Regularity and Variationality of Solutions to Hamilton-Jacobi Equations. part I: regularity

Andrea C. G. Mennucci
Scuola Normale Superiore
Piazza dei Cavalieri 7,
56126 Pisa, Italy*

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.

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.

*A preliminary abridged version of this paper was published in the proceedings [20] of the symposium "Advances in System Theory", by Kluwer Academic Publishers

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*.

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 $\mathcal{H}$ on a manifold M, see sec. §C.iii). See [8, 21] 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.ii); 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 1.9), 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.i 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 that is described in the appendixes). By assuming that H' is strongly convex in the argument $D_x w$, that $H' \in C^{R+1}$, in thm 4.10 and 4.12 in §4.3 in [4] it was proven that

Theorem 1.3 (4.10, 4.12 in [4]) *If $R \geq 2$, if $w_0 \in W^{1,\infty} \cap C^{R+1}(M')$ then the set Γ is $\mathcal{H}^{n-1}$ rectifiable.*

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

Theorem 1.4 (4.17 in [4]) *If $R \geq 2$, if $w_0 \in W^{1,\infty} \cap C^{R+1}(M')$ then the set $\Gamma \setminus \Sigma_u$ has negligeable $\mathcal{H}^{n-1+2/(R-1)}$ measure.* (1)

these results are improved in 5.1 in §5.i in this paper.

§1.ii 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 problem

$$\begin{cases} |du(x)|^2 - 1 = 0 & \text{on } M \setminus K \\ u = 0 & \text{on } K \end{cases} \tag{1.5}$$

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

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

$$\Sigma_u \cup \Gamma = \overline{\Sigma}_u = \mathrm{Cut}_K \tag{1.6}$$

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; otherwise, if K is not C^2, then the set Γ cannot be defined, and in this case

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

The above sets are known by many different names: the *skeleton*, the *ridge*, the *set of medial axes*.

Problem (1.5) has been extensively studied in many papers. Recently in [16] Mantegazza and M. have proved that the distance function d_K to the set K

$$d_K(x) = d(x,K) \doteq \min_{z \in K} d(x,z) \tag{1.7}$$

is the unique viscosity solution to (1.5), 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 (thm 3.4), and then d_K is C^1 in the open set $M \setminus (K \cup \overline{\Sigma}_u)$ (thm 3.5), and

Theorem 1.8 ([16] 4.7) *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 improved in section §5.ii.

In section §3 of the above paper there is also an example

(1) there is an errata in thm 4.17 in [4]: where it states "$\mathcal{H}^{n-1+2/R}$..." it should read "$\mathcal{H}^{n-1+2/(R-1)}$..."; the reason is that the map $\tilde{X}$ is of class C^{R-1}

Example 1.9 ([16] §3) *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.*

More recently, in [11], Itoh and Tanaka prove that, if K is a smooth submanifold, then

Theorem 1.10 ([11] 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 4.4 in this paper counts the focal points by rank, and can prove then a better result for the focal points of problem (1.1). Itoh and Tanaka prove also that

Theorem 1.11 ([11] B) *the distance to the cutlocus is locally Lipschitz;*

it is also interesting to note that, in the above example 1.9, the distance to the cutlocus is not Lipschitz (since Σ_u has infinite lenght in the region inside K).

In a unpublished work, Li and Nirenberg have proven the same result of 1.11 , 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).

Then in 2001 in [19] some regularity results where proved for Cauchy horizons in Lorentzian manifolds; deduced, as a special case, for the cutlocus, these results state that $\mathcal{H}^n(\mathrm{Cut}_K) = 0$, regardless of the regularity of K. This interesting result, as well as the above 1.11, is not covered by our theorems on problem (1.1).

§1.iii 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

$$W(t, x) = \inf \int_0^t L(\xi, \dot{\xi})dt + u_0(\xi(0)) \tag{1.12}$$

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 T^*M, such that, for any $x \in M$, the intersection $T^*_xM \cap \mathcal{U}$ is convex; minimize

$$W_{\mathcal{U}}(x) \doteq \inf \int_0^t L(\xi, \dot{\xi})dt + u_0(\xi(0)) \tag{1.13}$$

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 [9], or §1.4 in [15]); by adding some hypotheses, we may also assume that the above problem (1.12) has a minimum path ξ^*, for any x.

Similarly, the minimum $W_{\mathcal{U}}(x))$ of a problem like (1.13) provides usually a viscosity solution to a dual problem (1.1) (see chap 5 in [15], or chap 1 in [9]).

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 (1.12) 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 (3.7)); 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 (1.12), or (1.13), 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 [18]; other results may be found, e.g., in chap 5 in [15].

- ***(Regularity of (possibly non variational) H-J problems)*** the studies on regularity of viscosity solutions done in [4] and [16] were based mainly on properties of the characteristics' flow associated to H; these properties are not related to the variationality of the problem [(2)]; 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.iii)

§1.iv Layout

This is the layout of this paper.

