ESAIM: Control, Optimisation and Calculus of Variations Will be set by the publisher
URL: `http://www.emath.fr/cocv/`

REGULARITY AND VARIATIONALITY OF SOLUTIONS TO HAMILTON-JACOBI EQUATIONS. PART I: REGULARITY (2ND EDITION) *

Andrea C. G. Mennucci

Abstract. We formulate an Hamilton–Jacobi partial differential equation

$$H(x, Du(x)) = 0$$

on a n dimensional manifold M, with assumptions of convexity of $H(x, \cdot)$ and regularity of H (locally in a neighborhood of $\{H = 0\}$ in T^*M); we define the "*min* solution" u, a generalized solution; to this end, we view T^*M as a *symplectic manifold*.

The definition of "*min* solution" is suited to proving regularity results about u; in particular, we prove in the first part that the closure of the set where u is not regular may be covered by a countable number of $n-1$ dimensional manifolds, but for a $\mathcal{H}^{n-1}$ negligeable subset. These results can be applied to the cutlocus of a C^2 submanifold of a Finsler manifold.

1991 Mathematics Subject Classification. 49L25, 53C22, 53C60.

The dates will be set by the publisher.

1. Introduction

In this article we will study the Dirichlet type Hamilton–Jacobi PDE

$$\begin{cases} H(x, Du(x)) = 0 & \text{in } M \setminus K \\ u(x) = u_0(x) & \text{when } x \in K. \end{cases} \tag{1.1}$$

where M is a smooth manifold, u_0 is a continuous real function on K, H is a continuous real function on T^*M, and $K \subset M$ is a closed subset; when we will write "$K \in C^r$", though, K will be an embedded submanifold.

The main aim of this first part is to prove results on the regularity of u.

Let Ω be the subset of M where the generalized solution u is defined; assume that u is continuous.

We sketch here the type of results that we are interested in, for the benefit of this introductory discussion, without providing complete definitions.

Let $\Sigma_u \subset \Omega$ be the set where u is not differentiable; if $u_0, K \in C^2$, let $\Gamma \subset \Omega$ be the set of *conjugate points*.

Keywords and phrases: Hamilton–Jacobi equations, cutlocus, conjugate points
Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy
* *This 2nd version corrects one error in the in 2004's version*

We want to formulate a *structured regularity* result, in which we divide the set Ω in three subsets, and state that

- u is as regular as the data H, K, u_0 on $\Omega \setminus (\Sigma_u \cup \Gamma)$
- $\Sigma_u \setminus \Gamma$ can be covered locally by finitely many submanifolds that are as regular as H, K, u_0
- Γ is a closed *rectifiable set*, that is, there exists $N \subset \Gamma$ such that $\Gamma \setminus N$ can be covered by countably many regular $n-1$ dimensional manifolds, and N has zero $n-1$ dimensional measure.

There are various variants of the concept of *rectifiable set*; for example, if $\Gamma \setminus N$ can be covered by countably many C^r regular $n-1$ dimensional manifolds and $\mathcal{H}^{n-1}(N) = 0$, then we will say that Γ is C^r*-*$\mathcal{H}^{n-1}$*-rectifiable*; we will be more specific when needed. (For a precise definition of the Hausdorff measure $\mathcal{H}$ and Hausdorff dimension on a manifold M, see sec. A.3). See [7,20] for a complete discussion of the notion of rectifiability.

As a consequence of the above structured regularity, there follows that u is in the space SBV^2 of functions whose second derivative can be expressed as a measure $D^2u = D^2_a u + D^2_j u$, where $D^2_a u$ is absolutely continuous w.r.t a n-dimensional Hausdorff measure on M, and $D^2_j u$ is absolutely continuous w.r.t a $(n-1)$-dimensional Hausdorff measure restricted to $\Sigma_u \cup \Gamma$.

It is possible to obtain a result similar to the above in quite general hypotheses. Suppose that (1.1) has a viscosity solution u on M and that this solution u is *semiconcave* in $M \setminus K$ (see 2.2); then, by results in [1], we know that Σ_u is C^2-$\mathcal{H}^{n-1}$-rectifiable; moreover, by semiconcavity, we know that u is C^1 in the open set $M \setminus (K \cup \overline{\Sigma}_u)$. Unfortunately the set $\overline{\Sigma}_u$ can be in general quite larger than Σ_u (see 1.5), unless some further regularity is imposed on K and u_0.

If instead K, u_0 are regular enough then $\Sigma_u \cup \Gamma = \overline{\Sigma}_u$; and we will prove that these sets are rectifiable.

Two subcases of the above equation have been studied before: the Cauchy type equation, and the eikonal equation. We outline the known results.

1.1. Cauchy type Hamilton–Jacobi equation

The corresponding Cauchy type Hamilton–Jacobi equation

$$\begin{cases} \frac{\partial}{\partial t} w(t,x) + H'(t,x,D_x w(t,x)) = 0 & \text{for } t > 0,\ x \in M' \\ w(0,x) = w_0(x) & \forall x \in M'. \end{cases} \tag{1.2}$$

has been studied by Cannarsa, Sinestrari, M. in [4], for the case $M' = \mathbb{R}^n$ (but the proofs in [4] may be easily adapted to the general case of having M a manifold, using the symplectic formalism). By assuming that H' is strongly convex in the argument $D_x w$, that $H' \in C^{R+1}$, that $w_0 \in W^{1,\infty} \cap C^{R+1}(M')$, and $R \geq 2$, in thm 4.10 and 4.12 in §4.3 in [4] it was proven that

Theorem 1.1 (Ths. 4.10, 4.12 in [4])**.** *the set Γ is $\mathcal{H}^{n-1}$ rectifiable.*

and in thm 4.17 in §4.3 in [4] it was proven that

Theorem 1.2 (Th. 4.17 in [4])**.** *the set $\Gamma \setminus \Sigma_u$ has negligeable $\mathcal{H}^{n-1+2/R}$ measure.*

these results are improved in Proposition 5.1 in 5.1 in this paper.

1.2. Eikonal equation, cutlocus

Suppose that M is a Riemannian manifold with a norm on TM that is dual to the norm $|p|$ on T^*M; let $H(x,p) = |p|^2 - 1$; assume that M is complete.

When $u_0 = 0$, we obtain the *eikonal equation*

$$\begin{cases} |du(x)|^2 - 1 = 0 & \text{on } M \setminus K \\ u = 0 & \text{on } K \end{cases} \tag{1.3}$$

This problem is a very important example; the reader can use it as a guideline to understanding the results we propose.

Problem (1.3) has been extensively studied in many papers. Recently in [14] Mantegazza and Mennucci have proved that the distance function d_K to the set K

$$d_K(x) = d(x, K) \stackrel{\text{def}}{=} \min_{z \in K} d(x, z) \tag{1.4}$$

is the unique viscosity solution to (1.3), in the class of continuous functions bounded from below (thm 3.1), and that, if K is a C^3 manifold, then d_K enjoys the kind of regularity that we discussed above: in thm 3.4 it is shown that d_K is semiconcave in $M \setminus K$, for any closed set K, and then d_K is C^1 in the open set $M \setminus (K \cup \overline{\Sigma}_u)$ (thm 3.5), and

Theorem 1.3 (Th. 4.7 [14])**.** *If $K \in C^r$ with $r \geq 3$, then Γ is a C^{r-2}-$\mathcal{H}^{n-1}$-rectifiable set.*

This result is slightly improved in section 5.2.

In 2002 C. Pignotti in [19] proved the above result for a class of optimal exit time problems.

1.2.1. *The cutlocus*

If K is a C^2 regular embedded submanifold, then the conjugate points Γ are called *optimal focal points*; moreover

$$\Sigma_u \cup \Gamma = \overline{\Sigma}_u = \mathrm{Cut}_K \tag{1.5}$$

where Cut_K is the *cutlocus* of K, that is the locus of points where the geodesics starting orthonormally from K stop being optimal for the distance d_K. The above sets are known by many different names: the *skeleton*, the *ridge*, the *set of medial axes*.

If K is not a C^2 embedded submanifold, then the set Γ cannot be defined, and in this case

$$\overline{\Sigma}_u \supset \mathrm{Cut}_K$$

moreover, the cutlocus may fail to be a closed set, as in this simple example

Example 1.4. *Let $M = \mathbb{R}^2$ and $K \stackrel{def}{=} \{(0,0)\} \cup \{(0, 2^n), n \in \mathbb{Z}\}$; then*

$$Cut_K = \{(x, 2^n 3), x \in \mathbb{R}, n \in \mathbb{Z}\}, \qquad \overline{Cut_K} = Cut_K \cup \{(x, 0), x \in \mathbb{R}\}.$$

In section §3 of the paper [14] there is also an example

Example 1.5 (§3 [14])**.** *There is a $C^{1,1}$ curve $K \subset \mathbb{R}^2$ such that $\overline{\Sigma}_{d_K}$ has positive Lebesgue measure. K is the border of a convex set.*

1.2.2. *Other regularity results*

More recently, in [10], Itoh and Tanaka prove that, if K is a smooth submanifold, then

Theorem 1.6 (Th. [10] A)**.** *If $\lambda_k(x, v)$ is the time it takes for a geodesic starting from $x \in K$ with initial velocity v to reach its k-th focal point; then λ_k is locally Lipschitz*

The theorem 4.1 in this paper counts the focal points by rank, and obtain a better regularity result for the focal points of problem (1.1); the result 4.1 is though only a local result, so it does not imply thm. A in [10]. Itoh and Tanaka prove also that

Theorem 1.7 (Th. [10] B)**.** *the distance to the cutlocus is locally Lipschitz;*

it is also interesting to note that, in the above example 1.5, the distance to the cutlocus is not Lipschitz (since Σ_u has infinite length in the region inside K).

In [12], Li and Nirenberg have proven the same result of 1.7 , and have asserted that the minimum regularity for this to happen is $K \in C^{2,1}$; they have also extended it to the problem (1.1) and to the distance in Finsler spaces.

This interesting result is not covered by our theorems on problem (1.1) (see the discussion in 5.2).

Note that in general $\mathcal{H}^n(\mathrm{Cut}_K) = 0$, regardless of the regularity of K: this is proved in proposition 14 of the paper [18] on regularity results for Cauchy horizons in Lorentzian manifolds.

1.3. **Calculus of variation, viscosity solutions, vs *min* solutions**

Before we end the introduction, we would like to describe the framework and origin of these studies.

Consider, as an example, a classical problem in *Calculus of variation*: consider a function $L : TM \to \mathbb{R}$ and, for any a fixed $t > 0$, the problem

$$W(t,x) = \inf \int_0^t L(\xi,\dot{\xi})dt + u_0(\xi(0)) \tag{1.6}$$

where the infimum is to be found in the class of all absolutely continuous curves $\xi : [0,t] \to M$ such that $\xi(0) \in K$, $\xi(t) = x$; $W(t,x)$ is called the *value function*.

Alternatively, we may consider an *optimal exit time problem*: take a closed subset $\mathcal{U}$ of TM, such that, for any $x \in M$, the intersection $T_xM \cap \mathcal{U}$ is convex; minimize

$$W_{\mathcal{U}}(x) \stackrel{\text{def}}{=} \inf \int_0^t L(\xi,\dot{\xi})dt + u_0(\xi(0)) \tag{1.7}$$

in the class of all $t \geq 0$ and all absolutely continuous curves $\xi : [0,t] \to M$ such that $\xi(0) \in K$, $\xi(t) = x$ and moreover $(\xi,\dot{\xi}) \in \mathcal{U}$ for all times.

Under some reasonable hypothesis (including convexity and superlinearity of $L(x,\cdot)$) there exists a dual Hamiltonian H, such that W is a viscosity solution to the problem (1.2) (see e.g. §1.8 in [8], or §1.4 in [13]); by adding some hypotheses, we may also assume that the above problem (1.6) has a minimum path ξ^*, for any x.

Similarly, the minimum $W_{\mathcal{U}}(x)$) of a problem like (1.7) provides usually a viscosity solution to a dual problem (1.1) (see chap 5 in [13], or chap 1 in [8]).

We can moreover say that, if L is regular enough, then for any $x \in M$, the minimal curve $\xi^*(s)$ to the problem $W(t,x)$ is a *"characteristic"*, that is, it solves the Euler equations

$$\frac{d}{dt}\frac{\partial L}{\partial v}(\xi,\dot{\xi}) = \frac{\partial L}{\partial x}(\xi,\dot{\xi})$$

associated to the integrand $L(x,v)$ (and, if u_0, K are regular, then $\xi^*(s)$ leaves K with a prescribed angle); then, in the formula (1.6) defining $W(t,x)$, we may decide to search the minimum only in the class of all characteristics ending in x: this leads to the definition of the *min* solution u (*min* solution that will be defined in (3.5)); the *min* solution is then a kind of generalized solution, loosely based on "Cauchy's method of characteristics".

