}_0$ such that $\mathrm{Id}_{\mathbb{R}} \times \pi^*$ induces a diffeomorphism from that neighborhood onto its image.*

The form α_H restricted to $\mathcal{V}_0$ is exact and there exists a C^2 function $\hat{S}_{v_0} : \mathcal{V}_0 \to \mathbb{R}$ such that $\alpha_h \mid \mathcal{V}_0 = d\hat{S}_{v_0}$ and $\hat{S}_{v_0} | \{0\} \times \mathrm{Graph}(dv_0)$ is nothing but $v_0 \circ \pi^*$.

For each C^1 section $\sigma : U \to \mathbb{R} \times T^*M$ of $\mathrm{Id}_{\mathbb{R}} \times \pi^*$, defined on the open subset U of $\mathbb{R} \times M$, with $\sigma(U) \subset \mathcal{V}_0$, the function $v = \hat{S}_{v_0}\sigma$ is a C^3 solution of the Hamilton-Jacobi equation

$$\frac{\partial v}{\partial t} + H(t, q, \frac{\partial v}{\partial q}) = 0,$$

with $v(0,q) = v_0(q)$ for each $q \in V$ satisfying $(0,q) \in U$. The relationship between v and σ is given by $\sigma(t,q) = (t, q, \dfrac{\partial v}{\partial q})$, for each $(t,q) \in U$.

To prove the analogous facts when v_0 is merely to be $C^{1,1}$, we will need to recall a Lipschitz form of the implicit function theorem.

THEOREM 5.3.— *Let $\|\cdot\|$ be some fixed norm on $\mathbb{R}^k$. If $\lambda < 1$ and $\varphi : B(x_0, r) \to \mathbb{R}^k$ is some Lipschitz map with Lipschitz constant $\leqslant \lambda$, then the map $h = \mathrm{Id} + \varphi : \mathring{B}(x_0, r) \to \mathbb{R}^k$ has an open image, and is a bi-Lipschitz homeomorphism onto its image. The Lipschitz constant of its inverse is $\leqslant (1-\lambda)^{-1}$ and the image contains the ball $\mathring{B}(h(x_0), (1-\lambda)r)$.*

Moreover, if $\varphi_n : B(x_0, r) \to \mathbb{R}^k$ is a sequence of Lipschitz maps, all with Lipschitz constant $\leqslant \lambda$, such that φ_n converges uniformly to φ, and K is some compact subset of $h(\mathring{B}(x_0, r))$, then for n large enough the inverse of the map $h_n = \mathrm{Id} + \varphi_n$ is defined on K and the sequence of inverses h_n^{-1} converges uniformly on K to h^{-1}. In fact, we have for each n such that $h_n(\mathring{B}(x_0, r)) \supset K$

$$\sup_{y \in K} \|h_n^{-1}(y) - h^{-1}(y)\| \leqslant (1-\lambda)^{-1} \sup_{x \in B(x_0, r)} \|\varphi_n(x) - \varphi(x)\|.$$

The proposition below finishes the proof of Theorem 5.1.

PROPOSITION 5.4.— *Let $u_0 : V \to \mathbb{R}$ be a $\mathrm{C}^{1,1}$ function, where V is some open set in M. There is an open neighborhood $\hat{U}$ of $\{0\} \times \mathrm{Graph}(du_0)$ in the propagated graph $\mathcal{U}_0$ of du_0 such that $\mathrm{Id}_{\mathbb{R}} \times \pi^*$ induces a homeomorphism from that neighborhood $\hat{U}$ onto an open subset U of $\mathbb{R} \times M$.*

If we call $\sigma : U \to \hat{U}$ the inverse of the restriction $\mathrm{Id}_{\mathbb{R}} \times \pi^* | \hat{U}$, we can find a $\mathrm{C}^{1,1}$ solution $v : U' \to \mathbb{R}$ of the Hamilton-Jacobi equation, defined on an open neighborhood U' of $\{0\} \times V$ in U, such that $v(0, q) = v_0(q)$, for each $q \in V$, and $\sigma(t, q) = (t, q, \dfrac{\partial v}{\partial q}(t, q))$, for $(t, q) \in U'$.

Proof.— We first need to work locally. Rather than introducing coordinates charts, we will assume as usual that M is an open subset of $\mathbb{R}^k$, and we will work in the canonical coordinates of $\mathbb{R}^k$ and $T^*\mathbb{R}^k = \mathbb{R}^k \times (\mathbb{R}^k)^*$, where $(\mathbb{R}^k)^*$ is the dual space of $\mathbb{R}^k$.