- in section 3 we define the concept of *min* solution, that we use as a tool to express the regularity results;
- in sec. 4 we prove results on singular sets by studying the Hamiltonian flow;
- in sec. §4.iii we use these to prove results on regularity of generalized solutions to (1.1), such as viscosity solutions

[(2)]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 characteristics' flow (see sec. 4)

- in sec. 5 we apply the above results to the Cauchy type equation and to the distance function
- in appendixes A and B we review some definitions on symplectic manifolds and Hamiltonian flows
- in appendix C we discuss a Sard–type result by Yomdin

2 Prelims

§2.i 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 [(3)]; 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 $v \in T_x M$.

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.ii Viscosity solutions

Now we introduce the definition of viscosity solutions of PDE on manifolds. The concept of viscosity solutions is a recent one; under mild hypotheses, it provides a theory rich of: existence, uniqueness and stability of solutions, for first-order and second-order partial differential equations; we refer the reader to the reference [6] for a general introduction, and to the book [15] for a specific discussion of Hamilton–Jacobi equations and viscosity solutions. 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.

[(3)]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 $C^{(R,\theta)}_{\rm loc}$, the space of C^R functions whose R-th derivatives are *locally* Höelder of exponent θ

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*

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

Similarly, the set

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

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*

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

in Ω if for every $x \in \Omega$,

$$\begin{cases} H(x,v) \leq 0 & \forall v \in \partial^+ u(x), \\ H(x,v) \geq 0 & \forall v \in \partial^- u(x). \end{cases} \qquad (2.2.\star)$$

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; (eg, it is possible to express the above concept (2.2.$\star$) in local coordinates, using a suitable pull back of H). See §2 in [16].

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 I
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

$$\omega\Big((\dot x,\dot p),\ (\dot y,\dot q)\Big)\doteq\sum_i \dot q_i\dot x_i-\sum_i \dot p_i\dot y_i$$

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

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

More details are in §A.

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 *characteristics' flow*

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

as the solution of the system of ordinary differential equations

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

that is, (X,P) is the *Hamiltonian flow* for the symplectic manifold for (T^*M,ω), and U is the integral of ζ. More details are in §B.

§3.i Hypotheses

We define

$$Z\doteq\{H\le 0\}\doteq\{(x,p)\mid H(x,p)\le 0\}$$

Hypotheses 3.2 *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*

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

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

$$Z_x \doteq \{p \in T^*_x M \mid H(x,p) \leq 0\}$$

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 $\frac{\partial}{\partial p} H(x,p) = 0$. In particular,

Lemma 3.3 ($\{H = 0\}$ manifold) *if* (H1,H2) *holds, and Z_x does never contain an isolated point, then the set ∂Z coincides with $\{H = 0\}$, and $Z = \overline{(\mathring{Z})}$; $\{H = 0\}$ is a $(2n-1)$ submanifold of T^*M; moreover, $\{H = 0\}$ is decomposable in a certain number of connected components, and each component intersects all the fibers $T_x M$: these connected components can be thought of as* horizontal tubes*.*

§3.ii *min* solutions

We define the u_0*-annihilator*

$$TK^{\perp u_0} \doteq \{(z,p) \in T^*M \mid \; z \in K, p|_{T_z K} = du_0(z)\} \tag{3.4}$$

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

$TK^{\perp u_0}$ is a submanifold of T^*M, of dimension n; if $u_0 \equiv 0$, $TK^{\perp u_0}$ is a Lagrangian submanifold (see [17], 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

$$\begin{aligned} O \doteq & TK^{\perp u_0} \cap \{H = 0\} = \\ = & \{(x,p) \in T^*M \mid \; x \in K, \quad H(x,p) = 0, \quad \forall v \in T_x K, p \cdot v = du_0(x) \cdot v\} \end{aligned} \tag{3.5}$$

(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*

$$\Omega \doteq \{x \in M \mid x = X(s,z,q) \text{ for } s \geq 0, (z,q) \in O\} \tag{3.6}$$

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

We define the *min* solution u on Ω as

$$u(x) \doteq \min_{\begin{cases} t \geq 0, \; (z,q) \in O \\ \text{s.t. } X(t,z,q) = x \end{cases}} U(t,z,q) + u_0(z) \tag{3.7}$$

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

Hypotheses 3.8 *we will always assume that*

($\exists$**u**) *for any $x \in \Omega$, the above minimum (3.7) 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 3.9 *the hypothesis* (CC) *directly implies that* $u = u_0$ *on* K

Remark 3.10 *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., [17], sec. 9.4, and ex. 3.50). Note that* Λ' *is subset of a larger Lagrangian manifold*

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

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 section §6.ii).*

§3.iii 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], [10, §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* $\lambda : \Omega \to \mathcal{P}(\mathbb{R})$

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

We point out a difference between the classical "method of characteristics", and our definition (3.7): 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 throught a simple example: let $M = \mathbb{R}$, and consider the eikonal problem

$$\begin{cases} |du(x)|^2 - 1 = 0 \\ u(0) = u(1) = 0 \end{cases} \tag{3.11}$$

(see (1.5)) the solution obtained by the method of characteristic 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 characteristic, we obtain the solution in frame 2; to obtain the *min* solution $u(x)$, we apply to frame 2 the "*min* choice criterion",

(4) which is obviusly closed, as prescribed by 3.25 in [17]

(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 4.17.