So, where is the interest in studying *min* solutions? The interest is twofold

- ***(Variational H-J problems)*** it may be interesting to "reverse" the previous reasoning: given the Hamilton–Jacobi problem (1.1), is it possible to find a problem like (1.6), or (1.7), that admits minimum path ξ^* for any x, and such that the *value function* $W(1,x)$ (or resp. $W_{\mathcal{U}}(x)$) is a viscosity solution (and also a *min* solution) to (1.1)?

  If it is possible, we will say that the problem (1.1) is *variational*.

  A necessary condition is that the *min* solution should be defined everywhere.

  A set of sufficient conditions will be shown in the second part [16]; other results may be found, e.g., in chap 5 in [13].

- ***(Regularity of (possibly non variational) H-J problems)*** the studies on regularity of viscosity solutions done in [4] and [14] were based mainly on properties of the characteristic flow associated to H; these properties are not related to the variationality of the problem [1]; the first part of this paper will study the regularity of the Hamiltonian flow, and deduce results on the regularity of the *min* solution u.

The comparison of "*min* solutions" and "viscosity solutions" suggests also a nice and geometrical interpretation of what these are (see section 3.3)

2. Prelims

2.1. Notation

We fix some notations.

- M will be a connected borderless differentiable manifold of class C^∞ and of dimension n, $n \geq 2$;
- K will be a C^1-regular closed embedded submanifold of M of dimension k with $0 \leq k \leq \dim(M) - 1$;
- H will be a continuous real function defined on the cotangent bundle T^*M, $C^{1,1}$ in a neighborhood of $\{H = 0\}$; we will moreover add some kind of convexity hypothesis to the function $p \mapsto H(x,p)$
- and u_0 will be a C^0 real function defined on K.

R will be a natural number, and $\theta \in [0,1]$; the class $C^{(R,\theta)}$ will characterize the regularity of the following problems [2]; moreover, whenever we will talk about the regularity of H, by writing "$H \in C^{(R,\theta)}$", we will always mean that "$H \in C^{(R,\theta)}$ in a neighborhood of $\{H = 0\}$".

We will use the notation $p \cdot v$ to mean that a covector $p \in T_x^*M$ is applied to a vector $v \in T_xM$.

If $f : M \to \mathbb{R}$ is a regular function, we will write $df(x)$ or $Df(x)$ for its differential in the point x; if $g : \mathbb{R} \times M \to \mathbb{R}$ is a regular function, $g = g(t,x)$, we will write $\dot{g}$ for $\frac{\partial g}{\partial t}$ (and not g', which will be a different function).

2.2. Viscosity solutions

Now we introduce the definition of viscosity solutions of PDE on manifolds. As in the standard case $M = \mathbb{R}^n$, we begin with the definition of the following generalized differentials. Let Ω be an open subset of M.

Definition 2.1. *Given a continuous function $u : \Omega \to \mathbb{R}$ and a point $x \in M$, the* superdifferential *of u at x is the subset of T_x^*M defined by*

$$\partial^+ u(x) = \left\{ d\varphi(x) \,|\, \varphi \in C^1(M), \varphi(x) - u(x) = \min_M(\varphi - u) \right\}. \tag{2.1}$$

Similarly, the set

$$\partial^- u(x) = \left\{ d\psi(x) \,|\, \psi \in C^1(M), \psi(x) - u(x) = \max_M(\psi - u) \right\}$$

is called the subdifferential *of u at y. Notice that it is equivalent to replace the* max *(*min*) on all M with the maximum (minimum) in an open neighborhood of x in M.*

Definition 2.2. *We say that a continuous function u is a* viscosity solution *of equation*

$$H(x, Du(x)) = 0$$

[1] in particular, the definitions of "singular points" and "conjugate points", which in [4] where based on properties of an associated optimal control problem, have a natural equivalent definition that is based only on the properties of the characteristic flow (see sec. 4)

[2] when $\theta = 0$, the space $C^{(R,0)}$ will be identified with C^R. Note that, since we are working inside a generical differential manifold, the space $C^{(R,\theta)}$ is actually $C^{(R,\theta)}_{\mathrm{loc}}$, the space of C^R functions whose R-th derivatives are *locally* Hölder of exponent θ

in Ω *if for every* $x \in \Omega$,

$$\begin{cases} H(x,v) \le 0 & \forall v \in \partial^+ u(x), \\ H(x,v) \ge 0 & \forall v \in \partial^- u(x). \end{cases} \tag{2.2}$$

If only the first condition is satisfied (resp. the second), u is called a viscosity subsolution *(resp. a* viscosity supersolution*).*

The study of viscosity solutions on manifolds requires the development and usage of appropriate tools. See §2 in [14].

We will not use the concept of viscosity solutions directly; we will use a definition of "*min* solution", that is loosely based on the classical "Cauchy's method of characteristics".

A tool often used in connection with viscosity solutions is *semiconcavity*, that is defined as

Definition 2.3. *v is a semiconcave function in the open set A if, for any $x \in A$, there is a neighborhood $B \subset A$ of x, a choice of local coordinates $\phi : \mathbb{R}^n \to B$, with $\phi(0) = x$, and a C^2 function f on B such that $(v + f) \circ \phi$ is concave on the unit ball $B_1(0)$ in $\mathbb{R}^n$*

There is also another, more general, definition of semiconcavity: see [3], and references therein.

Part 1. *min* solutions

In this first part, we will define the Hamilton-Jacobi problem; we will impose some regularity hypotheses on it, so that we may define the *min* solution $u(x)$ (a kind of generalized solution) and study its regularity.

3. Setting of the problem

We consider T^*M as a symplectic manifold: we define the symplectic 2–form

$$\omega\Big((\dot{x},\dot{p}),\ (\dot{y},\dot{q})\Big) \stackrel{\text{def}}{=} \sum_i \dot{q}_i \dot{x}_i - \sum_i \dot{p}_i \dot{y}_i$$

and the duality $\omega^\#$ between TT^*M and TTM, given by

$$\omega^\#(\nu) \cdot \nu' = \omega(\nu, \nu') \quad \forall \nu' \in T^*M \quad .$$

We will also make use of the concept of *Lagrangian submanifold* Λ of (T^*M, ω), that is a n-dimensional submanifold such that, for any $y \in \Lambda$, for any two $\eta, \eta' \in T_y\Lambda$, $\omega(\eta, \eta') = 0$.

We define the *characteristic flow* [3]

$$(X(\cdot, z, q), P(\cdot, z, q), U(\cdot, z, q))$$

as the solution of the system of ordinary differential equations

$$\begin{cases} \dot{X}(s) = \frac{\partial H}{\partial p}(X(s), P(s)) \\ X(0) = z \\ \dot{P}(s) = -\frac{\partial H}{\partial x}(X(s), P(s)) \\ P(0) = q \\ \dot{U}(s) = P(s) \cdot \frac{\partial H}{\partial p}(X(s), P(s)) \\ U(0) = 0 \end{cases} \tag{3.1}$$

(X, P) is the *Hamiltonian flow* for the symplectic manifold for (T^*M, ω).

[3] the classical method of characteristics would require to define $\dot{\widetilde{U}}(s) = P(s) \cdot \frac{\partial H}{\partial p}(X(s), P(s))$ and $\widetilde{U}(0) = u_0(z)$; since H does not depend on u, then $\widetilde{U} = U + u_0(z)$ at all times: hence follows the definition (3.5) of the solution.

3.1. Hypotheses

We define

$$Z \stackrel{\text{def}}{=} \{H \le 0\} \stackrel{\text{def}}{=} \{(x,p) \mid H(x,p) \le 0\}$$

Hypotheses 3.1. *We will always suppose, in this first part, that*

(u_0K1): *K is a C^1-regular closed embedded submanifold of M of dimension k with $0 \le k \le \dim(M) - 1$, and u_0 is a C^1 real function defined on K, and*

(H1): *H is $C^{(1,1)}$ in a neighborhood of the zero set*

$$\{H = 0\} \stackrel{def}{=} \{(x,p) \mid H(x,p) = 0\}$$

in T^*M, and $\partial Z = \{H = 0\}$;

we will moreover suppose, where stated, that

(H2): *$H \in C^2$, and H is strongly convex in the p variable, that is, the Hessian is positive definite,*

$$\frac{\partial^2 H}{\partial pp}(x,p) > 0$$

for any x, p in a neighborhood of $\{H = 0\}$

We define also the sections

$$Z_x \stackrel{\text{def}}{=} \{p \in T_x^*M \mid H(x,p) \le 0\}$$

Note that the hypothesis (H2) implies that any connected component of Z_x is either a strictly convex set (with regular boundary) or a point (x,p), that is isolated in T_x^*M; the latter happening iff $\frac{\partial}{\partial p}H(x,p) = 0$.

3.2. *min* solutions

We define the u_0-*annihilator*

$$TK^{\perp u_0} \stackrel{\text{def}}{=} \{(z,p) \in T^*M \mid \ z \in K, p|_{T_zK} = du_0(z)\} \tag{3.2}$$

where we write $p|_{T_zK} = du_0(z)$ to mean that $\forall v \in T_zK,\ p \cdot v = du_0(z) \cdot v$.

$TK^{\perp u_0}$ is a submanifold of T^*M, of dimension n; if $u_0 \equiv 0$, $TK^{\perp u_0}$ is a Lagrangian submanifold (see [15], ex. 3.22).

We define the set $O \subset T^*M$ of covectors that are based on K and compatible with H and u_0

$$\begin{aligned} O \stackrel{\text{def}}{=} & TK^{\perp u_0} \cap \{H = 0\} = \\ = & \{(x,p) \in T^*M \mid \ x \in K, \quad H(x,p) = 0, \quad \forall v \in T_xK, p \cdot v = du_0(x) \cdot v\} \end{aligned} \tag{3.3}$$

(that is a closed subset of T^*M). In the following, we will often look at the flow (X, P) restricted to $\mathbb{R}^+ \times O$, and its derivatives as such; to emphasize this fact, we will use the variable $y = (z,q) \in O$ when necessary.

We define the *reachable set*

$$\Omega \stackrel{\text{def}}{=} \{x \in M \mid x = X(s,z,q) \text{ for } s \ge 0, (z,q) \in O\} \tag{3.4}$$

Note that this set is, a priori, not necessarily open.

We define the *min* solution u on Ω as

$$u(x) \stackrel{\text{def}}{=} \min_{\left\{\begin{array}{l} t \ge 0,\ (z,q) \in O \\ \text{s.t. } X(t,z,q) = x \end{array}\right.} U(t,z,q) + u_0(z) \tag{3.5}$$

We will use the *min* solution as a tool to study the regularity of generalized solutions; so

Hypotheses 3.2. *we will always assume that*

- ($\exists$**u**): *for any* $x \in \Omega$, *the above minimum* (3.5) *is attained and is not* $-\infty$; *and that*
- **(CC):** *the* compatibility condition: $\forall (z,q) \in O, t > 0$ *such that* $x = X(t,z,q) \in K$ *we have that* $U(t,z,q) \geq u_0(x) - u_0(z)$;

this last is actually only related to the second requirement we imposed in the definition of the problem (1.1), indeed

Proposition 3.3. *the hypothesis* (CC) *directly implies that* $u = u_0$ *on* K

Remark 3.4. *It is worthwhile to mention the relationship between the above definition and the theory of symplectic manifolds. Indeed, consider any open set* $A \subset M$ *where* u *is regular; consider the manifold* $\Lambda' \subset T^*M$ *that is the graph of the* 1*-form* du*:* Λ' *is a* Lagrangian submanifold *of* (T^*M, ω)*. In particular, since* Λ' *is the graph of the* 1*-form* du*,* [4] *then,* Λ' *is called* exact*, and* u *is called the* generating function *of* Λ *(see, e.g., [15], sec. 9.4, and ex. 3.50). Note that* Λ' *is subset of a larger Lagrangian manifold*

$$\Lambda \overset{def}{=} \{(X(t,y), P(t,y)) \mid t \in \mathbb{R}, y \in O\}$$

that is spanned by the flow (see 6.1) (X,P) *in* T^*M *(with initial conditions in* O*).* Λ *is* not *in general the graph of a form based on* M*: to study the regularity of the* min *solution, we will study how and why* Λ *is not exact (in proposition 4.4).*

3.3. Multi–valued solutions and criterion of choice

It is well known that *Cauchy's method of characteristics* provides a way to solve first order PDEs (see [5], [9, §VI.7]); the only relevant problem is the *ambiguity* of the solution, that is, which value $U(s,y)$ to choose for $u(x)$ when we have multiple choices of s, y such that $x = X(s,y)$; we may otherwise say that the method of characteristics defines the solution as a *multivalued function* $\lambda : \Omega \to \mathcal{P}(\mathbb{R})$

$$\lambda(x) \overset{\text{def}}{=} \{U(t,y) \mid \exists t \in \mathbb{R}, y \in O \text{ s.t. } \; X(t,z,q) = x\}$$

We point out a difference between the classical "method of characteristics", and our definition (3.5): in the former, there is no provision for having $t \geq 0$. For this reason, our *min* solutions will not be regular in a neighbourhood of K, in general. [5]

Sticking to the (more used) one–valued solutions, there is the mathematical problem that (even if $\Omega = M$) there would be in general *no* global regular solution, and, in contrast, an infinite number of *almost everywhere* solutions. The stalemate was solved by the introduction of the concept of *viscosity solutions*.