We fix some $x \in V$ and choose $r > 0$ such that $\bar{B}(x, 2r) \subset V$. We will call K a Lipschitz constant for the map $x \mapsto d_x u_0 \in (\mathbb{R}^k)^*$, on the compact set $\bar{B}(x, 2r)$. Using an approximation by convolution we can find a sequence v_0^n of C^∞ functions all defined on a neighborhood of $\bar{B}(x, r)$, and such that

1) the sequence v_0^n converges uniformly to u_0 on $\bar{B}(x, r)$;

2) the sequence of derivatives dv_0^n converges uniformly to du_0 on $\bar{B}(x, r)$;

3) For each n, the map $y \mapsto d_y v_0^n$ has a Lipschitz constant on $\bar{B}(x, r)$ which is $\leqslant K$;

The first two properties imply that there exists some constant A such that for each n and each $y \in \bar{B}(x,r)$ we have $\|d_y v_0^n\| \leqslant A$, and also $\|d_y u_0\| \leqslant A$. We can find some $\epsilon > 0$ such that the map $(s,y,v) \mapsto \Phi_s(0,y,v) = (s, \phi_s^*(y,p))$ is defined on a neighborhood of $[-\epsilon,\epsilon] \times \bar{B}(x,r) \times \{p \in (\mathbb{R}^k)^* \mid \|p\| \leqslant A\}$. Since the map $(s,y,p) \mapsto \phi_s^*(y,p)$ is C^2 and $\phi_0^*(y,p) = (y,p)$, we can write $\phi_s^*(y,p) = (y,p) + s\theta(s,y,p)$ with θ of class C^1. In particular, the map θ is Lipschitz on the compact set $[-\epsilon,\epsilon] \times \bar{B}(x,r) \times \{p \in (\mathbb{R}^k)^* \mid \|p\| \leqslant A\}$. We will denote by K_1 a Lipschitz constant of θ on that last set.

We claim that there exist $\epsilon_0 > 0$ and $r_0 > 0$ such that

(i) the map $(s,y,0) \mapsto (\mathrm{Id}_{\mathbb{R}} \times \pi^*) \circ \Phi_s(0,y,d_y u_0) = (s, \pi^* \phi_s^*(0,y,d_y u_0))$ is a homeomorphism h from the open set $]-\epsilon_0,\epsilon_0[\times \mathring{B}(x,r)$ to some open subset of $\mathbb{R} \times \mathbb{R}^*$ containing the compact set $[-\epsilon_0/2, \epsilon_0/2] \times \bar{B}(x,r_0)$;

(ii) for each n, the map $(s,y,0) \mapsto \pi^* \Phi_s(0,y,d_y v_0^n) = (s, \pi^* \phi_s^*(0,y,d_y v_0^n))$ is a homeomorphism h_n from the open set $]-\epsilon_0,\epsilon_0[\times \mathring{B}(x,r)$ to some open subset of $\mathbb{R} \times \mathbb{R}^*$ containing the compact set $[-\epsilon_0/2, \epsilon_0/2] \times \bar{B}(x,r_0)$;

(iii) the sequence of homeomorphisms h_n^{-1} converges uniformly to h^{-1} on the compact set $[-\epsilon_0/2, \epsilon_0/2] \times \bar{B}(x,r_0)$.

In fact, we have $h(s,y) = (s, y + s\theta(s,y,d_y u_0))$, and $h_n(s,y) = (s, y + s\theta(s,y,d_y v_n))$. For s fixed, the maps $y \mapsto s\theta(s,y,d_y u_0)$ and $y \mapsto s\theta(s,y,d_y v_n)$ have Lipschitz constant $\leqslant |s| K_1 K$. If we choose $\epsilon_1 > 0$ such that $\epsilon_1 K_1 K < 1$ and we set $2r_0 = (1 - \epsilon_1 K_1 K) r$, we can apply the Lipschitz form of the implicit function theorem to conclude that, for $s \in]-\epsilon_1, \epsilon_1[$, each one of the maps $y \mapsto sy + \theta(s,y,d_y u_0)$ and $y+ \mapsto s\theta(s,y,d_y v_n)$ is a homeomorphism from $\mathring{B}(x,r)$ onto some open subset of $\mathbb{R}^k$ that contains the open ball of radius $2r_0$ around respectively $x + s\theta(s,x,d_x u_0)$ and $x + s\theta(s,x,d_x v_0^n)$. Since $d_x u_0$ and $d_x v_0^n$ are both of norm $\leqslant A$ and θ is continuous we can find $\epsilon_0 \in]0, \epsilon_1]$ such that $s\theta(s,x,d_x u_0)$ and $s\theta(s,x,d_x v_0^n)$ are both of norm $< r_0$ for $|s| < \epsilon_0$. It follows easily that with these choices we do satisfy both (i) and (ii) above. Property (iii) follows from the last part of the Lipschitz form of the implicit function theorem, and what we just imposed on $s\theta(s,x,d_x u_0)$ and $s\theta(s,x,d_x v_0^n)$.