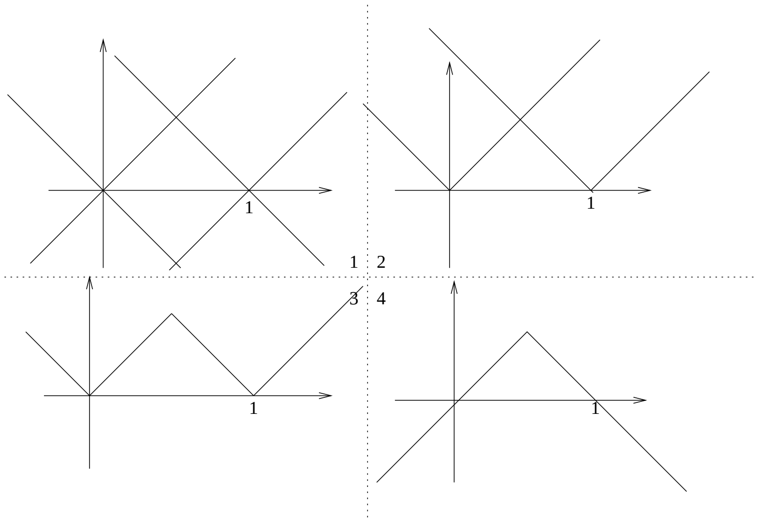


Figure 1: solutions to (3.11)

namely, we choose the minimum solutions proposed by the characteristic, 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 [18]) 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 [13]; see also [7], §3.3).

Rougly 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.iv Conditions

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

Conditions 3.12 *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;

[(6)]this is a condition on compatibility of H, K, u_0: see the condition in §5.1 in [15]

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

$$\{(t, y) \in \mathbb{R}^+ \times O \mid \ X(t, y) \in C, \ (U(t, y) + u_0(z)) \leq a\} \tag{3.12.$\star$}$$

is compact;

we will say that (OXUp) *holds in a open region $A \subset M$ if (3.12.$\star$) 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 3.13 *If* (OXUp) *holds, then* ($\exists\mathbf{u}$) *holds, that is, the* min *solution exists.*

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

Proposition 3.14 *if* (CC0) (H1,H2)*, hold, if $R \geq 1$, 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).*

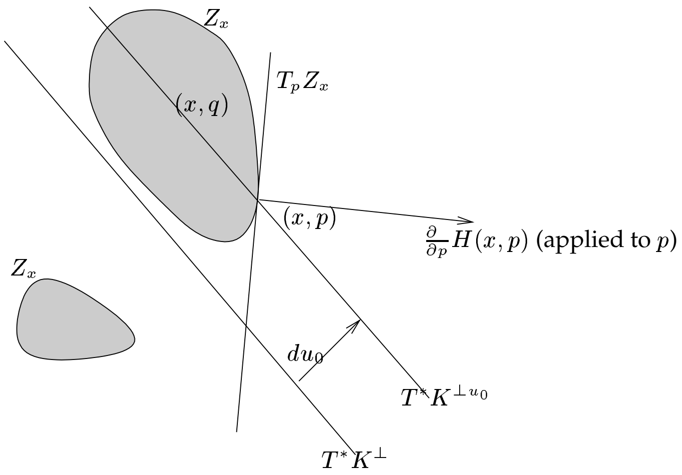


Figure 2: condition CC0, and examples of Z_x and $TK^{\perp u_0}$ in T_x^*M (the set Z_x is gray); (the element $\frac{\partial}{\partial p}H(x, p)$ of T_xM is represented by using scalar product duality)

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 3.17.
On condition (CC0), see also 4.16.

Remark 3.15 (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 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 3.16 *u is lower semicontinuous where* (OXUp) *holds locally.*

Proof. *Suppose that there are points $x_i \in \Omega \underset{i}{\longrightarrow} x$, then there are points $(t_i, y_i) \in \mathbb{R}^+ \times O$ (and $y_i = (z_i, q_i)$) s.t. $x_i = X(t_i, y_i)$ and $u(x_i) = U(t_i, y_i) + u(z_i)$; we extract a subsequence i_j such that $\lim_j u(x_{i_j}) = \liminf_i u(x_i)$; when $\liminf_i u(x_i) \neq \infty$, by* (OXUp) *we may assume (eventually passing to a new subsequence) that*

$$(t_{i_j}, y_{i_j}) \underset{j}{\longrightarrow} (t, y)$$

and then by continuity of X, U, u_0, we obtain $u(x) \leq U(t,y) + u_0(z) = \liminf_i u(x_i)$. □

§3.v 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 [18]). This example is based on the eikonal equation, that was defined in §1.ii.

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 (1.5).