The fact that the viscosity solution v does often coincide with the *min* solution u points out the fact that the *viscosity solution* is obtained when we a priori decide for a *choice criterion*, to obtain some one–valued solution v and u from the multi–valued solution λ: v and u are (somehow) a *sheet* of v.

We summarize this through a simple example: let $M = \mathbb{R}$, and consider the eikonal problem

$$\begin{cases} |du(x)|^2 - 1 = 0 \\ u(0) = u(1) = 0 \end{cases} \tag{3.6}$$

(see (1.3)) the solution obtained by the method of characteristics is multivalued, and has 4 values for any x, as in frame 1 in figure 1; when we add the condition $t \geq 0$ to the method of characteristics, we obtain the solution

[4] which is obviusly closed, as prescribed by 3.25 in [15]

[5] The choice of having $t \geq 0$ was necessary, to have, as stated below, that u be a viscosity solution to (1.1) (under suitable hypotheses). See 4.14.

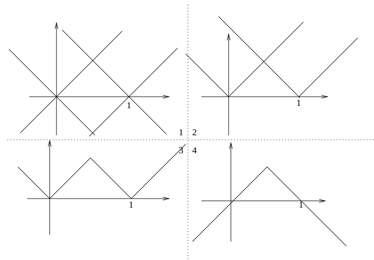

FIGURE 1. solutions to (3.6)

in frame 2; to obtain the *min* solution $u(x)$, we apply to frame 2 the "*min* choice criterion", namely, we choose the minimum solutions proposed by the characteristics, see frame 3; this $u(x)$ is a viscosity solution.

The "criterion of choice" to isolate a viscosity solution is instead different; for Hamilton-Jacobi equations s.t. $p \mapsto H(x,p)$ is convex (or, at least, satisfying the (H3) in the second part [16]) it is equivalent to say that

- v is viscosity solution, or that
- v is locally semiconcave (as defined in 2.3) and it solves the equation almost everywhere

(This definition of *generalized solution* was proposed in [11]; see also [6], §3.3).

Roughly speaking, this means that, when the solution bends, the kink must always look concave. This said, the viscosity solution is not, in general, unique, and it may differ from the *min* solution: indeed, the functions in frame 3,4 in fig. 1 are both viscosity solutions.

3.4. **Conditions**

We list now a number of conditions we will use in the results we will propose, and discuss the relationship between these.

Conditions 3.5. *We will possibly suppose in the following that*

(CC0): (compatibility condition [6])*:*
$\forall x \in K$, for each connected component A of Z_x, if there is a $p \in A$ such that $p|_{T_xK} = du_0(x)$, then there is a $\exists q \in A$ such that $q|_{T_xK} = du_0(x)$ and $H(x,q) < 0$; that is, if $TK^{\perp u_0}$ intersects A, then it intersects the interior of A, see figure 2;

(OXUp): (properness-coercivity condition)*:*
if C is a compact subset of M, $a \in \mathbb{R}$, then

$$\{(t,y) \in \mathbb{R}^+ \times O \mid \ X(t,y) \in C, \ (U(t,y) + u_0(z)) \le a\} \tag{3.7}$$

is compact;

we will say that (OXUp) *holds in a open region $A \subset M$ if* (3.7) *is compact for any C compact subset of A.*

If the problem is variational, then (OXUp) is usually true; so we are not amazed by finding out that:

Remark 3.6. *If* (OXUp) *holds, then* ($\exists$**u**) *holds, that is, the* min *solution exists.*

The *compatibility condition* (CC0) implies for example, that

Proposition 3.7. *if* (CC0) (H1,H2)*, hold, if $R \ge 1$, if $K, u_0 \in C^{R+1}$, $H \in C^R$, then O is a $(n-1)$-dimensional, embedded submanifold of T^*M, of class C^R.*

Proof. The proof is straightforward: indeed, we can write O as $O = \{H = 0\} \cap TK^{\perp u_0}$; $TK^{\perp u_0}$ is a submanifold; by hypotheses (CC0), $\{H = 0\}$ is a submanifold where it intersects $TK^{\perp u_0}$: indeed, if $DH(x,p) = 0$, then in particular (x,p) would be an isolated point in Z_x but then (CC0) would be unsatisfiable in (x,p).

[6]this is a condition on compatibility of H, K, u_0: see the condition in §5.1 in [13]

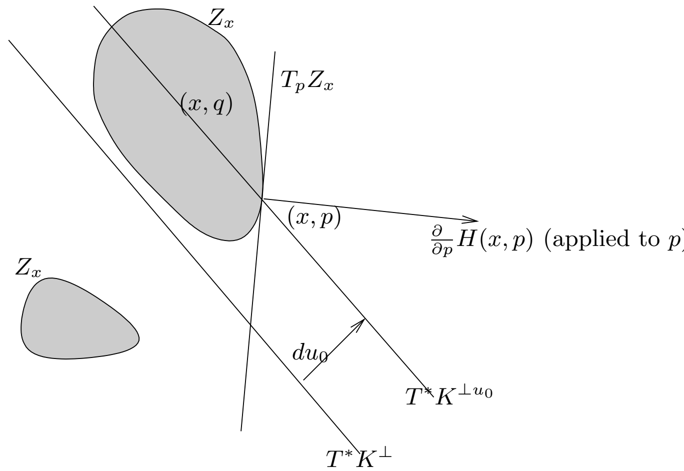

FIGURE 2. condition CC0, and examples of Z_x and $TK^{\perp u_0}$ in T^*_xM (the set Z_x is gray); (the element $\frac{\partial}{\partial p}H(x,p)$ of T_xM is represented by using scalar product duality)

So both parts are submanifolds, of class C^R; since these two submanifolds are transversal, it follows that O is a submanifold. □

Sometimes the reverse of this proposition holds: see 3.10.

On condition (CC0), see also 4.13.

Remark 3.8 (noncharacteristic initial data). *We also obtain that the conditions* (CC0,H2) *imply that* "the initial data K, u_0 is noncharacteristic for the problem (1.1)"*: this is an important condition, necessary for applying Cauchy's method of characteristics* ([7]*); it means that the characteristic curve is transversal to K; indeed, $\frac{d}{dt}X = \frac{\partial H}{\partial p}$, but, by the above reasoning, $\frac{\partial H}{\partial p}$ is not in the annihilator $(T_xK^\perp)^\perp$ of $T_xK^\perp$: since $(T_xK^\perp)^\perp \equiv T_xK$, this is like saying that $\frac{d}{dt}X(0,x,p)$ is not contained in T_xK*

Proposition 3.9. *u is lower semicontinuous where* (OXUp) *holds locally.*

3.5. Example: geodesics and distance

In this section we will show an example, to clarify the role played by the conditions that we have shown (and the ones that we will show in the second part [16]). This example is based on the eikonal equation, that was defined in 1.2.

For the eikonal equation, the set O takes the special form

$$O = \{(x,p) \mid x \in K,\ \forall v \in T_xK,\ p \cdot v = 0,\ |p| = 1\}$$

while the characteristics X are geodesics curves; so, d_K is the *min* solution to (1.3).

Remark 3.10. *In the special case of the eikonal problem, the condition that the set O be a regular submanifold of T^*M implies that u_0 must satisfy the condition $|du_0| < 1$, that coincides with* (CC0)*). So, in the case of the eikonal problem, the inverse of proposition 3.7 is true.*

Example 3.11. *Let $M = \mathbb{R}^2$ and $H(x,p) = |p|^2 - 1$ as above, so that the flow is*

$$X(t,z,p) = z + 2tp \quad , \qquad P(t,z,p) = p \quad , \qquad U(t,z,p) = 2t|p|^2 \quad ,$$

[7] which applies to the case $\dim(K) = \dim(M) - 1$; see eq. 7.16 in [9, § VI]

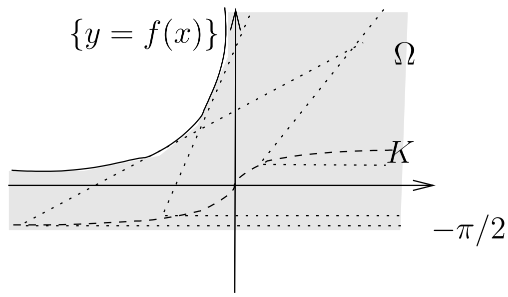

FIGURE 3. Example 3.11; characteristics are dotted, K is dashed, Ω is gray

so that if $|p| = 1$*, then* $U(t,z,p) = d(z,\ X(t,z,p))$*; let*

$$K = \{(x_1, x_2) \mid x_2 = \arctan(x_1)\} \quad ;$$

suppose that

$$u_0(x) = x_1$$

then by direct computation

$$O = \left\{ (x,p) \mid x \in K, p = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \ \text{or } p = \begin{pmatrix} (1+x_1^2)^2 - 1 \\ 2(1+x_1^2) \end{pmatrix} \frac{1}{1+(1+x_1^2)^2} \right\} .$$

We obtain that the min *solution* u *is defined in a set*

$$\Omega = \left\{ x \mid x_1 < 0, -\pi/2 < x_2 \le f(x) \right\} \cup \\ \cup \left\{ x \mid x_1 \ge 0, -\pi/2 < x_2 \right\}$$

where $f : (-\infty, 0) \to (-\pi/2, \infty)$ *is a regular strictly increasing function (and then* Ω *is not open), and that* $u(x) = x_1$ *when* $-\frac{\pi}{2} < x_2 \le \arctan(x_1)$*. (see 4.14).*

Note that the global regular solution of the above problem $v(x) = x_1$, is not equal to the *min* solution. Since $\Omega \neq \mathbb{R}^2$, this equation is clearly not variational.

In this example, the hypotheses (H1,H3), (CC), (CC0), ($\exists$u) are satisfied; and (OXUp) is not satisfied.

($\exists\underline{u}$) (def. in the second part [16]) is satisfied, by choosing

$$\underline{u}(x) = -\sqrt{|x|^2 + 1} \tag{3.8}$$

but (G$\underline{u}$) does not hold.

Example 3.12. *By adding a small circle to* K*, and defining* $u_0 = 0$ *on the circle, we build a second example from the above example; we obtain that the reachable set of this second example is* $\Omega = \mathbb{R}^2$*, and the* min *solution* u *of this second example is lower semicontinuous across the part* $\{x | x_2 = f(x_1)\}$ *of* Γ *(where* (OXUp) *holds locally).*

4. Regularity results

In the following we will assume that $n \doteq \dim(M) \ge 2$.

We start by defining the singularity sets.

We define the set Σ to be the set of $x \in \Omega$ s.t. the minimum (3.5) that defines $u(x)$ is given by at least two points in $\mathbb{R}^+ \times O$, that is

$$\Sigma \stackrel{\text{def}}{=} \left\{ x \in \Omega \mid \quad \begin{array}{c} \exists (s,y),(s',y') \in \mathbb{R}^+ \times O, y=(z,q), y'=(z',q'), \quad \text{s.t.} \\ (s,y) \neq (s',y') \quad x = X(s,y) = X(s',y'), \\ u(x) = U(s,y) + u_0(z) = U(s',y') + u_0(z') \end{array} \right\} \tag{4.1}$$

If the problem (1.1) comes from a variational problem, then Σ is the set of points x such the value function $u(x)$ has (at least) two minima curves.

If $X \in C^1$ and O is a C^1 submanifold, we define the set Γ to be the set of $x \in \Omega$ s.t. the minimum that defines $u(x)$ is given by at least a point in $\mathbb{R}^+ \times O$ that is critical for X, ie

$$\Gamma \stackrel{\text{def}}{=} \left\{ x \in \Omega \mid \quad \begin{array}{c} \exists s \geq 0, \exists y = (z,q) \in O, \quad x = X(s,y), \\ u(x) = U(s,y) + u_0(z), \frac{\partial X}{\partial (s,y)}(s,y) \text{ has not rank } n \end{array} \right\} \tag{4.2}$$

(where, in the definition, X is viewed as a map from $\mathbb{R} \times O$ to M).