We now return to the case where M is a general manifold.

It follows from (i) above that $\mathrm{Id}_{\mathbb{R}} \times \pi^*$ restricted to the propagated graph $\mathcal{U}_0$ is a local homeomorphism in a neighborhood of $\{0\} \times \mathrm{Graph}(du_0)$. Using the same topological argument as before (see [La, pages 109–110]) there is an open neighborhood $\hat{U}$ of $\{0\} \times \text{Graph}(du_0)$ in $\mathcal{U}_0$ such that $\text{Id}_{\mathbb{R}} \times \pi^*$ induces a homeomorphism from $\hat{U}$ onto some open set U. This of course finishes the first part of the proof.

We want now to show that for each $q_0 \in V$ there is an open neighborhood N_{q_0} of q_0 in V and some $\epsilon_{q_0} > 0$ such that we can find a C^1 function $u^{q_0} :]-\epsilon_{q_0}, \epsilon_{q_0}[\times N_q \to \mathbb{R}$, with $]-\epsilon_{q_0}, \epsilon_{q_0}[\times N_{q_0} \subset U$, and

$$\forall (t,q) \in]-\epsilon_{q_0}, \epsilon_{q_0}[\times N_{q_0}, \sigma(t,q) = (t, q, \frac{\partial u^{q_0}}{\partial q}(t,q))$$

$$\text{and} \quad \frac{\partial u^{q_0}}{\partial t} = -H((t, q, \frac{\partial u^{q_0}}{\partial q}(t,q)),$$

where $\sigma : U \to \hat{U}$ is the inverse of $Id_{\mathbb{R}} \times \pi^* | \hat{U}$. To show this we apply (ii) above to find some open sets $O, N_{q_0} \subset V$, some ϵ_{q_0}, and a sequence of C^∞ functions $v_0^n : O \to \mathbb{R}$ with

a) the sequence $v_0^n : O \to \mathbb{R}$ converges uniformly to $u_0 | O$

b) each of the propagated graphs $\mathcal{V}_n^0$ of the dv_0^n contains an open set $\hat{U}^n$ such that $Id_{\mathbb{R}} \times \pi^*$ induces a homeomorphism onto some open subset U^n of $\mathbb{R} \times M$ with $U^n \supset]-\epsilon_{q_0}, \epsilon_{q_0}[\times N_q$;

c) if we call $\sigma_n : U^n \to \hat{U}^n$ the inverse of $Id_{\mathbb{R}} \times \pi^* | \hat{U}^n$ then σ_n converges uniformly to σ on $]-\epsilon_{q_0}, \epsilon_{q_0}[\times N_q$.

By Lemma 5.2 above for each n we can find a C^3 (in fact C^∞) function $v^n :]-\epsilon_{q_0}, \epsilon_{q_0}[\times N_{q_0}$ such that for each $(t,q) \in]-\epsilon_{q_0}, \epsilon_{q_0}[\times N_{q_0}$

$$\sigma_n(t,q) = (t, q, \frac{\partial v^n}{\partial q}(t,q)) \text{ and } \frac{\partial v^n}{\partial t}(t,q) = -H((t, q, \frac{\partial v^n}{\partial q}(t,q)),$$

and $v^n(0,q) = v_0^n(0,q)$, for each $q \in N_{q_0}$. By condition a) and b) above it follows that the sequence v^n converges uniformly in the C^1 topology to the required function $u^{q_0} :]-\epsilon_{q_0}, \epsilon_{q_0}[\times N_{q_0}$. Remark that two such local solutions corresponding to the points q_0 and q_0' are equal on the intersection of their domain of definition because this intersection is of the form $]-\epsilon, \epsilon[\times N$ with $N = N_{q_0} \cap N_{q_0'}$ and $\epsilon = \inf(\epsilon_{q_0}, \epsilon_{q_0'})$, the two functions coincide on $\{0\} \times V$, and they have the same derivative which is entirely determined by the section σ. We can then define u on $\bigcup_{q \in V}]-\epsilon_q, \epsilon_q[\times N_q$ by $u |]-\epsilon_q, \epsilon_q[\times N_q = u^q$. □