Remark 3.17 *The condition that the set O (defined in (3.5)) be a regular submanifold of T^*M, implies that u_0 must satisfy the condition $|du_0| < 1$ (that is* (CC0)*). In this case, proposition 3.14 is inverted.*

Indeed, we can characterize O in the following way: for $x \in K$, we decompose T_xM in a normal part N_xK and a tangent part T_xK; we extend $du_0(x)$ to T_xM to be zero on N_xK; then we see that

$$O = \{(x,p) \mid p = du_0(x) + p', p' \in N^*K, |p'|^2 = 1 - |du_0(x)|^2\}$$

where

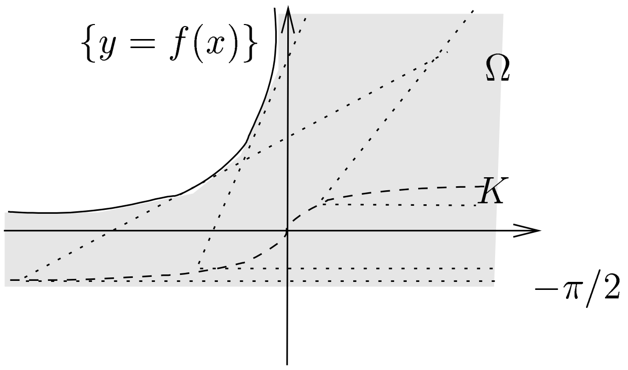
$$N^*K \doteq \{(x,p) \mid x \in K,\ \forall v \in T_xK,\ p \cdot v = 0\} \quad .$$

Note that O is, in a sense, a system of cylindrical coordinates *around K.*

---

[7] which applies to the case $\dim(K) = \dim(M) - 1$; see eq. 7.16 in [10, § VI]

Figure 3: Example 3.18; characteristics are dotted, K is dashed, Ω is gray



Example 3.18 *Let $M = \mathbb{R}^2$ and $H(x,p) = |p|^2 - 1$ as above, so that the flow is*

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

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

$$du_0(x) = \begin{pmatrix} (1+x_1^2)^2 \\ 1+x_1^2 \end{pmatrix} \frac{1}{1+(1+x_1^2)^2} \quad ;$$

and, by definition,

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

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

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

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 4.17).

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) [(8)] are satisfied; and (OXUp) is not satisfied [(9)] (but (OXUp) holds inside the region Ω) ; ($\exists\underline{\text{u}}$)

[(8)]for any $x \in \Omega$, there is only a limited span of characteristics crossing it: indeed if $y = X = x + 2tp$ and $x_1 \to -\infty$ while y_1 is kept fixed, then $y_2 \to -\pi/2$

[(9)](take $C \subset M$ to be a compact set containing a point $(x, -\pi/2)$ in its interior, and $a = 2|x| + 1$, in the equation (3.12.$\star$) in the definition of (OXUp)) (def. in the second part [18]) is satisfied, by choosing

$$\underline{u}(x) = -\sqrt{|x|^2+1} \tag{3.19}$$

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

Example 3.20 *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 $n \doteq \dim(M) \geq 2$.

We start by defining the singularity sets

We define the set Σ to be the set of $x \in \Omega$ s.t. the minimum (3.7) that defines $u(x)$ is given by at least two points in $\mathbb{R}^+ \times O$, that is

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

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

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

(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 $\Sigma_u \doteq \{x \mid \not\exists du(x)\}$ 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

$$G^{(i)} \doteq \left\{ (s,y) \in \mathbb{R} \times O \mid \frac{\partial X}{\partial(s,y)}(s,y) \text{ has rank } n-i \right\} \tag{4.3}$$

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 $\frac{\partial}{\partial t} X = \frac{\partial}{\partial p} H(X,P) \neq 0$ (see 3.3).

We will use

- the Hausdorff measure $\mathcal{H}$ 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 §C.iii.

We will always assume (∃u) and (CC) (for prop. 3.9).

§4.i Rectifiability of conjugate points

We will now prove that the set Γ of conjugate points is rectifiable, by proving that the set of all focal points is rectifiable: we will prove that

Theorem 4.4 *Assume* (CC0,H1,H2), $u_0, K \in C^{2+s}$, *with* $s \in \mathbb{N}$. *If* H *is as regular as specified in the following table, then 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* $h = 1 \dots$*) where* $A_{i,h} \subset O$ *are open sets; the regularity of* $\lambda_{i,h}$ *is related to the regularity of* H*, and to the dimension* $\dim(M) = n$*, as follows:*

	H	λ
$n=2$	$C^{(2,1)}$	$loc.Lipshitz$
$n=2$	C^{s+3}	C^{s+1}
$n \geq 3$	C^{s+n}	C^{s+1}

(4.4.⋆)

So there follows from C.6 that the set of focal points is rectifiable, and so is the set Γ of conjugate points: for example, when $n \geq 3$ and $H \in C^{n+s}$, the focal points can be covered by C^{s+1} submanifolds of M, but for a set that has Hausdorff measure $\mathcal{H}^\alpha$ zero, where $\alpha \doteq n - 2 + 1/(s+1)$; and similarly for the other cases.

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.ii.1.

In the next section we will instead prove in 4.12 that, under certain hypotheses, the set Γ has dimensionality at most $n - 2 + 2/(r + \theta - 1)$.