This latter set Γ is called *set of conjugate points* in Calculus of Variation; it can also be called *set of optimal focal points*, adapting a terminology from Riemannian geometry.

The former set Σ will play the role that the set $\Sigma_u \stackrel{\text{def}}{=} \{x \mid \nexists du(x)\}$ was playing in the introduction; and indeed, in mild hypothesis, $\Sigma_u = \Sigma$, as will be discussed in a forthcoming paper. We will call Σ *set of singular points*.

We will cover the set Γ with

$$X\left(\bigcup_{i \geq 1}^{n} G^{(i)}\right) = \{X(x) \mid x \in \cup_i G^i\}$$

that is the image of the sets

$$G^{(i)} \stackrel{\text{def}}{=} \left\{ (s,y) \in \mathbb{R} \times O \;\middle|\; \frac{\partial X}{\partial (s,y)}(s,y) \text{ has rank } n-i \right\} \tag{4.3}$$

under the map X. We will call the image points $\{X(x) \mid x \in \cup_i G^i\}$ *focal points*.

Note that, if (H2) holds and Z_x has no isolated points, then $G^{(n)} = \emptyset$, since $\frac{\partial}{\partial t} X = \frac{\partial}{\partial p} H(X,P) \neq 0$.

Let $s \geq 0$. We will use

- the Hausdorff measure $\mathcal{H}^s$ and Hausdorff dimension $\dim^H$
- the set function $\mathcal{M}^s(A)$
- the "entropy dimension", or "box dimension" $\dim^M(A)$, that is the least s such that $\mathcal{M}^s(A) < \infty$.

These concept can be defined for any A subset of a generic manifold M, as argued, and to the limits specified, in appendix §A; we will use them through the Sard-type theorem A.4.

We will always assume (∃u) and (CC) (for prop. 3.3).

4.1. Regularity of conjugate points

We will prove in this section results regarding the set of *focal points*; each following result extends to the set Γ of *conjugate points* that is a subset of the focal points. In the next section we will instead prove in 4.9 that, under certain hypotheses, the set $\Gamma \setminus \Sigma$ has dimensionality strictly less than $n-1$.

Theorem 4.1. *Assume* (CC0,H1,H2). *If u_0, K, H are regular enough , then, by lemma 4.4, there is a (at most) countable number of $n-1$ dimensional submanifolds of $\mathbb{R} \times O$ that cover all the sets G^i; these submanifolds are graphs of functions $\lambda_{i,h} : A_{i,h} \to \mathbb{R}$ (for $h = 1 \ldots$) where $A_{i,h} \subset O$ are open sets. The least regular case is*

$i = n-1$, and the regularity of the λ functions is related to the regularity of u_0, K, H, and to the dimension $\dim(M) = n$ *as in the following table:*

$\dim(M)$	u_0, K	H	λ
$n=2$	$C^{(R+2,\theta)}$	$C^{(R+2,\theta)}$	$C^{(R,\theta)}$
$n\geq 3$	$C^{(R+2,\theta)}$	$C^{(R+n-1,\theta)} \cap C^n$	$C^{(R,\theta)}$

(4.4)

where $R \in \mathbb{N}, \theta \in [0,1]$.

We now infer some explanatory results on the regularity of the *focal points* $X(\cup_i G^i)$ from the above theorem.

At the lowest regularity, when $u_0, K \in C^2$, $H \in C^n$, we know that $X \in C^1$ and that the sets G^i are graphs; we conclude that the set of focal points has measure zero. When $u_0, K \in C^{(2,\theta)}$, $H \in C^n \cap C^{(2,\theta)}$, we know that the dimension of the sets G^i does not exceed $n-\theta$; so again we conclude that the set of focal points has dimension at most $n-\theta$. In the case $\theta = 1$, we can obtain the set of all focal points is *rectifiable*; that is, if $u_0, K \in C^{(2,1)}$, $H \in C^n \cap C^{(2,1)}$, then the sets G^i are covered by Lipshitz graphs, so (by known results in [7]) the set of focal points may be covered by $(n-1)$-dimensional C^1 regular submanifolds of M, but for a set of Hausdorff $\mathcal{H}^{n-1}$ measure zero.

When we further raise the regularity, we may suppose that $u_0, K \in C^{s+3}$, $H \in C^{s+n}$ (with $s \in \mathbb{N}$) [8]; then the sets G^i are covered by graphs $(\lambda(y), y)$ inside $\mathbb{R} \times O$ of regularity C^{1+s} ; while $X \in C^{2+s}$ (at least), and we restrict it to those graphs; we can then apply Thm. A.4 to state that the focal points are covered by C^{1+s} regular submanifolds of M but for a set of $\mathcal{H}^\alpha$ measure zero, where $\alpha \doteq n - 2 + 1/(1+s)$.

In some special cases we can prove that the set of focal points is actually C^{R+1}-M^{n-1}-rectifiable: see in §5.2.1.

In the example 3.11 we have that $n = 2$ and $u_0, K, H \in C^\infty$, and the curve $\{x | x_2 = f(x_1)\}$ is contained in Γ, so that 4.1 is fairly optimal.

We will also prove that

Theorem 4.2. *Suppose $R \geq 2$, and $u_0, K, H \in C^{(R+1,\theta)}$, and* (CC0,H1,H2) *holds; then it is possible to cover Γ by a countable family of C^1 hypersurfaces of dimension $(n-1)$, with the exception of a set that has $\mathcal{H}^{n-1}$ measure equal to zero.*

When $R = 1$, that is, when $H \in C^{(2,\theta)}$, $u_0, K \in C^{2,\theta}$, then the set of focal points has dimension at most $n-\theta$.

(In this theorem's hypotheses we have written $R \geq 2$; but, if $R \geq n-1$, the previous theorem is stronger.)

This section is devoted to the proof of these results.

We will use this version of Dini's theorem

Lemma 4.3. *Let $r \in \mathbb{N}$, $\theta \in [0,1]$. Consider a function $F : \mathbb{R}^n \to \mathbb{R}$; we write $F = F(t,x)$, with $t \in \mathbb{R}$, $x \in \mathbb{R}^{n-1}$; suppose $F \in C^{r,\theta}$ and $\frac{\partial F}{\partial t}$ exists and is continuous; if $\frac{\partial F}{\partial t}(t,x) \neq 0$, then in a neighbourhood V of (t,x) the zero set $\{F = 0\}$ coincides with the graph $\{(\phi(x), x)\}$ of a Hölder function $\phi \in C^{(r,\theta)}$. If $r + \theta > 0$, we will then say that $V \cap \{F = 0\}$* can be covered *by a manifold of class $C^{(r,\theta)}$; if $r = 0$ and $\theta > 0$, $V \cap \{F = 0\}$ has Hausdorff dimension at most $(n-\theta)$, and, if $\overline{V}$ is compact, then $\mathcal{M}^{n-\theta}(V \cap \{F = 0\}) < \infty$; if $r = \theta = 0$, we can say that the Lebesgue measure of the zero set $\{F = 0\}$ is zero.*

The main tool is this lemma; the complete proof of the lemma is in section 6.

Lemma 4.4. *We assume that the hypotheses* (CC0,H1,H2) *hold.*

We set the regularity of the data u_0, K, H by defining parameters $R, R' \in \mathbb{N}$, $\theta, \theta' \in [0,1]$, and assuming that

$$u_0 \in C^{(R'+2,\theta')}, \quad K \in C^{(R'+2,\theta')}, \quad H \in C^{(R+2,\theta)} \quad ;$$

[8] a similar result may be obtained when $u_0, K \in C^{(s+3,\theta)}$, $H \in C^{(s+n,\theta)}$

by 3.7, the flow $\Phi = (X, P)$ is $C^{(R+1,\theta)}$ regular; and O is a $C^{(R'+1,\theta')} \cup C^{(R+2,\theta)}$ manifold (that is, the least regular of the two).

Lets fix $i \geq 1$, $i \leq n-1$, and fix a point $(s', y') \in \mathbb{R} \times O$, such that $(s', y') \in G^{(i)}$.

Let $\mathcal{U}$ be a neighbourhood of 0 in $\mathbb{R}^{n-1}$ and let $\phi : \mathcal{U} \to O$ be a local chart to the neighbourhood $\phi(\mathcal{U})$ of $y' = \phi(0)$. The map ϕ has regularity $C^{(R'+1,\theta')} \cup C^{(R+2,\theta)}$. In the following, y will be a point in $\phi(\mathcal{U})$.

To study $G^{(i)}$, we should study the rank of the Jacobian of the map $(t, x) \mapsto X(t, \phi(x))$; since the regularity of X is related only to the regularity of H, it will be useful to decouple this Jacobian in two parts. To this end, we define a n-form α on $\mathbb{R} \times O$, with requirement that $\alpha(t, y) = \alpha(y)$ (that is, α does not depend on t).

Writing $X^{(t,y)}$ for $X(t, y)$, let

$$X^{(t,y)^*}\alpha$$

be the push–forward of α along X ; $X^{(t,y)^*}\alpha$ is then a tangent form defined on $T_{X(t,y)}M$; it will be precisely defined in eq. (6.2). We remark that $X^{(t,y)^*}\alpha = 0$ iff $(t, y) \in \bigcup_j G^j$. Note that the pushforward $X^{(t,y)^*}$ is $C^{(R,\theta)}$ regular, while the form α is as regular as TO, that is, α is $C^{(R',\theta')} \cup C^{(R+1,\theta)}$.

Note that, since X solves an O.D.E., then X and $\frac{\partial}{\partial t}X$ have the same regularity; note moreover that

$$\frac{\partial^j}{\partial t^j}\left(X^{(s',y')^*}\alpha\right) = \left(\frac{\partial^j}{\partial t^j}X\right)^{(s',y')^*}\alpha$$

since α does not depend on t. So, by hypotheses and by the definition (6.2) of $X^{(t,y)^*}\alpha$, the forms $X^{(t,y)^*}\alpha$ and $\frac{\partial}{\partial t}(X^{(t,y)^*}\alpha)$ have regularity $C^{(R,\theta)} \cap C^{(R',\theta')}$ (see also eq. (6.3)); the derivates $\frac{\partial^j}{\partial t^j}X^{(s',y')^*}\alpha$ with $j \geq 1$ have regularity $C^{(R-j+1,\theta)} \cup C^{(R',\theta')}$.

Then, when $R + 1 \geq i$, we prove (in section 6) that

$$X^{(s',y')^*}\alpha = 0, \quad \frac{\partial}{\partial t}X^{(s',y')^*}\alpha = 0, \quad \ldots \frac{\partial^{i-1}}{\partial t^{i-1}}X^{(s',y')^*}\alpha = 0$$

whereas

$$\frac{\partial^i}{\partial t^i}X^{(s',y')^*}\alpha \neq 0$$

We define eventually the map $F : \mathbb{R} \times \mathbb{R}^{n-1} \to \mathbb{R}$ given by

$$F(t, x) = \frac{\partial^{i-1}}{\partial t^{i-1}}X^{(t,\phi(x))^*}\alpha \quad ;$$

since

$$\frac{\partial}{\partial t}F(t, x) \stackrel{def}{=} \frac{\partial^i}{\partial t^i}X^{(t,\phi(x))^*}\alpha \neq 0$$

the above Dini lemma implies that the set $G^{(i)}$ is locally covered by the graph of a function λ_i defined on a open subset of O; λ has the same regularity of F, so, if $i = 1$ then λ is in $C^{R,\theta} \cup C^{(R',\theta')}$ while for $i \geq 2$ it is $C^{(R-i+2,\theta)} \cup C^{(R',\theta')}$.

The above directly implies 4.1.

If H is not enough regular, we may balance this lack imposing more regularity on O (that is, on u_0, K). We use the Sard–type theorem A.4, due to Yomdin:

Lemma 4.5. *let $i \geq 1, i \leq n-1, R + \theta > 1$, assume $O \in C^{R,\theta}$ and $H \in C^{R+1,\theta}$, and consider the map X as a $C^{R,\theta}$ map from $\mathbb{R} \times O \to M$: then*

- *the Hausdorff dimension of $X(G^{(i)})$ is at most $(n - i + i/(R + \theta))$ (where $n = \dim(M)$);*
- *moreover, in case $\theta = 0$, the $\mathcal{H}^{n-i+i/R}$ measure of $X(G^{(i)})$ is zero*

This result does not provide much information for $X(G^{(1)})$: indeed, it just states that the Hausdorff dimension of $X(G^{(1)})$ does not exceed $n-1+1/(R+\theta)$; whereas, by the above lemma 4.4, if $R+\theta \geq 2$ then $X(G^{(1)})$ is rectifiable; otherwise, if $R=1$, $\theta \in (0,1)$, then $X(G^{(1)})$ has dimension at most $n-\theta$, which is always less than $n-1+1/(1+\theta)$.