Remark 5.5.— A crucial step in the proof above was to show that in a neighborhood of $\{0\} \times V$ the propagated graph $\mathcal{U}_0$ is in fact a graph above some open set of the base for the projection $\mathrm{Id}_{\mathbb{R}} \times \pi^* : \mathbb{R} \times T^*M \to \mathbb{R} \times M$. One can ask what such a property implies when u_0 is merely C^1, compare with [Li, Remark 1.1, pages 15–16, and Proposition 15.1, page 265]. The following theorem settles this question.

THEOREM 5.6.— *Let $u_0 : V \to \mathbb{R}$ be a C^1 function, where V is some open set in M. Suppose that the projection $\mathrm{Id}_{\mathbb{R}} \times \pi^*$ restricted to the propagated graph $\mathcal{U}_0$ is a local homeomorphism in a neighborhood of $\{0\} \times \mathrm{Graph}(du_0)$. If the Lagrangian L is almost complete, then the function u_0 is in fact $\mathrm{C}^{1,1}$.*

Proof.— Since this is a local result, we can assume that $M = \mathbb{R}^k$, that $\bar{V}$ is compact, and that u_0 is Lipschitzian and bounded in V. For future reference, we will denote by K a Lipschitz constant for u_0 on V. We can also assume, restricting V if necessary, that there exists a $\delta > 0$ such that the projection $\mathrm{Id}_{\mathbb{R}} \times \pi^*$ induces a (global) homeomorphism from $\mathcal{U}_0 \cap]-\delta, \delta[\times T^*M$ onto some open subset U of $]-\delta, \delta[\times M$.

Since $\bar{V}$ is compact, for each $A \geqslant 0$, we can find a constant $C(A)$ such that

$$\forall t \in [-1,1], \forall (q,v) \in TV, L(t,q,v) \geqslant A\|v\| + C(A). \tag{5.1}$$

We use the well-known Lax-Oleinik formula to show that there exists a solution of the Hamilton-Jacobi equation, see for example [Be, Theorem 5.1 page 66].

We will need the notion of absolutely continuous curve. A curve $\gamma : [a,b] \to \mathbb{R}^k$ is said to be absolutely continuous, if the derivative $\gamma'(s)$ exists almost everywhere, the integral $\int_a^b \|\gamma'(s)\|\, ds$ is finite, and $\gamma(t) = \int_a^t \gamma'(s)\, ds$, for each $s \in [a,b]$.

We define a function $u : [-1,1] \times V \to \mathbb{R}$ by

$$\forall q \in V, \forall t \in]0,1]; u(t,q) = \inf_{\gamma} \int_0^t L(s,\gamma(s),\dot{\gamma}(s))\, ds + u_0(\gamma(0)),$$

where the infimum is taken over all absolutely continuous curves $\gamma : [0,t] \to V$ with $\gamma(t) = q$;

$$\forall q \in V, \forall t \in [-1,0[, u(t,q) = \sup_{\gamma} u_0(\gamma(t)) - \int_t^0 L(s,\gamma(s),\dot{\gamma}(s))\, ds,$$

where the supremum is taken over all absolutely continuous curves $\gamma : [0,t] \to V$ with $\gamma(0) = q$; and

$$\forall q \in V, u(q,0) = u_0(q).$$

Claim. — The function u is indeed finite everywhere. There is an open neighborhood U' of $\{0\} \times V$ in $[-1,1] \times V$ such that $u|U'$ is continuous, and u restricted to $U' \setminus \{0\} \times V$ is locally Lipschitzian. Moreover, at each point $(t,q) \in U' \setminus \{0\} \times V$ where u is differentiable we have

$$(t, q, \frac{\partial u}{\partial q}(t,q)) \in \mathcal{U}_0 \text{ and } \frac{\partial u}{\partial t} = -H(t, q, \frac{\partial u}{\partial q}(t,q)). \tag{5.2}$$