In the example 3.18 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 4.4 is fairly optimal.

We will also prove that

Theorem 4.5 *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* Γ *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 4.6 *Let $r \in \mathbb{N}$. 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 $\frac{\partial F}{\partial t}$ exists, and $F \in C^{r,\theta}, \frac{\partial F}{\partial t} \in C^0$; if $\frac{\partial F}{\partial t}(t,x) \neq 0$, in a neighbourhood V of (t,x) the zero set $\{F = 0\}$ coincides with the graph $\{(\phi(x),x)\}$ of a Hölder function $\phi \in C^{(r,\theta)}$: we will then say that* $V \cap \{F = 0\}$ can be covered by a manifold of class $C^{(r,\theta)}$. *As a corollary, if $r = 0$, $V \cap \{F = 0\}$ has Hausdorff dimension at most $(n - \theta)$, and, if $\overline{V}$ is compact, then $\mathcal{M}^{n-\theta}(V \cap \{F = 0\}) < \infty$.*

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

Lemma 4.7 *Assume* (CC0,H1,H2); *assume $R' \in \mathbb{N}, R' \geq 1$, $u_0 \in C^{R'+1}$, $K \in C^{R'+1}$; assume $R \in \mathbb{N}, R \geq 1$, $H \in C^{(R+1,\theta)}$; then by 3.14, O is a $C^{R'} \cap C^{R+1}$ manifold, and the flow $\Phi = (X,P)$ is $C^{(R,\theta)}$ regular as a map from $\mathbb{R} \times T^*M \to T^*M$.*

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

To study $G^{(i)}$, we define locally n indipendent regular vector fields

$$e_1 \dots e_n$$

in a neighbourhood of (s',y') in $\mathbb{R} \times T^*M$; we define then the tangent n-form

$$\alpha = e_1 \wedge \cdots \wedge e_n$$

We choose $e_1 \dots e_n$ such that $\alpha(t,y')$ spans the tangent space $T_{(t,y')}(\mathbb{R} \times O)$, and such 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. (6.6). Note that, by hypotheses and by the definition (6.6) of $X^{(t,y)^*}\alpha$, the forms $X^{(t,y)^*}\alpha$ and $\frac{\partial}{\partial t}(X^{(t,y)^*}\alpha)$ have regularity $C^{(R-1,\theta)}$ (as shown by eq. (6.7)): we then define

$$r = (R-1) \wedge (R-i+1) = \begin{cases} R-i+1 & \textit{for } i \geq 2, \\ R-1 & \textit{when } i = 1 \end{cases}$$

so we can always say that the form $\frac{\partial^{i-1}}{\partial t^{i-1}} X^{(s',y')^*}\alpha$ has regularity $C^{(r,\theta)}$.

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

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

whereas

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

To restrict y to move on O, we choose a local coordinate system $\phi : \mathbb{R}^{n-1} \to O$ and define the map $F : \mathbb{R} \times \mathbb{R}^{n-1} \to \mathbb{R}$ given by by

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

which has regularity $C^{r,\theta} \cap C^{R'}$ with

Since

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

the above Dini lemma implies that the set $G^{(i)}$ is locally covered by the graph of a function λ_i of class $C^{r,\theta} \cup C^{R'}$ (that is, the least regular of the two), and defined on a open subset of O.

The above directly implies 4.4.

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 C.6, due to Yomdin:

Lemma 4.8 *let $i \geq 1, i \leq n$, assume $O \in C^R$ 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, 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 $n - (R - 1 + \theta)$, which is always less than $n - 1 + 1/(R+\theta)$.

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, the set $X(G^{(i)})$ is rectifiable, so it is at most $n-1$ dimensional; but, by lemma 4.8, we also obtain that its dimension actually does not exceed $n - i + i/(R+\theta)$, which decreases rapidly when $R + \theta$ is large
- whereas, if $R < i$, then we can only use lemma 4.8; 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 $O \in C^1$, 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 4.5, we can apply the lemma 4.8 to

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

and apply the lemma 4.7 to

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

(and then apply lemma 4.8 to their image under X). This proves 4.5.

Note that the above is actually a study of the "regularity" of the lagrangian submanifold Λ (defined in 3.10), 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 [17].

Remark 4.9 *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.ii 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 $\mathring{\Omega}$ (10). 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 4.7, the sets $G^{(i)}$ are locally covered by graphs.