When we consider $i \geq 2$, though, this lemma does provide new information for $X(G^{(i)})$: indeed,

- if $R \geq i$ then, by the lemma 4.4, the set $X(G^{(i)})$ is rectifiable, so it is at most $n-1$ dimensional; but, by lemma 4.5, we also obtain that its dimension actually does not exceed $n-i+i/(R+\theta)$, that is less than $n-1$ when $R+\theta$ is large
- whereas, if $R < i$, then we can only use lemma 4.5; in particular, if $R \geq 2$, then the $\mathcal{H}^{n-1}$–measure of $X(G^{(i)})$ is always zero, and the dimension is $n-i(1-1/(R+\theta))$, which decreases when i increases.
  When $R=1$, that is, when $H \in C^{(2,\theta)}$, if $O \in C^{(1,\theta)}$, then $X(G^{(i)})$ has dimension at most $n-i\frac{\theta}{1+\theta}$ (note that the set $X(G^{(1)})$ has dimension at most $n-\theta$, which is higher)

So, in the hypotheses of theorem 4.2, we can apply the lemma 4.5 to

$$X(G^{(R+1)}) \dots X(G^{(n-1)})$$

and apply the lemma 4.4 to

$$G^{(1)} \dots G^{(R)}$$

(and then apply lemma 4.5 to their image under X). This proves Thm. 4.2.

Note that the above is actually a study of the "regularity" of the lagrangian submanifold Λ (defined in 3.4), where we call "regular" all points of Λ where its tangent is not vertical (i.e. the regular points for the canonical projection π_M^* when restricted to Λ). See also 3.32 in [15].

Remark 4.6. *in the definition of $x \in \Gamma$, it is said that there must exist a pair (s,y) s.t. $X(s,y)=x$, satisfying these two conditions:*

- *"$u(x) = U(s,y) + u_0(z)$", that is, (s,y) is the minimum for $u(x)$*
- *"the map X is critical in (s,y), that is, $\frac{\partial X}{\partial(s,y)}(s,y)$ is not invertible".*

In all the above discussions we have only used the second condition, so we have actually proved results regarding the set of focal points; in what follows, instead, we will use both conditions, and prove a result that is specific to conjugate points.

4.2. **Stricter estimates on conjugate points**

Let now $\mathcal{A}$ be the set of points x where u is continuous and that are in the internal part $\mathring{\Omega}$ [9]. We want now to prove that, if we only look inside $\mathcal{A}$, then the set $\Gamma \cap \mathcal{A} \setminus \Sigma$ of points that are conjugate (but not singular), is actually of dimension strictly less than $n-1$.

Let $\overline{x} = (t,y)$ in what follows; we will write DX for $\frac{\partial}{\partial \overline{x}} X$.

As a consequence of lemma 4.4, the sets $G^{(i)}$ are locally covered by regular graphs.

Fix $(t',y') \in G^{(1)}$; then we may choose a neighbourhood V of y' and $e>0$ s.t. $([t'-e,t'+e] \times V) \cap G^{(j)} = \emptyset$ for $j>1$ (using the lower semicontinuity of the rank), and we may choose a ϕ s.t. for $y \in V$, for t in $[t'-e,t'+e]$,

$$\det DX(t,y) = 0 \iff t = \phi(y) \quad .$$

[9] we will show in next section that $\mathcal{A} \supset \Omega \setminus (\Gamma \cup K)$; the remark 3.12 shows that u may be discontinuous across Γ

Then, for $y \in V$, (possibly choosing a smaller V) there is a function $\nu(y)$, $\nu : V \to T(\mathbb{R} \times O)$ with $\nu(y) \in T_{(0,y)}(\mathbb{R} \times O)$, [10] such that $DX(\phi(y), y))\nu(y) = 0$, that is, $\nu(y) \in \mathrm{Ker}DX(\phi(y), y))$, and ν is as regular as DX. [11]

First, we prove a lemma

Lemma 4.7. *Suppose $H, X \in C^2, \phi \in C^1$. Let $\overline{x} = (t, y)$; let ϕ, ν, V be as above; for sake of simplicity, assume that y has local coordinates $y_2 ... y_n$ in V, so that $\overline{x}_1 = t, \overline{x}_2 = y_2, ... \overline{x}_n = y_n$; then*

$$\frac{\partial^2 H}{\partial q^2}\left(\nu_1 - \sum_{i=2}^{n} \frac{d\phi}{dy_i}\nu_i\right) = \sum_{h=1}^{n} \frac{\partial^2 X_h}{\partial \nu^2} q_h \stackrel{def}{=} q \cdot \frac{\partial^2 X}{\partial \nu^2} \tag{4.5}$$

where

$$q_h(y) = \sum_j \frac{\partial P_h}{\partial \overline{x}_j}(\phi(y), y)\nu_j(y) \tag{4.6}$$

and

$$\frac{\partial^2 H}{\partial q^2} = \sum_{r,s=1}^{n} q_r(y) q_s(y) \frac{\partial^2 H}{\partial p_r \partial p_s}(X(\phi(y), y), P(\phi(y), y))$$

and similarly for $\frac{\partial^2 X_h}{\partial \nu^2}$; whereas (by 6.2)

$$q \cdot \frac{\partial X}{\partial \nu} \stackrel{def}{=} \sum_h \frac{\partial X_h}{\partial \overline{x}_j} q_h = \sum_{h,r} \frac{\partial X_h}{\partial \overline{x}_j}\frac{\partial P_h}{\partial \overline{x}_r}\nu_r = \sum_{h,r} \frac{\partial X_h}{\partial \overline{x}_r}\frac{\partial P_h}{\partial \overline{x}_j}\nu_r = 0 \tag{4.7}$$

Proof. The proof goes exactly as in lemma 4.18 in [4]. □

Then we prove that

Lemma 4.8. *Suppose that* (OXUp) *holds; choose a point $x' \in \Gamma$, $x' \notin \Sigma$, and $x' \in \mathring{\Omega}$, such that u is continuous in x'; let $x' = X(t', y')$ where t', y' minimizes $u(x')$, and suppose that $(t', y') \in G^{(1)}$; let V, ϕ, ν be defined as above, and $q(y)$ be defined as in* (4.6)*; then*

$$q \cdot \frac{\partial^2 X}{\partial \nu^2} \stackrel{def}{=} \sum_i \frac{\partial^2 X_i}{\partial \nu^2}(\phi(y'), y') q_i(y') = 0$$

Proof. Suppose that $\left(q \cdot \frac{d^2 X}{d\nu^2}\right) \neq 0$, whereas, we know that for any ν', $q \cdot \frac{dX}{d\overline{x}}\nu' = 0$, (see (4.7) and (6.5)): so, in a sense, the map X folds along the hyperplane orthogonal to q.

Suppose that $q \cdot \frac{d^2 X}{d\nu^2} > 0$, for simplicity; let $e > 0$; let γ be a curve in M with $\gamma(0) = x'$, $\dot\gamma(0) = \frac{d^2 X}{d\nu^2}$; then, possibly restricting V, by lemma 2.5 in [2], (used as 4.19 in [4]) the equation $X(t, y) = \gamma(s)$ has two solutions (t, y) in $(t' - e, t' + e) \times V$ for small positive s, none for small negative s.

So, let $x_k = \gamma(-1/k)$ and let (t_k, y_k) be the minimizer of $u(x_k)$; by the lemma, $(t_k, y_k) \notin (t' - e, t' + e) \times V$.

On the other hand, by OXUP, we may assume that (t_k, y_k) converges to some point (t, y), so that (t, y) minimizes $u(x')$ (since u is continuous in x'): but then, $(t, y) \neq (t', y')$, so that $x' \in \Sigma$. □

We obtain from the above

---

[10] ν is a vector field on the graph of ϕ

[11] if $i = 1$, if we choose local coordinates around x, we may build the n-covector-form

$$\frac{\partial X_1}{\partial \overline{x}} \wedge \cdots \wedge \frac{\partial X_n}{\partial \overline{x}}$$

on $T(\mathbb{R} \times O)$; this is, actually, a $n - 1$ form, when evaluated for $(t, y) = (\phi(y), y)$ so it admits a dual 1-tangent-form $\nu(y)$

Theorem 4.9. *Let $\mathcal{A}$ be the set of points x where u is continuous and that are in the internal part $\mathring{\Omega}$; suppose* (OXUp, CC0, H1,H2) *hold, $K, u_0, H \in C^{(R+1,\theta)}$, $R \ge 2$; let $\alpha = n - 2 + 2/(R+\theta)$; then the set $(\Gamma \cap \mathcal{A} \setminus \Sigma)$ has Hausdorff dimension at most α, and if $\theta = 0$ moreover $\mathcal{H}^\alpha(\Gamma \cap \mathcal{A} \setminus \Sigma) = 0$.*

Note that if $R + \theta > 2$, then $\alpha < n - 1$; so this theorem is not implied by 4.2.

Proof. Let $x \in (\Gamma \cap \mathcal{A}) \setminus \Sigma$, and let (t', y') be the optimal point for $u(x)$; if $(t', y') \in G^{(2)} \cup \dots G^{(n)}$, then we directly apply Sard–type Lemma A.4, so that the Hausdorff dimension is at most $n - 2 + 2/(R+\theta)$; if (t', y') is in $G^{(1)}$, then, by the lemma above,

$$\frac{\partial^2 X}{\partial \nu^2}(\nu(y), y) \cdot q = 0$$

so that, by the lemma 4.7,

$$\sum_{i=2} \frac{\partial \phi}{\partial y_i}(y)\nu_i(y) = \nu_1 \tag{4.8}$$

since $\frac{\partial}{\partial q^2} H > 0$, by (H2).

Let $\tilde{X}(y) = X(\phi(y), y)$; then, by choosing $\nu = \nu(y)$

$$\sum_{i=2}^{n} \frac{\partial \tilde{X}}{\partial y_i}\nu_i = \frac{\partial X}{\partial t}(\phi(y), y) \sum_{i=2}^{n} \frac{\partial \phi}{\partial y_i}(y)\nu_i + \sum_{i=2}^{n} \frac{\partial X}{\partial y_i}(\phi(y), y)\nu_i$$

but this, by (4.8) above, is equal to

$$\sum_{i=2}^{n} \frac{\partial \tilde{X}}{\partial y_i}\nu_i = \frac{\partial}{\partial t} X(\phi(y), y)\nu_1 + \sum_{i=2}^{n} \frac{\partial X}{\partial y_i}(\phi(y), y)\nu_i = DX\nu = 0$$

which brings to

$$\text{rank}\frac{d}{dy}\tilde{X} \le n - 2$$

[12] Suppose now that $K, u_0 \in C^{(R'+1,\theta)}$, $H \in C^{(R+1,\theta)}$, with $R, R' \ge 2$. The manifold O has regularity $C^{R+1,\theta} \cup C^{R',\theta}$; (by lemma 4.4 in case $i = 1$) ϕ has regularity $C^{R-1,\theta} \cup C^{R'-1,\theta}$; so the map $\tilde{X} : O \to M$ has class $C^{R-1,\theta} \cup C^{R'-1,\theta}$. So we conclude by setting $R = R'$ and applying the lemma A.4 to $\tilde{X}$, that entails

$$\mathcal{H}^{n-2+1/(R+\theta-1)}(\tilde{X}(V)) < \infty$$

so that the dimensionality of $(\overline{X}(G') \setminus \Sigma)$ is at most $n - 2 + 1/(R+\theta-1)$; then $\alpha = \max\{n - 2 + 2/(R+\theta),\ n - 2 + 1/(R+\theta-1)\}$ and the result follows. ☐

Note that, if the problem (1.1) is variational, then the set $\mathcal{A} = \Omega = M$, and u is continuous (as is well known, and is shown also by the results in the second part [16] of this paper); if the problem is not variational, then the theorem is false outside of $\mathcal{A}$, as shown by remark 3.12.

Remark 4.10. *The above three theorems 4.1, 4.2 and 4.9, need the hypothesis* (H2)*; but, it is clear from the proofs that this may be weakened to these hypotheses:* "O is a manifold" *and* "$\forall \mu \in \ker DX$, $\mu \ne 0$, if $\nu = DP\mu$, then $\frac{d^2 H}{d\nu^2} > 0$"

[12] there is an errata in the proof of thm 4.17 in [4]: where it states "$\mathcal{H}^{n-1+2/R}(\overline{X}(G') \setminus \Sigma) = 0$" it should read "$\mathcal{H}^{n-1+1/(R-1)}(\overline{X}(G') \setminus \Sigma) = 0$"; the thesis holds nonetheless, since $R \ge 2$. The correct proof follows exactly the last steps of this proof. This error has been corrected also in [17].