Let us show how to finish the proof of the theorem using the claim. Since u is locally Lipschitzian on $U' \setminus \{0\} \times V$, it is differentiable almost everywhere. Using (5.2), and the fact that $\mathrm{Id}_{\mathbb{R}} \times \pi^*$ induces a (global) homeomorphism from $\mathcal{U}_0 \cap]-\delta, \delta[\times T^*M$ onto the open subset U, we see that the derivative map $(t,q) \mapsto d_{(t,q)}u$, which is defined almost everywhere can be continuously extended to $U \cap U'$. Since the map u restricted to $U \cap U' \setminus \{0\} \times V$ is locally Lipschitzian, we see that it is in fact C^1 on $U \cap U' \setminus \{0\} \times V$; since the subset $\{0\} \times V$ is contained in a hyperplane, we can then conclude that u is in fact C^1 on $U \cap U'$. By continuity of the derivative and (5.2), the function u is a C^1 solution of the Hamilton-Jacobi equation, it follows from Theorem 3.1 that u is $C^{1,1}$, but $u_0(q) = u(0,q)$, for $q \in V$. Hence, the function u_0 must also be $C^{1,1}$.

It remains to prove the claim. We will do that for points (t,q) with $t \geqslant 0$, leaving the case of negative t to the reader. The proof of the claim relies on arguments which are essentially well-known, see for example [Fl, Theorem 1, page 518] or [Be, §5 page 64].

If $\gamma : [0,t] \to V$ is a continuous piecewise C^1 curve, by integration of (5.1), we have

$$\int_0^t L(s, \gamma(s), \dot\gamma(s))\, ds \geqslant (K+1)\ell(\gamma) + C(K+1)t, \tag{5.3}$$

where $\ell(\gamma) = \int_0^t \|\gamma'(s)\|\, ds$ is the length of γ, and K is here as above the Lipschitz constant of u_0 on V. For a path such that $\gamma(t) = q$, using $|u_0(q) - u_0(\gamma(0))| \leqslant K d(q, \gamma(0)) \leqslant K\ell(\gamma)$, we obtain

$$\int_0^t L(s, \gamma(s), \dot\gamma(s))\, ds + u_0(\gamma(0)) \geqslant u_0(q) + \ell(\gamma) + C(K+1)t. \tag{5.4}$$

This gives

$$u(q,t) \geqslant u_0(q) + C(K+1)t.$$

If we consider the constant path $\gamma_q : [0,t] \to V, s \mapsto q$, and set $A = \sup\{L(t,q,0) \mid t \in [-1,1], q \in \bar{V}\} < \infty$, we find

$$u(q,t) \leqslant At + u_0(q).$$

It follows that

$$|u(q,t) - u_0(q)| \leqslant t \max(|A|, |C(K+1)|).$$

This shows that u is finite everywhere, and continuous at each point of $\{0\} \times V$, since u_0 is itself continuous.

We fix now $q_0 \in V$ and choose $r > 0$ such that $\bar{B}(q_0, 3r) \subset V$. Using what we have done above, to compute $u(t,q)$ we can restrict to absolutely continuous curves $\gamma : [0,t] \to V$ such that $\int_0^t L(s,\gamma(s),\dot\gamma(s))\,ds + u_0(\gamma(0)) \leqslant At + u_0(q)$. From (5.4), such a curve must satisfy $\ell(\gamma) \leqslant (A - C(K+1))t$. If we now fix $t_0 > 0$ such that $(A - C(K+1))t_0 \leqslant r$, we find that for $t \in [0,t_0]$ and $q \in \bar{B}(q_0, r)$, we have

$$u(t,q) = \inf_{\gamma} \int_0^t L(s,\gamma(s),\dot\gamma(s))\,ds + u_0(\gamma(0)),$$

where the infimum is now taken over all absolutely continuous curves $\gamma : [0,t] \to \bar{B}(q_0, 2r)$ with $\gamma(t) = q$. Since $\bar{B}(q_0, 2r)$ is compact, Tonelli's theorem, see [Cl, page 30], implies that, for each $q \in \bar{B}(q_0, r)$, there exits an absolutely continuous curve $\gamma_{(t,q)} : [0,t] \to \bar{B}(q_0, 2r)$ with $\gamma_{(t,q)}(t) = q$, and such that

$$u(t,q) = \int_0^t L(s,\gamma_{(t,q)}(s),\dot\gamma_{(t,q)}(s))\,ds + u_0(\gamma_{(t,q)}(0)).$$

By definition of the function u, such a curve minimizes action among all absolutely continuous curves with values in V and having the same endpoints. Thus such a $\gamma_{(t,q)}$ is what is called a local minimizer. Although generally for time dependent Lagrangians an absolutely continuous local minimizer is not necessarily C^2, this is indeed the case when the Lagrangian is almost complete. In fact, the classical argument (see for example [Ma, page 175]) that shows that absolutely continuous local minimizers are C^2 extremal curves for a *complete* Lagrangian works also for almost complete Lagrangians.