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

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

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 $\nu(y) \in T_{(0,y)}(\mathbb{R} \times O)$, (11) 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. (12)

First, we prove a lemma

Lemma 4.10 *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*

$$\frac{\partial^2 H}{\partial q^2} \left(\nu_1 - \sum_{i=2}^{n} \frac{d\phi}{dy_i} \nu_i \right) = \sum_{h=1}^{n} \frac{\partial^2 X_h}{\partial \nu^2} q_h \doteq q \cdot \frac{\partial^2 X}{\partial \nu^2} \tag{4.10.$\star$}$$

where

$$q_h(y) = \sum_j \frac{\partial P_h}{\partial \overline{x}_j}(\phi(y), y)\nu_j(y) \tag{4.10.$\star\star$}$$

and

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

---

(10) we will show in next section that $\mathcal{A} \supset \Omega \setminus (\Gamma \cup K)$; the remark 3.20 shows that u may be discontinuous across Γ

(11) ν is a vector field on the graph of ϕ

(12) 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 $\nu(y)$

and similarly for $\frac{\partial^2 X_h}{\partial \nu^2}$

Proof. *Indeed, we differentiate*

$$\sum_j \frac{\partial X_h}{\partial \overline{x}_j}(\phi(y), y)\nu(y) = 0, \quad \forall h$$

by y_i*, then we multiply by* $q_h \nu_i$ *and sum over* h, i*, and note that*

$$\sum_h \frac{\partial X_h}{\partial \overline{x}_j} q_h = \sum_{h,r} \frac{\partial X_h}{\partial \overline{x}_j} \frac{\partial P_h}{\partial \overline{x}_r} \nu_r = \sum_{h,r} \frac{\partial X_h}{\partial \overline{x}_r} \frac{\partial P_h}{\partial \overline{x}_j} \nu_r = 0 \qquad (4.10.\diamond)$$

(by 6.3); so we prove, (exactly as in lemma 4.18 in [4]) that

$$\left(\sum_{r,s=1}^{n} \frac{\partial^2 H}{\partial p_r \partial p_s} q_r q_s \right) \sum_{i=2}^{n} \frac{d\phi}{dy_i} \nu_i + \sum_{h,j=1}^{n} \sum_{i=2}^{n} \frac{\partial^2 X_h}{\partial y_i \partial \overline{x}_j} \nu_i \nu_j q_h = 0 \qquad (4.10.\diamond\diamond)$$

but it is also easy to prove (using again (4.10.⋄)) that

$$\sum_{h,j=1}^{n} \frac{\partial^2 X_h}{\partial t \partial \overline{x}_j} \nu_1 \nu_j q_h = \nu_1 \sum_{h,j=1}^{n} \frac{\partial}{\partial \overline{x}_j} \left[\frac{\partial H}{\partial p_h}(X, P) \right] \nu_j q_h = \nu_1 \sum_{r,s=1}^{n} \frac{\partial^2 H}{\partial p_r \partial p_s} q_r q_s$$

and we just need to sum and subtract. □

then we prove that

Lemma 4.11 *if* (OXUp)*, if* $x' = X(t', y') \in \Gamma$*, if* t', y' *minimizes* $u(x')$*, if* $(t', y') \in G^{(1)}$*, if* V, ϕ, ν *are defined as above, and* $q(y)$ *is defined as in (4.10.⋆⋆), and if* $x' \notin \Sigma$*, and if* $x' \in \mathring{\Omega}$*, and if* u *is continuous in* x'*, then*

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

Proof. *Suppose that* $\left(q \cdot \frac{d^2 X}{d\nu^2}\right) \neq 0$*, whereas, we know that for any* ν'*,* $q \cdot \frac{dX}{dx} \nu' = 0$*, (see (4.10.⋄) and (6.9)): 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

Theorem 4.12 *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,* $K, u_0 \in C^{R'+1}$*,* $H \in C^{R+1,\theta}$*,* $R = R' \geq 1$*;*
then the set $\Gamma \cap \mathcal{A} \setminus \Sigma$ *has dimensionality at most* $n - 2 + 2/(R + \theta - 1)$

Proof. Let $x \in (\Gamma \cap \mathcal{A}) \setminus \Sigma$, and let (t', y') be the optimal point for $u(x)$; if $(t', y') \in G^{(2)} \cup \ldots G^{(n)}$, then we directly apply Sard–type Lemma C.6; 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 4.10,

$$\sum_{i=2} \frac{\partial \phi}{\partial y_i}(y)\nu_i(y) = \nu_1 \tag{4.12.$\star$}$$

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 (4.12.$\star$) 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

$$rank \frac{d}{dy}\tilde{X} = -1 + rank DX = n - 2$$

The map ϕ has regularity $C^{R-1,\theta} \cap C^{R'}$, while $O \in C^{R+1} \cap C^{R'}$, so the map $\tilde{X} : O \to M$ has class $C^{R-1,\theta} \cap C^{R'}$, so we conclude by applying lemmaC.6 to $\tilde{X}$. □

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 [18] of this paper); if the problem is not variational, then the theorem is false outside of $\mathcal{A}$, as shown by remark 3.20.