4.3. **Structured regularity of u**

The above discussion can be combined with a simple regularity result for the *min* solutions to obtain a good understanding of the regularity (and lack of it) of u:

Theorem 4.11. *assume* (CC0,OXUp, H1,H2)*; assume also that $H \in C^{R+1}$, $u_0, K \in C^{R+1}$, with $R \geq 1$; suppose that u is locally bounded from below; then*

1. *$\Gamma \cup K$ is closed in Ω, and $\Omega \setminus (\Gamma \cup K)$ is open;*
2. *the closure of Σ in Ω is contained in $\Sigma \cup \Gamma \cup K$, which is closed in Ω* [13]

   So we split the open set $\Omega \setminus (\Gamma \cup K)$ in parts, where we can state that
3. $$u \in C^{R+1}(\Omega \setminus (\Gamma \cup \Sigma \cup K))$$

   and u is a regular solution of (1.1) *in the open set $\Omega \setminus (\Gamma \cup \Sigma \cup K)$;*
4. *if $x \in \Sigma \setminus (\Gamma \cup K)$, then there exists a neighborhood B of x (containing no points of $\Gamma \cup K$) and a finite number of C^{R+1} functions $u_1, \ldots, u_k : B \to \mathbb{R}$ such that*
   - *the graphs $\{(x', u_i(x')) \mid x' \in B\}$ are transversal*
   - *$u(x') = \min_i u_i(x')$ for $x' \in B$*
   - *the u_i solve $H(x, du_i(x)) = 0$ in B*

   We obtain, in particular, that $\Sigma \setminus (\Gamma \cup K) \subset \{x \mid \nexists \nabla u(x)\}$

These results are localizable to any open region where (OXUp) *holds.*

Corollary 4.12. *The* min *solution u is locally semiconcave in $\Omega \setminus (\Gamma \cup K)$ (as defined in 2.3); then, u is a viscosity solution of* (1.1) *in $\Omega \setminus \Gamma$.*

Remark 4.13. *It is somewhat unfortunate that, if u is a $C^1(M)$ function that solves* (1.1)*, and if K is $(n-1)$-dimensional and orientable, then* (CC0) *must be false: regular solutions are not covered by the above theorem.*

Remark 4.14 (On $t \geq 0$)**.** *Suppose that* $\dim K = \dim M - 1$*, and that K has an orientation in M, and* (CC0) *holds: then, the manifold O is composed by two connected components; we pick one, that we call O^+.*

We may build the min *solution without the requirement that $t \geq 0$ in* (3.5)*, and using O^+ instead of O: this alternative definition u^+ of* min *solution would be then*

$$u^+(x) \stackrel{def}{=} \min_{\begin{cases} t \in \mathbb{R}, \ (z,q) \in O^+ \\ s.t. \ X(t,z,q) = x \end{cases}} U(t,z,q) + u_0(z)$$

The above theorem, when applied to u^+, would change slightly: the third and fourth statement would extend to the sets $\Omega \setminus (\Gamma \cup \Sigma)$ and $\Sigma \setminus \Gamma$, respectively. [14] *In particular, in the example 3.11, if we choose $O^+ = \{(z,q) \mid z \in K, q = (1,0)\}$, we obtain $u^+(x) = x_1$ in the region $-\frac{\pi}{2} < x_2 < \frac{\pi}{2}$.*

5. Applications

5.1. **the Cauchy problem**

We show now how the above theorems may be used for the Cauchy problem (1.2)

$$\begin{cases} \frac{\partial}{\partial t} w(t,x') + H'(t,x',\frac{\partial}{\partial x'} w(t,x')) = 0 & \text{for } t > 0,\ x' \in M' \\ w(0,x') = w_0(x') & \forall x' \in M'. \end{cases} \tag{1.2}$$

---

[13] Note that the example 3.11 shows that we may have $\overline{\Sigma} \neq \Sigma \cup \Gamma$.

[14] even if $\dim K < \dim M - 1$, we could anyway build a *min* solution using all $t \in \mathbb{R}$, that is, without the requirement that $t \geq 0$ in (3.5); but we have would have anyway that $\Gamma \supset K$, so that in general u would not be regular on K

As customary, we define $M = \mathbb{R} \times M'$, $K = \{0\} \times M'$, $u_0(0, x') = w_0(x')$, and then we split variables and differentials (accordingly to the product structure $M = \mathbb{R} \times M'$), as $x = (t, x')$, $dw = (\frac{\partial}{\partial t} w, \frac{\partial}{\partial x'} w)$, $p = (\widetilde{p}, p')$; we define H by

$$H(x, p) = \widetilde{p} + H'(t, x', p')$$

so that

$$H(x, dw) = \frac{\partial}{\partial t} w(t, x') + H'\Big(t, x', \frac{\partial}{\partial x'} w(t, x')\Big)$$

and then $w(x)$ solves (1.1).

The set O is, in this case,

$$O = \Big\{ (t, x', \widetilde{p}, p') \in T^*M \;\big|\; (t, x') \in K,\ \widetilde{p} = -H'(t, x', p'),\ p' = \frac{\partial}{\partial x'} w(t, x') \Big\} =$$

$$= \Big\{ \Big(0, x', -H'\Big(0, x', \frac{\partial}{\partial x'} w_0(x')\Big), \frac{\partial}{\partial x'} w_0(x') \Big) \in T^*M \;\big|\; x' \in M' \Big\}$$

which is obviously a manifold.

The characteristic strips $(X(t, z, q), P(t, z, q))$ for 1.1 (defined in (3.1)) are easily related to the characteristic strips $(X'(t, z', q'), P'(t, z', q'))$ for 1.2, as follows. If we write the equation (3.1) by dividing the first component (in $\mathbb{R}$) from the second component (in M'), and similarly for P, we obtain $X = (T, X')$, $P = (\widetilde{P}, P')$. This generates four O.D.E. whose initial conditions are coded by the set O:

$$T(0) = 0, \quad X'(0) = x', \quad \widetilde{P}(0) = \widetilde{p}, \quad P'(0) = p' = \frac{\partial}{\partial x'} w_0(x')$$

So two (trivial) O.D.E. drive the characteristic curves T and P'' so that $T(t) = t$ and $\widetilde{P}(t) = -H'(t, X'(t), P'(t))$; and two O.D.E. drive X' and P'.

The function H does not satisfy (H2): indeed, $\frac{\partial^2}{\partial \widetilde{p}^2} H = 0$; we therefore use remark 4.10 to apply the results in this paper to the Cauchy problem above.

For any $\nu \in TT^*M$ we decompose $\nu = (\tilde{\nu}, \nu')$ with $\tilde{\nu} \in \mathbb{R}$, $\nu' \in TT^*M'$, and we obtain $\frac{\partial^2}{\partial \nu^2} H = \frac{\partial^2}{\partial \nu'^2} H'$. Let us suppose that $H' \in C^2(T^*M)$, and

$$\frac{\partial^2}{\partial p_i' p_j'} H'(t, x', p')$$

is positive definite, for any (t, x', p') in $\mathbb{R} \times T^*M'$: then $\frac{\partial^2}{\partial \nu^2} H \neq 0$ iff $\nu' \neq 0$.

Let $\mu \in \mathrm{Ker} DX$ and set $\nu = DP\mu$. We decompose DX, DP as $DX = (DT, DX')$, and $DP = (D\widetilde{P}, DP')$. Since $T = t$ and $\mu \in \mathrm{Ker} DT$, then $DT\mu = \tilde{\mu} = 0$. Since $H(X, P) = \widetilde{P} + H'(t, X', P') = 0$, then

$$\frac{d}{d\mu} H = 0 = D\widetilde{P}\mu + \frac{\partial H'}{\partial t} \tilde{\mu} + \frac{\partial H'}{\partial x'} DX'\mu + \frac{\partial H'}{\partial p'} DP'\mu = \tilde{\nu} + \frac{\partial H'}{\partial p'} \nu' = 0$$

So if $\mu \in \mathrm{Ker} DX, \mu \neq 0$ then (by the rank argument Lem. 6.1) $\nu = DP\mu \neq 0$: then $\nu' \neq 0$, and then $\frac{\partial^2}{\partial \nu^2} H \neq 0$. Summarizing

Proposition 5.1. *To apply the theorems 4.1, 4.2 and 4.9 to the Cauchy problem, we substitute the hypothesis* (H2) *by the hypothesis "H' is C^2 and $\frac{\partial^2}{\partial p_i' p_j'} H'$ is positive definite".*

This improves the results of 4.10 4.12 4.17 in [4] (reported as 1.2 and 1.1 in this paper); to provide for an easy comparison, we summarize these results

- if $n' = \dim(M')$, $n = n' + 1$, if $H' \in C^s$ with $s = n \vee 3$ and $w_0 \in C^2$, then the set Γ has measure zero, so the set $\overline{\Sigma}_u = \Sigma \cup \Gamma$ has measure zero;
- if $H, w_0 \in C^{(2,1)}$, then the set Γ is rectifiable, so the set $\overline{\Sigma}_u = \Sigma \cup \Gamma$ is rectifiable;
- and when $H' \in C^{R+1,\theta}$, $w_0 \in C^{R+1,\theta}$, $R \geq 2$, w is continuous, we prove that the Hausdorff dimension of $\Gamma \setminus \Sigma$ is at most β, and moreover $\mathcal{H}^\beta(\Gamma \setminus \Sigma) = 0$ if $\theta = 0$, where $\beta = n' - 1 + 2/(R + \theta)$. [15]

In the counterexample in §4.4 in [4], w_0 is $C^{1,1}(M')$ and not $C^2(M')$; so our results close the gap between the counterexample, where w_0 is $C^{1,1}(M')$, and the theorem, where w_0 is $C^2(M')$; and actually, studying the counterexample, it is quite clear that, if w_0 is smoothed to become a $C^2(M')$ function, then the counterexample would not work.

5.2. **Eikonal equation and cutlocus**

As in 3.5, consider a smooth Riemannian manifold M, and a closed set $K \subset M$ and let $d_K(x) = d(x, K)$ be the distance to K. We set $u_0 = 0$: then O is the bundle of unit covectors that are normal to TK, and $d_K(x)$ coincides with the *min* solution $u(x)$.

We define

$$\Sigma_{d_K} \stackrel{\text{def}}{=} \{x \mid \nexists \nabla d_K(x)\}$$

If K is C^1, then Σ_{d_K} coincides with Σ as defined in (4.1).

Since d_K is semiconcave in $M \setminus K$, Σ_{d_K} is always rectifiable.

This primal problem is a good test bed to discuss the differences and synergies of the results in this paper and the results in Itoh and Tanaka [10] and Li and Nirenberg [12].

- In the example in §3 in [14] (see 1.5 here), there is a curve $K \subset \mathbb{R}^2$, $K \in C^{1,1}$ such that $\overline{\Sigma}_{d_K}$ has positive Lebesgue measure. Note that in this example $\overline{\Sigma}_{d_K} \neq \mathrm{Cut}(K) = \Sigma_{d_K}$, so the cutlocus $\mathrm{Cut}(K)$ is rectifiable (but not closed).

  We do not know if there is a curve $K \in C^{1,1}$ such that $\mathrm{Cut}(K)$ is not rectifiable. (We recall that, by prop. 14 in [18], $\mathrm{Cut}(K)$ has always measure zero).
- The theorem 4.1 states that if K is C^2, then Γ has measure zero, so by (1.5) and 4.11.4, we obtain that $\overline{\Sigma}_{d_K} = \mathrm{Cut}(K)$ has measure zero; so 4.1 closes the gap between the counterexample §3 [14] and the previous available results.
- In example in remark 1.1 in [12], for all $\theta \in (0, 1)$ there is a compact curve $K \in C^{2,\theta}$ such that the distance to the cut locus is not locally Lipschitz; by thm. 4.1, the cutlocus has dimension at most $n - \theta$.

  We do not know if there exists an example of a compact curve $K \in C^{2,\theta}$ such that $\mathcal{H}^{n-1}(\mathrm{Cut}(K)) = \infty$
- By the results in Itoh and Tanaka [10] and Li and Nirenberg [12], when $K \in C^3$, the distance to the cut locus is locally Lipschitz and the cutlocus is rectifiable, and moreover (by cor 1.1 in [12]), for any B bounded $\mathcal{H}^{n-1}(\mathrm{Cut}(K) \cap B) < \infty$. By Theorem **??**, the set of (non optimal) focal points is rectifiable as well.

5.2.1. *Improvements*

We want to show an improvement, based on theorem 4.1, for the special case of the distance function and the eikonal equation on a 2-dimensional Riemannian manifold.

Consider a 2-dimensional Riemannian manifold M; then there is only a type of focal points: those points $(t, y) \in O$ such that the rank of $DX(t, y)$ is 1; those are the points $(t, y) \in G^1$.