It follows that each curve $\gamma_{(t,q)}$ is a C^2 extremal curve for the Lagrangian. Using the first variation formula, and the fact that $u(q,t)$ is defined as an infimum, it is not difficult to show that the derivative of u_0 at $\gamma_{(t,q)}(0)$ is $\frac{\partial L}{\partial v}(0, \gamma_{(t,q)}(0), \dot\gamma_{(t,q)}(0))$. Since u_0 is Lipschitzian with Lipschitz constant K, we conclude that the points $(\gamma_{(t,q)}(0), \dot\gamma_{(t,q)}(0))$ are all contained in the compact subset $\mathcal{K}$ of TM defined by

$$\mathcal{K} = \{(\tilde q, v) \in TM \mid \tilde q \in \bar B(q_0, 2r), \|\frac{\partial L}{\partial v}(\tilde q, v)\| \leqslant K\}.$$

Let $t_0' > 0$ be such that $\phi_s(0, \tilde q, v)$ be defined for all $(s, \tilde q, v) \in [0, t_0'] \times \mathcal{K}$. If we consider only points $(t, q) \in [0, \min(t_0, t_0')]$ and numbers $s \in [0, t]$, we conclude that the points $(\gamma_{(t,q)}(s), \dot\gamma_{(t,q)}(s))$ are all contained in the same compact subset $\bigcup_{s\in[0,t_0']} \phi_s(\{0\}\times\mathcal{K})$. Using a similar (but simpler) argument than the one used in the proof of Theorem 3.2, relying basically on the fact that L is Lipschitzian on a neighborhood of the compact set $[0, \min(t_0, t_0')] \times \bigcup_{s\in[0,t_0']} \phi_s(\{0\}\times\mathcal{K})$, it can be shown that u is a Lipschitz map on each subset of the form $[\delta, \min(t_0, t_0')] \times \bar B(q_0, r)$ with $\delta > 0$. It remains to identify the derivative of u when it exists at a point $(t, q) \in [\delta, \min(t_0, t_0')] \times \bar B(q_0, r)$. It follows from the first variation formula, and the definition of u, that $\frac{\partial u}{\partial q}(t, q) = \frac{\partial L}{\partial v}(t, q, \dot\gamma_{(t,q)}(t))$. Since the curve $\gamma_{(t,q)}$ is an extremal curve for the Lagrangian L, we have $(q, \dot\gamma_{(t,q)}(t)) = \phi_t(0, \gamma_{(t,q)}(0), \dot\gamma_{(t,q)}(0))$, hence, by the definition (2.4) of ϕ_s^*, we obtain $(q, \frac{\partial u}{\partial q}(t, q)) = (q, \frac{\partial L}{\partial v}(t, q, \dot\gamma_{(t,q)}(t)) = \phi_t^*(0, \gamma_{(t,q)}(0), \frac{\partial L}{\partial v}(0, \gamma_{(t,q)}(0), \dot\gamma_{(t,q)}(0))$. By what we saw above, we have the equality $\frac{\partial L}{\partial v}(0, \gamma_{(t,q)}(0), \dot\gamma_{(t,q)}(0)) = d_{\gamma_{(t,q)}(0)}u_0$, therefore $(t, q, \frac{\partial u}{\partial q}(t, q))$ is indeed in the propagated graph $\mathcal{U}_0$. To compute the partial derivative $\frac{\partial u}{\partial t}(t, q)$, we first observe that the definition of u as an infimum forces the equality

$$\forall t' \in [0, t], u(t', \gamma_{(t,q)}(t')) = \int_0^{t'} L(s, \gamma_{(t,q)}(s), \dot\gamma_{(t,q)}(s))\, ds + u_0(\gamma_{(t,q)}(0)).$$

Taking derivatives at $t' = t$, we obtain

$$\frac{\partial u}{\partial t}(t, q) + \frac{\partial u}{\partial q}(t, q)[\dot\gamma_{(t,q)}(t)] = L(t, \gamma_{(t,q)}(t), \dot\gamma_{(t,q)}(t))).$$

But we have already seen that $\frac{\partial u}{\partial q}(t, q) = \frac{\partial L}{\partial v}(t, q, \dot\gamma_{(t,q)}(t))$. Since $H(t, q, p) = \frac{\partial L}{\partial v}(t, q, v)[v] - L(t, q, v)$, when $p = \frac{\partial L}{\partial v}(t, q, v)$, we indeed obtain the equality $\frac{\partial u}{\partial t}(t, q) = -H(t, q, \frac{\partial u}{\partial q}(t, q))$. □

It is possible to give a particular geometric formulation of some of the results given here.