Remark 4.13 *The above three theorems 4.4, 4.5 and 4.12, need the hypothesis* (H2)*; but, it is clear from the proofs that this may be weakened to these hypotheses: "O* is a manifold" *and "$\forall \mu \in \ker DX$, $\mu \neq 0$,* if $\nu = DP\mu$, then $\frac{d^2 H}{d\nu^2} > 0$"

§4.iii 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 4.14 *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*

i). $\Gamma \cup K$ is closed in Ω, and $\Omega \setminus (\Gamma \cup K)$ is open;

ii). 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

iii).

$$u \in C^{R+1}(\Omega \setminus (\Gamma \cup \Sigma \cup K))$$

and u is a regular solution of (1.1) in the open set $\Omega \setminus (\Gamma \cup \Sigma \cup K)$;

iv). 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.*

Proof. *We define $\underline{U}(t, z, q) = U(t, z, q) + u_0(z)$ for convenience.*

We will use in the following this line of reasoning: if $x \notin \Gamma$, then we choose $(t, z, p) \in \mathbb{R}^+ \times O$ s.t. $X(t, z, p) = x$, $\underline{U}(t, z, p) = u(x)$; we have that $DX(t, z, p)$ is invertible, and, having $R \geq 1$ and $X \in C^R$, then X is invertible in a neighbourhood of (t, z, p); if moreover $x \notin K$, then we know that $t > 0$, so we may assume that this neighbourhood does not contain negative times.

This proves that $\Omega \setminus (\Gamma \cup K)$ is open.

We prove first the last points

iii) *The crucial point (see [4, thm 3.9]) is that, if $x \notin \Gamma$, $x \notin K$, then the set*

$$A \doteq \{(t, z, p) \in \mathbb{R}^+ \times O \mid X(t, z, p) = x,\ U(t, z, p) + u_0(z) = u(x)\}$$

is finite; indeed, A is compact, (by (OXUp)*) and it is discrete, by the above reasoning on invertibility.*

For the first claim, we note that when $x \notin \Sigma$ we have that A is a singleton: the result follows then from the second claim.

iv) *For the second claim, we invert $X(t, y)$ in a neighborhood B_i of any point (t_i, y_i) of A, to obtain a function X_i^{-1}; we can then define $u_i(x') = \underline{U}(X_i^{-1}(x'))$ in a neighborhood $X(B_i)$ of x; from the "Cauchy's method of characteristics", we have $H(x, du_i(x)) = 0$ (see B.6).*

We prove we can choose $B \subset X(\cup_i B_i)$ s.t. $u(x') = \min_i u_i(x')$ for $x' \in B$. For any sequence $x_n \underset{n}{\longrightarrow} x$, let (t'_n, y'_n) be s.t. $u(x_n) = \underline{U}(t'_n, y'_n)$ and $x_n = X(t'_n, y'_n)$; assume on the contrary that there is a sequence $x_n \to x$ s.t. $u(x_n) < \min_i u_i(x_n)$; then $(t'_n, y'_n) \notin \cup_i B_i$; by 3.16 u is l.s.c, so we get $\lim_n u(x_n) = u(x)$.

By (OXUp)*, up to a subsequence, we may assume that $(t'_n, y'_n) \underset{n}{\longrightarrow} (t', y')$; so $u(x) = \underline{U}(t', y')$; then we would have $(t', y') = (t_i, y_i)$ for a i, contradicting the fact that $(t'_n, y'_n) \notin \cup_i B_i$.*

[13] Note that the example 3.18 shows that we may have $\overline{\Sigma} \neq \Sigma \cup \Gamma$.

Since $du_i(x) = P(X_i^{-1}(x))$, then $u_i \in C^{R+1}$; moreover if we had $du_i(x) = du_j(x)$, then we would have $P(t_i, y_i) = P(t_j, y_j)$, $X(t_i, y_i) = X(t_j, y_j) = x$, so that by uniqueness of solutions of ODEs, $(t_i, y_i) = (t_j, y_j)$: possibly choosing a smaller neighborhood, we obtain that if $i \neq j$, $du_i(x') \neq du_j(x')$ in B.

Now we prove the two points i), ii)

i) *let x_n be a sequence of points in $\Gamma \cup K$ and suppose that $x_n \to x \in \Omega$. Since K is closed, we may assume that $x_n \notin K$ definitively (otherwise, $x \in K$).*

There exist pairs t_n, y_n such that $X(t_n, y_n) = x_n$, and $\underline{U}(t_n, y_n) = u(x_n)$.

We may assume (passing to a subsequence, and by (OXUp)*) that $\lim u(x_n) = \liminf u(x_n)$, $t_n, y_n \to t, y$: then, $u(x) \leq \lim u(x_n)$*

If $\lim u(x_n) = u(x)$, then we are done: indeed, by continuity of the determinant, $\frac{\partial}{\partial t,y} X(t, y)$ is not full rank, so that $x \notin \Gamma$

We seek a contradiction by supposing that $x \notin \Gamma \cup K$, and $\lim u(x_n) < u(x)$: then, u would not be continuous in x, contrary to the above reasonings.

ii) *let x_n be a sequence of points in Σ and suppose that $x_n \to x \in \Omega$; if we have $x \in \Gamma \cup K$, we have nothing to prove.*

There exist two different pairs t_n, y_n and t'_n, y'_n such that $X(t_n, y_n) = x_n = X(t'_i, y'_i)$, and $\underline{U}(t_n, y_n) = u(x_n) = \underline{U}(t'_n, y'_i)$.