Suppose now that K is compact: then O is compact. Note that, actually, O has a very simple structure: for example, if K is a connected 1-dimensional curve, then O is a fiber bundle on K, with a discrete fiber; if K is a collection of points, then $O \sim K \times S^1$.

We can define the function $c : O \to \mathbb{R}^+ \cup \{\infty\}$ to be the first time $t = c(y)$ such that the rank of $DX(t, y)$ is 1; we define $c = \infty$ if the rank of DX is 2 for all times; we define $B = \{c < \infty\}$, so that the image of $X(c(y), y)$, for $y \in B$, covers all of the optimal focal points Γ.

---

[15] that is: $\beta = \alpha - 1$ if α is defined as in 4.9

Let $s \in \mathbb{N}$. By theorem 4.1, we know that, if $K \in C^{3+s}$, then B is open, and $c \in C^{1+s}$ on B.

Consider an open bounded set $A \subset M$; let $T = \sup_A d_K$; then there is a compact set $C = \{c \leq T\} \subset B$ such that $X(c(y), y)$, for $y \in C$, covers all of the optimal focal points $A \cap \Gamma$. So we can apply Yomdin's Sard–type theorem A.4 to the function c on C, to state that

Corollary 5.2. *Consider a 2-dimensional smooth Riemannian manifold M; suppose that K is a compact C^{3+s} embedded submanifold.*

Then, for any open bounded set $A \subset M$, the set $A \cap \Gamma$ is C^{s+1}-$M^{1/(s+1)}$-rectifiable: that is, it can be covered by at most countably many C^{s+1} curves, but for a set E such that $\mathcal{M}^{1/(s+1)}(E) = 0$.

The above discussion can be extended to the distance function in Finsler manifolds where the Finsler metric F is regular and the Hessian of F is positive definite.

6. Proof of 4.4

We will now prove some results needed for 4.4 and the following. We will silently assume that, everywhere we talk about (X, P) and their derivatives, we are using canonical local coordinates. We assume (CC0); by 3.7, O is a $n-1$–dimensional submanifold of T^*M.

Lemma 6.1. *The derivative*

$$\left(\frac{\partial X}{\partial(t,y)}, \frac{\partial P}{\partial(t,y)} \right)$$

is solution to a linear system, obtained by deriving (3.1)*; by well known properties of linear systems (see e.g. [9, pg. 46]), for any $z, q \in O$*

$$\forall t \geq 0, \qquad rank \begin{pmatrix} \frac{\partial X}{\partial(t,y)}(t,z,q) \\ \frac{\partial P}{\partial(t,y)}(t,z,q) \end{pmatrix} = rank \begin{pmatrix} \frac{\partial X}{\partial(t,y)}(0,z,q) \\ \frac{\partial P}{\partial(t,y)}(0,z,q) \end{pmatrix} \tag{6.1}$$

At time $t = 0$ the flow Φ_0 is the identity; if (CC0) *holds, then the derivative of the flow (X, P) wrt $y \in O$ is the derivative of the injection of O in T^*M; its rank is $n-1$, the dimension of O. The derivative of the flow (X, P) wrt t, at $t = 0$, is transversal to TO, since the curve X is transversal to K (see 3.8). We have proved that the rank in* (6.1) *is n, for any $t \geq 0$.*

Note that $\frac{\partial X}{\partial t,y}(0, y)$ is invertible if and only if $\dim K = n-1$.

Proposition 6.2. *We have that*

$$A_{i,j} \stackrel{def}{=} \sum_j \frac{\partial P_h}{\partial \overline{x}_i} \frac{\partial X_h}{\partial \overline{x}_j}$$

is symmetric

Proof. an easy proof comes from the theory of symplectic geometry: consider the *Lagrangian submanifold* Λ' of T^*M (introduced in 3.4), then

$$\begin{bmatrix} DX \\ DP \end{bmatrix}$$

is a *Lagrangian frame*: that is, its span is the tangent space to Λ'; then the thesis is a known fact, see lemma 2.28 in [15].

$\square$

Now we prove 4.4

We want to define the n form α so that α does not depend on t; and so that $\alpha = e_1 \wedge \cdots \wedge e_n$ where the vectors fields $e_{n-i+1} \dots e_n$ span the kernel of $\frac{\partial}{\partial \overline{x}} X$ at the point (s', y') (kernel that we will call V) while $\frac{\partial}{\partial \overline{x}} X$ is full rank on $e_1 \dots e_{n-i}$ (that generate the space W).

One possible way to to this is to fix the local chart $\phi : \mathcal{U} \subset \mathbb{R}^{n-1} \to O$, define

$$\hat{e}_1 \stackrel{\text{def}}{=} \phi^* \frac{\partial}{\partial x_1} \ , \ \ldots \hat{e}_{n-1} \stackrel{\text{def}}{=} \phi^* \frac{\partial}{\partial x_{n-1}} \ , \ \hat{e}_n \stackrel{\text{def}}{=} \frac{\partial}{\partial t}$$

and then choose a $n \times n$ constant matrix A, so that

$$e_h \stackrel{\text{def}}{=} \sum_k A_{h,k} \hat{e}_k$$

satisfy the requirements.

We write $\overline{x} = (t, y)$, and $DX^{(t,y)}$ for the differential $\frac{\partial X}{\partial \overline{x}}$ computed in t, y.

The push forward $X^{(s,y)^*}\alpha$ has the form

$$X^{(s,y)^*}\alpha \stackrel{\text{def}}{=} DX^{(s,y)} e_1 \wedge \cdots \wedge DX^{(s,y)} e_n \tag{6.2}$$

Now, if we derive the push-forward once, we obtain n terms of the form

$$\begin{aligned} \frac{\partial}{\partial s} X^{(s,y)^*}\alpha = \sum_j DX^{(s,y)} e_1 \wedge \cdots \wedge DX^{(s,y)} e_{j-1} \wedge \frac{\partial}{\partial s}(DX^{(s,y)} e_j) \wedge \\ \wedge DX^{(s,y)} e_{j+1} \wedge \cdots \wedge DX^{(s,y)} e_n = \\ = \sum_j DX^{(s,y)} e_1 \wedge \cdots \wedge DX^{(s,y)} e_{j-1} \wedge \left(\frac{\partial^2 H}{\partial p x} \frac{\partial X}{\partial \overline{x}} + \frac{\partial^2 H}{\partial p p} \frac{\partial P}{\partial \overline{x}} \right) e_j \wedge \\ \wedge DX^{(s,y)} e_{j+1} \wedge \cdots \wedge DX^{(s,y)} e_n \end{aligned} \tag{6.3}$$

If $i > 1$, it is quite clear that all the terms in the sum are zero in (s', y'): indeed, if $j < n$ in the sum, the term $DX^{(s',y')} e_n = 0$ in the wedge product, otherwise if $j = n$, $DX^{(s',y')} e_{n-1} = 0$

Similarly, this happens for any derivation $\frac{\partial^h}{\partial s^h} X_s^* \alpha$, as long as $h < i$: if we derive formally, we will notice, that in any n-form of the sum we always find a term like $DX^{(s',y')} e_k$ for $k \geq n - i + 1$.

This drastically changes if we consider the ith derivative

$$\frac{\partial^i}{\partial s^i} X^{(s,y)^*}\alpha$$

if we derive formally, we may isolate an unique term (which is repeated $i!$ times), namely

$$\begin{aligned} i!\ DX^{(s,y)} e_1 \wedge \cdots \wedge DX^{(s,y)} e_{n-i} \wedge \frac{\partial}{\partial s}(DX^{(s,y)} e_{n-i+1}) \wedge \ldots \\ \cdots \wedge \frac{\partial}{\partial s}(DX^{(s,y)} e_n) \end{aligned} \tag{6.4}$$

which may be nonzero at (s', y'). We now study this term at point (s', y') (and we omit the superscripts such as $^{(s',y')}$), for sake of simplicity).

We may substitute $\frac{\partial}{\partial s}(DX) e_j = \frac{\partial^2 H}{\partial pp} DP e_j$, being $\frac{\partial^2 H}{\partial px} \frac{\partial X}{\partial \overline{x}} e_h = 0$ when $h \geq n - i + 1$; so that

$$\begin{aligned} i!\ DX e_1 \wedge \cdots \wedge DX e_{n-i} \wedge \left(\frac{\partial^2 H}{\partial pp} DP e_{n-i+1}\right) \wedge \ldots \\ \cdots \wedge \left(\frac{\partial^2 H}{\partial pp} DP e_n\right) \end{aligned}$$

The rank of $DXe_1 \wedge \cdots \wedge DXe_{n-i}$ is $n-i$, by our choice of $e_1 \dots e_n$; and similarly, the rank of

$$(\frac{\partial^2 H}{\partial pp} DPe_{n-i+1}) \wedge \cdots \wedge (\frac{\partial^2 H}{\partial pp} DPe_n)$$

is i: otherwise, there would be a non–null vector v in the kernel V of DX such that $DPv = 0$: but this would contradict the lemma 6.1, since then

$$\begin{pmatrix} \frac{\partial X}{\partial(t,y)}(t,z,q) \\ \frac{\partial P}{\partial(t,y)}(t,z,q) \end{pmatrix} v = 0$$

The span of $\frac{\partial^2 H}{\partial pp} DP$ on V is transversal to the span of DX on W: otherwise if there would be vectors $v \in V, w \in W$ such that $\frac{\partial^2 H}{\partial pp} DPv = DXw$, then,

$$(DPv) \cdot (\frac{\partial^2 H}{\partial pp} DPv) = (DPv) \cdot (DXw)$$

that is an indexless version of the formula

$$\sum_{i,j,h,k} \frac{\partial P_h}{\partial \overline{x}_i}(\overline{x}) v_i \frac{\partial^2 H}{\partial p_h p_j} \frac{\partial P_j}{\partial \overline{x}_k}(\overline{x}) v_k = \sum_{i,j,k} \frac{\partial X_h}{\partial \overline{x}_i}(\overline{x}) w_i \frac{\partial P_h}{\partial \overline{x}_k}(\overline{x}) v_k$$

in local coordinates; by 6.2, the RHS is

$$\sum_{i,j,k} \frac{\partial X_h}{\partial \overline{x}_i}(\overline{x}) w_i \frac{\partial P_h}{\partial \overline{x}_k}(\overline{x}) v_k = \sum_{i,j,k} \frac{\partial X_h}{\partial \overline{x}_i}(\overline{x}) v_i \frac{\partial P_h}{\partial \overline{x}_k}(\overline{x}) w_k = 0 \tag{6.5}$$

while the LHS is obviously non zero, by (H2).

Then the above form (6.4) is non degenerate at (s', y').

This concludes the proof of 4.4.

Appendix A. Metric entropy and Yomdin's theorem

Let E be an open domain in $\mathbb{R}^m$; and let $f : E \to \mathbb{R}^n$ be a function, of class at least C^1. We will write $f \in C^{k,\theta}$, for $\theta \in [0,1]$ and $k \in \mathbb{N}$: if $\theta = 0$, we identify $C^{k,\theta} = C^k$, whereas if $\theta > 0$, $C^{k,\theta}$ is the usual space of functions $f \in C^k$ such that $D^k f$ is Hölder continuous of exponent θ.

In the following, $B^n_r(y)$ will be the open ball in $\mathbb{R}^n$, of radius $r > 0$, centered at y; we will often omit the superscript and the point y.

We will denote by $\Sigma = \Sigma(f)$ the set of *critical points* of f, that is, the set of all points x where $Df(x)$ has not maximal rank; $\Sigma^l = \Sigma^l(f)$ will be the subset of E where $Df(x)$ has at most rank l, where $l \in \{0, \dots, ((m \wedge n) - 1)\}$.

We call $\Delta = f(\Sigma)$ the *critical values*; if $B \subset E$, $\Delta(f|B) = f(\Sigma(f) \cap B)$ the critical values coming from B and similarly $\Delta^l = f(\Sigma^l)$ and $\Delta^l(f|B) = f(\Sigma^l(f) \cap B)$.