THEOREM 5.7.— *Suppose that M is a Riemannian manifold, and $N \subset M$ is a C^1 codimension 1 submanifold, and assume that the submanifold topology is the same as the topology induced from M. We denote by $d_N : M \to \mathbb{R}$ the distance function $d_N(q) = d(q, N) = \inf_{q' \in N} d(q, q')$. The distance function d_N is differentiable at every point of $V \setminus N$, where V is some open neighborhood of N in M if and only if N is $C^{1,1}$. In that case, the function d_N itself is $C^{1,1}$.*

It follows from [PR] that d_N is $C^r, r \geqslant 2$ on $V \setminus N$, where V is some open neighborhood of N in M, if and only if N is itself C^r. This is of course false for $k = 1$ by Theorem 5.7. The case where M is 2-dimensional can be proven in a different way see [Ze, §1.1].

Proof of Theorem 5.7.— The argument is essentially local, so we can assume that M is an open subset of $\mathbb{R}^k$ and that N is closed in M, connected and orientable. Replacing M by a neighborhood of N, we may assume that $M \setminus N = O_+ \cup O_-$, where O_+ and O_- are two disjoint open subsets of M. We then define $u : M \to \mathbb{R}$ by $u|N = 0$, $u|O_+ = d_N$, $u|O_- = -d_N$. It is clear that u is Lipschitzian with Lipschitz constant 1, and it is differentiable at $q \in M \setminus N$ if and only if d_N is. Suppose that u is differentiable at $q \in O_+$ (resp. $q \in O_-$) if q is close enough to N, there exists a point $q' \in N$ such that $d_N(q) = d(q', q)$, and a geodesic curve $\gamma : [0, d_N(q)] \to M$, parametrized by arc-length, such that $\gamma(0) = q'$ and $\gamma(d_N(q)) = q$. It is clear that we must have $u(\gamma(s)) = s$ (resp. $u(\gamma(s)) = -s$), and that $\dot\gamma(0)$ is a unit vector orthogonal to $T_{q'}N$ and pointing toward O_+ (resp. O_-). Differentiating $u(\gamma(s)) = s$ (resp. $u(\gamma(s)) = -s$), at $s = d_N(q)$, we see that $d_q u(\dot\gamma(d_N(q)) = 1$ (resp. $d_q u(\dot\gamma(d_N(q)) = -1$. Since both $\|\dot\gamma(d_N(q))\| \leqslant 1$ and $\|d_q u\| \leqslant 1$, we must have that $\|d_q u\| = 1$, and $-\dot\gamma(d_N(q))$ (resp. $-\dot\gamma(d_N(q)))$ is the only vector $v \in T_q M$ such that $d_q u = \langle v, \cdot \rangle$. When u is differentiable on $V \setminus N$, where V is some open neighborhood of N in M, it then follows that the derivative $q \mapsto d_q u$ extends continuously to N. The value at $q' \in N$ being $\langle \nu(q'), \cdot \rangle$, where $\nu(q') \in T_{q'}M$ is the unit vector perpendicular to $T_{q'}N$ and pointing to O_+. Since u has a derivative of norm 1 everywhere on V, it must be $C^{1,1}$, and therefore $q' \mapsto \nu(q')$ is a locally Lipschitz map on N, hence N itself is $C^{1,1}$.

We let the reader prove, using the Lipschitz version of the implicit function theorem, that when N is $C^{1,1}$, the exponential map of the Riemannian metric restricted to the normal bundle $\nu(N)$ of N in M is a homeomorphism from an open neighborhood of the zero section in $\nu(N)$ onto some open neighborhood of N in M. This in turn implies that the derivative of the Lipschitz map u, which exists almost everywhere by Rademacher theorem, extends continuously to a neighborhood of N, hence u must be C^1. □