We define the dimensional constant

$$\alpha = \alpha(m,l,s) = l + \frac{m-l}{s} \tag{A.1}$$

A.1. **Metric entropy**

We recall that, given a set $A \in \mathbb{R}^n$ and a real number $\gamma \in [0,n]$, the *γ–dimensional Hausdorff measure* of A is defined by

$$\mathcal{H}^\gamma(A) = \frac{\mathcal{V}_\gamma}{2^\gamma} \sup_{\delta>0} \inf \left\{ \sum_{i=1}^{+\infty} (\text{diameter } A_i)^\gamma \;\middle|\; A \subset \bigcup_{i=1}^{+\infty} A_i, \;\; \text{diameter } A_i < \delta \right\},$$

where $\mathcal{V}_0 = 1$, $\mathcal{V}_\gamma$ is the Lebesgue measure of the unit ball in $\mathbb{R}^\gamma$ if $\gamma \geq 1$ is an integer, and $\mathcal{V}_\gamma$ is a suitable positive constant otherwise. The *Hausdorff dimension* of A is given by

$$\dim^{\mathcal{H}}(A) = \inf\{\gamma > 0 \,:\, \mathcal{H}^\gamma(A) = 0\}.$$

We similarly recall that the *metric entropy* $M(\varepsilon, A)$ is the minimum number of balls of radius ε that cover A. Obviously

$$\mathcal{H}^\gamma(A) \leq \mathcal{V}_\gamma \liminf_{\varepsilon \to 0} \varepsilon^\gamma M(\varepsilon, A) \tag{A.2}$$

where $\mathcal{V}_s$ is the same dimensional constant that is used in defining the Hausdorff measure. We then define

$$\mathcal{M}^\gamma(A) \doteq \mathcal{V}_\gamma \liminf_{\varepsilon \to 0} \varepsilon^\gamma M(\varepsilon, A) \ .$$

We define the *entropy dimension*

$$\dim^M(A) = \inf\{\gamma > 0 \,:\, \mathcal{M}^\gamma(A) = 0\}.$$

Sometimes $\dim^M(A)$ is called *box dimension*, since it can be obtained by dividing $\mathbb{R}^m$ in cubic boxes of size ε, and counting the number of boxes that intersect the set A, and confronting it with $\varepsilon^{-\gamma}$.

If A is bounded, then $\dim^M(A) \leq n$.

We remark that the map $A \mapsto \mathcal{M}^s(A)$ is *not* a measure[16]; but it is more interesting than $\mathcal{H}^s(A)$ when one is concerned with studying the topological structure of A: for example,

- $\mathcal{M}^s(A) = \mathcal{M}^s\left(\overline{A}\right)$, and then $\dim^M(A) = \dim^M\left(\overline{A}\right)$
- let $t > 0$, and let $B \subset \mathbb{R}$ be defined by $B \doteq \{0\} \cup \{k^{-t} \;:\; k \in \mathbb{N}\}$; then, by a result of Besicovitch and Taylor (see thm 5.2 in [22], or cor. 3.11 in [21])

  $$\dim^M(B) = 1/(t+1)$$

  (whereas the Hausdorff dimension of B is zero).

By (A.2), the entropy dimension $\dim^M(A)$ is always grater or equal than the Hausdorff dimensionality $\dim^{\mathcal{H}}(A)$; for this reason, any upper bound on $\dim^M(A)$ is in general a more stringent information than an upper bound on $\dim^{\mathcal{H}}(A)$.

A.2. **Yomdin's Sard–type statement**

To state the needed theorem, we need a more precise version of Taylor's theorem

Proposition A.1. *Let $f \in C^k$, let A be a bounded open set with $\overline{A} \subset E$, and let $\omega(r) = \omega(r, f|A, k)$ be the modulus of continuity of $D^k f$ on A, that is, the least positive concave increasing function satisfying*

$$\|D^k f(x) - D^k f(y)\| \leq \omega(|x-y|) \qquad \forall x, y \in A \ \ ; \tag{A.3}$$

[16]indeed, $\mathcal{M}^s(\cdot)$ is finitely subaddittive, but it is not addittive: if $n=1$, $A = [0,1] \cap \mathbb{Q}$, $B = [0,1] \setminus \mathbb{Q}$, then $\mathcal{M}^1(A) = \mathcal{M}^1(B) = \mathcal{M}^1([0,1]) = 1$; and similarly $\mathcal{M}^s(\cdot)$ is not σ–sub–addittive

let for convenience

$$R(r) = R(r, f|A, k) = \frac{1}{k!} r^k \omega(r, f|A, k) \tag{A.4}$$

Let p_y be the Taylor's polynomial of degree k centered at $y \in A$: then for all $x, y \in A$

$$|f(x) - p_y(x)| \leq R(|x - y|) = \frac{1}{k!} |x - y|^k \omega(|x - y|) \quad . \tag{A.5}$$

Yomdin uses the above theorem to finely approximate a generic function by polynomials; then he uses results in algebraic geometry (that he had proven in [23]) on the approximating polynomials, so that he proves, in section 4 in [22], that

Theorem A.2. *let $k \geq 1$, $\theta \in [0, 1]$, $f \in C^{k,\theta}$; for t small, if $\theta > 0$ then*

$$R(t)^\alpha M\Big(R(t), \Delta^l(f|B_r)\Big) \leq C' \|Df\|^l r^m \Big(\omega(t) t^{-\theta}\Big)^{\frac{m-l}{k+\theta}}$$

else if $\theta = 0$,

$$R(t)^\alpha M\Big(R(t), \Delta^l(f|B_r)\Big) \leq C' \|Df\|^l r^m \Big(\omega(t)\Big)^{\frac{m-l}{k}}$$

where $\alpha = \alpha(m, l, k + \theta)$, $C' = C'(m, n, k, l) > 0$; then, letting $t \to 0$, for any A such that $\overline{A}$ is compact and $\overline{A} \subset E$, we obtain

- *if $f \in C^k$, then $\mathcal{M}^\alpha(\Delta^l(f|A)) = 0$*
- *if $f \in C^{k,\theta}$, then $\omega(t) t^{-\theta} \leq \|D^k f\|_\theta$, and then $M(\varepsilon, \Delta^l(f|A)) = O(\varepsilon^{-\alpha})$ and*

$$\mathcal{H}^\alpha\big(\Delta^l(f|A)\big) \leq \mathcal{M}^\alpha\big(\Delta^l(f|A)\big) \leq C' \|Df\|^l r^m \big(\|D^k f\|_\theta\big)^{\frac{m-l}{k+\theta}} \quad .$$

As a corollary, if $f \in C^k$, we obtain $\mathcal{H}^\alpha(\Delta^l) = 0$, as prescribed by Federer's theorem 3.4.3 in [7].

A.3. **Extension to manifolds**

Suppose now that M, N are generic manifolds, of dimension m and n; then, if (N, g) is a Riemannian structure on N, if $A \subset N$, we can define $\mathcal{H}^s_g(A)$ and $\mathcal{M}^s_g(A)$ as above, using the distance on N provided by g.

Proposition A.3. *Suppose $F \subset N$ is compact. Let (N, g) and (N, g') be two Riemannian structures on N; there exists a constant $c > 0$ such that*

$$1/c\, \mathcal{H}^s_g(A) \leq \mathcal{H}^s_{g'}(A) \leq c \mathcal{H}^s_g(A)$$

and

$$1/c\, \mathcal{M}^s_g(A) \leq \mathcal{M}^s_{g'}(A) \leq c \mathcal{M}^s_g(A)$$

for all $A \subset F$

so we know that the "dimensionality" is not "locally" dependent on the particular Riemannian structure that we use; moreover, when the manifold N is not endowed with a Riemannian structure, the value $\mathcal{M}^\alpha(A)$ is not well defined, but the statements

- "$\mathcal{M}^\alpha(A) = 0$"
- "$0 < \mathcal{M}^\alpha(A) < \infty$"
- "$\mathcal{M}^\alpha(A) = \infty$"

are well defined; and similarly for $\mathcal{H}^\alpha$; then the Hausdorff dimension and the entropy dimension are well defined as well.

Similarly it can be shown that the class $C^{k,\theta}_{\text{loc}}$ is well defined, independently of the Riemannian structure used to define it.

Suppose $E \subset M$ is open, and $\overline{E}$ is compact. Let $f : E \to N$ be in $C^{k,\theta}_{\text{loc}}$; assign an arbitrary Riemannian structure (M, g) on M, and (N, g') on N; cover E and $f(E)$ by finitely many local charts; we can apply the previous results inside each local chart, and state that

Proposition A.4. *let $A \subset E$ be such that $\overline{A}$ is compact and $\overline{A} \subset E$; then*

- *if $f \in C^k(E \to M)$, then $\mathcal{M}^\alpha(\Delta^l(f|A)) = 0$ and $\mathcal{H}^\alpha(\Delta^l(f)) = 0$;*
- *if $f \in C^{k,\theta}(E \to M)$, then $\mathcal{M}^\alpha\big(\Delta^l(f|A)\big) < \infty$ and $\dim^{\mathcal{H}}(\Delta^l(f)) \le \alpha$*

Appendix B. Acknowledgement

The author thanks Alberto Abbondandolo, Piermarco Cannarsa and Carlo Sinestrari for the valuable discussions.

The author also thanks the Institut des Hautes Études Scientifiques (Bures-sur-Yvette, France), for the kind hospitality in Nov. 1999, and Prof M. Gromov, for the discussions on the topics that were at the origin of this appendix section.

The author thanks Prof Graziano Crasta for spotting the error that is corrected in this revised version.

References

[1] G. Alberti. On the structure of singular sets of convex functions. *Calc. Var.*, 2:17–27, 1994.
[2] A. Ambrosetti and G. Prodi. *A primer of nonlinear analysis.* Cambridge University Press, Cambridge, 1993.
[3] L. Ambrosio, P. Cannarsa, and H. M. Soner. On the propagation of singularities of semi–convex functions. *Ann. Sc. Norm. Sup. Pisa*, XX (4):597–616, 1993.
[4] Piermarco Cannarsa, Andrea Mennucci, and Carlo Sinestrari. Regularity results for solutions of a class of Hamilton–Jacobi equations. *Arch. Rat. Mech.*, 140:197–223, 1997. (or, preprint 13-95 Dip. Mat. Univ Tor Vergata, Roma).
[5] R. Courant and D. Hilbert. *Methods of Mathematical Physics*, volume II. Interscience, New York, 1963.
[6] Lawrence C. Evans. *Partial Differential Equations*, volume 19 of *Graduate Studies in Mathematics.* A.M.S.
[7] H. Federer. *Geometric measure theory.* Springer–Verlag, 1969.
[8] W. H. Fleming and H. M. Soner. *Controlled Markov processes and viscosity solutions.* Springer–Verlag, 1993.
[9] P. Hartman. *Ordinary Differential Equations.* Wiley, New York, 1964.
[10] J. Itoh and M. Tanaka. The Lipschitz continuity of the distance function to the cut locus. *Trans. A.M.S.*, 353(1):21–40, 2000.
[11] S. N. Kružhkov. The cauchy problem in the large for certain non-linear first order differential equations. *Soviet Math. Dockl.*, 1:474–475, 1960. MR 22 #12311.
[12] YanYan Li and Louis Nirenberg. The distance function to the boundary, Finsler geometry and the singular set of viscosity solutions of some Hamilton-Jacobi equations. *Comm. Pure Appl. Math.*, LVIII, 2005. (first received as a personal communication in June 2003).
[13] P. L. Lions. *Generalized Solutions of Hamilton–Jacobi Equations.* Pitman, Boston, 1982.
[14] C. Mantegazza and A. C. Mennucci. Hamilton–Jacobi equations and distance functions on Riemannian manifolds. *Applied Math. and Optim.*, 47(1):1–25, 2002. arXiv.org:math.AP/0201296.
[15] Dusa McDuff and Dietmar Salomon. *Introduction to Symplectic Topology.* Oxford Mathematical Monographs. Oxford University Press - Clarendon Press - Oxford, 1995.
[16] A. C. G. Mennucci. Regularity and variationality of solutions to Hamilton-Jacobi equations. part ii: variationality, existence, uniqueness. inviato Appl Math Optim 19/9/2005), 2004.
[17] C. Sinestrari P. Cannarsa. *Semiconcave functions, Hamilton-Jacobi equations and optimal control problems*, volume 58 of *Progress in Nonlinear Differential Equations and Their Applications.* Birkhauser Boston, 2004.
[18] G. J. Galloway P. T. Chruściel, J. H. G. Fu and R. Howard. On fine differentiability properties of horizons and applications to Riemannian geometry. *to appear*, 2001.
[19] C. Pignotti. Rectifiability results for singular and conjugate points of optimal exit time problems. *J. Math. Anal. Appl.*, 270(2):681–708, 2002.
[20] L. Simon. *Lectures on Geometric Measure Theory*, volume 3 of *Proc. Center for Mathematical Analysis.* Australian National University, Canberra, 1983.
[21] Y. Yomdin. β-spreads of sets in metric spaces and critical values of smooth functions.

[22] Y. Yomdin. The geometry of critical and near–critical values of differential mappings. *Mathematische Annalen*, 4(264):495–515, 1983.

[23] Y. Yomdin. Metric properties of semialgebraic sets and mappings and their applications in smooth analysis. In *Géométrie algébrique et applications, III (la Rábida, 1984)*, pages 165–183, Paris, 1987. Hermann.