Remark 5.8.— 1) The argument given above works even when N is merely a C^0 codimension 1 submanifold. In fact localizing as in the proof above, and using some arguments from Algebraic Topology (essential equivalent to Jordan's theorem) we see that $M \setminus N = O_+ \cup O_-$, where O_+ and O_- are two disjoint open subsets of M satisfying $N \subset \bar{O}_+$ and $N \subset \bar{O}_-$. If d_N is differentiable on $M \setminus N$, it can be shown that if $q \in N$ is fixed, then for any $r > 0$ small enough there exists $q_+ \in O_+$ and $q_- \in O_-$ such that $d_N(q_+) = d(q, q_+) = r$ and $d_N(q_-) = d(q, q_-) = r$. Moreover, if $q_+ \in O_+$ and $q_- \in O_-$ are such that $d_N(q_+) = d(q, q_+)$ and $d_N(q_-) = d(q, q_-)$, as soon as q, q_+, q_- are close enough to define a geodesic triangle this geodesic triangle must be degenerate with a flat angle at q, since the geodesic between q_+ and q_- must intersect N. It is then not difficult to see that the tangent unit vector at q to the geodesic joining q and q_+ depends only on q and not on the choice of $q_+ \in O_+$. This defines a normal vector at every point of $q \in N$. The arguments in the proof above can then be adapted to show that u is $C^{1,1}$, with a derivative of norm 1 everywhere. It follows that $N = u^{-1}(0)$ must be a $C^{1,1}$ submanifold.

2) If N is a closed subset of the Riemannian manifold M, its caustic $\mathcal{C}_N$ is the subset

$$\mathcal{C}_N = \{q \in M \mid \exists q_1, q_2 \in N, q_1 \neq q_2, d_N(q) = d(q, q_1) = d(q, q_2)\}.$$

When N is a C^0 codimension 1 submanifold, it is not difficult to obtain from what we have seen above that $N \setminus \bar{\mathcal{C}}_N$ is the (open) set of points $q \in N$ such that there is an open neighborhood O_q of q in N which is a $C^{1,1}$ submanifold of M. In particular, if N is nowhere $C^{1,1}$, then the caustics accumulate everywhere on N.

Bibliography

[Be] Benton Jr. (S. H.), *The Hamilton-Jacobi Equation: A Global Approach,* Academic Press, New York, San Francisco, and London, Mathematics in Science and Enggineering **131**, (1977).

[BP] Bialy (M.) & Polterovitch (L.), Hamiltonian Diffeomorphisms and Lagrangian Distributions, *Math. Zeitschrift* **2**, p. 173-210 (1992).

[CC] Caffarelli (L. A.) & Cabré (X.), *Fully Nonlinear Elliptic Equations,* Colloquium Publications **43**, AMS, Providence (1995).

[Cl] CLARKE (F. H.), Methods of Dynamic and Nonsmooth Optimization, CBMS-NSF Regional Conference Series in *Applied Mathematics* **57**, SIAM, Philadelphia (1989).

[Es] ESCHENBURG (J. H.), Horospheres and the Stable Part of the Geodesic Flow, *Math. Zeitschrift* **153**, p. 237-251 (1977).

[Fa] FATHI (A.), Théorème KAM faible et Théorie de Mather sur les systèmes lagrangiens, *C. R. Acad. Sci. Paris, Série I* **324**, p. 1043-1046 (1997).

[Fl] FLEMING (W. H.), The Cauchy Problem for a nonlinear First Order Partial Diferential Equation, *J. Diff. Equ.* **5**, p. 515-530 (1969).

[He] HERMAN (M. R.), Inégalités à priori pour des tores lagrangiens invariants par des difféomorphismes symplectiques, *Publ. Math. IHES* **70**, p. 47-101 (1989).

[Ki] KISELMAN (C. O.), Regularity Classes for Operations in Convexity Theory, *Kodai Math. J* **15**, p. 354-374 (1992).

[Kn] KNIEPER (G.), Mannigfaltigkeiten Ohne Konjugierte Punkte, *Bonner Mathematische Shriften* **168**, (1986).

[La] LANG (S.), *Differential and Riemannian Manifolds*, Third Edition, Graduate Texts in Mathematics **160**, Springer, New York, Berlin, Heidelberg (1995).

[Li] LIONS (P. L.), *Generalized Solutions of Hamilton-Jacobi Equations*, Research Notes in Mathematics **69**, Pitman, London (1982).

[Ma] MATHER (J. N.), Action Minimizing Meausures for Positive Definite Lagrangian Systems, *Math. Z.* **207**, p. 169-207 (1991).

[PR] POLY (J-B.) & RABY (G.), Fonction distance et singularités, *Bull. Sci. Math.* **108**, p. 187-195 (1984).

[Ze] ZEGHIB (A.), Lipschitz Regularity in some Geometric Problems, *preprint* ENS-Lyon 2000.