When we have $x \notin \Gamma \cup K$, we seek a contradiction by supposing also that $x \notin \Sigma$: let (t, y) be the unique point that gives the minimum $u(x)$.

We know by the previous discussions that u is continuous in a neighborhood of (t, y); so we have $u(x_n) \to u(x)$; so the limit point of (t_n, y_n) is a minimum (in the formula that defines $u(x)$), so we obtain that $(t_n, y_n) \to (t, y)$ and $(t'_n, y'_n) \to (t, y)$. But there is a neighborhood of (t, y) where X is a local diffeomorphism: then we should have $(t_n, y_n) = (t'_n, y'_n)$ definitively.

□

Corollary 4.15 *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 $\Omega \setminus \Gamma$.*

Proof. Indeed, we have seen that u is locally the minimum of a finite number of regular solutions, and then u is locally semiconcave. By the previous proposition, u is a classical solution almost everywhere in $\Omega \setminus \Gamma$; by (H2)*, $H(x, \cdot)$ is convex: the result follows from the theory of viscosity solutions.* □

Remark 4.16 *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 4.17 (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 (3.7), and using O^+ instead of O: this alternative definition u^+ of* min *solution would be then*

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

The above theorem, when applied to u^+, would change slightly: the third and fourth statement would extend to the sets $\Omega \backslash (\Gamma \cup \Sigma)$ and $\Sigma \backslash \Gamma$, respectively. [14] *In particular, in the example 3.18, 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.i the Cauchy problem

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

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

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 $x = (t',x')$, $dw = (\frac{\partial}{\partial t'} w, \frac{\partial}{\partial x'} w)$, $p = (\tau, p')$; we define H by

$$H(x,p) = \tau + H'(t',x',p')$$

so that

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

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

The set O is, in this case,

$$O = \left\{ (t',x',\tau,p') \in T^*M \;\middle|\; (t',x') \in K,\ \tau = H'(t',x',p'),\ p' = \frac{\partial}{\partial x'} w(t',x') \right\} =$$
$$= \left\{ \left(0, x', H'(t',x',p'), \frac{\partial}{\partial x'} w(0,x')\right) \in T^*M \;\middle|\; x' \in M' \right\}$$

which is obviously a manifold; the set O codes the initial conditions for the characteristics' curves of problem (1.2): $X'(0,x',p') = x'$, $P'(0,x',p') = p' = \frac{\partial}{\partial x'} w(t',x')$.

The function H does not satisfy (H2): indeed, $\frac{\partial^2}{\partial \tau^2} H = 0$; more in general, note that, if $\nu \in TT^*M$ we decompose $\nu = (\nu'', \nu')$ with $\nu'' \in \mathbb{R}$, $\nu' \in TT^*M'$, and we obtain $\frac{\partial^2}{\partial \nu^2} H = \frac{\partial^2}{\partial \nu'^2} H'$

[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 (3.7); but we have would have anyway that $\Gamma \supset K$, so that in general u would not be regular on K

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 $\nu' \neq 0$; we will use this to apply remark 4.13.

We set $\nu = DP\mu$, and $\nu = (\nu'', \nu')$ with $\nu'' \in \mathbb{R}$, $\nu' \in TT^*M'$. If $\mu \in KerDX$ and $\mu \neq 0$, then $\nu = DP\mu \neq 0$ (by the rank argument 6.1).

The characteristics' strips $(X(t,z,q), P(t,z,q))$ for 1.1 (defined in (B.4)) are easily related to the characteristics' strips $(X'(t',z',q'), P'(t',z',q'))$ for 1.2.

If we write the characteristics' equation (B.4) by dividing the first component (in $\mathbb{R}$) from the second component (in M') (and similarly for P), we obtain $X = (T, X')$, $P = (P^t, P')$: then $T(t,z,q) = t$, while

$$\dot{P}^t = -\frac{\partial}{\partial t'} H'(t', X', P')$$

We decompose DX as (DT, DX') and $\mu = (\mu'', \mu')$ as $\mu'' \in \mathbb{R}$, $\mu' \in TT^*M'$, and then $\nu'' = DT\mu = 0$, so that $\nu' \neq 0$; this shows that the remark 4.13 holds.

Summarizing

Proposition 5.1 *To apply the theorems 4.4, 4.5 and 4.12 to the Cauchy problem, we substitute the hypothesis* (H2) *by the hypothesis "$\frac{\partial^2}{\partial p'_i p'_j} H'$ is positive definite"*

This improves the results of 4.10 4.12 4.17 in [4] (reported as 1.4 and 1.3 in this paper); indeed, the above shows that,

- if $H' \in C^s(T^*M)$ with $s = n \vee 3$ and $w_0 \in C^2(M)$, then the set Γ is rectifiable;
- and when $H' \in C^{R+1,\theta}(T^*M)$, $w_0 \in C^{R+1,\theta}(M)$, $R \geq 1$, we prove that $\mathcal{H}^{n'-1+2/(R+\theta-1)}(\Gamma \setminus \Sigma) = 0$ where $n' = \dim(M')$.

Note that, indeed, 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.ii Eikonal equation and cutlocus

As in §3.v, 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

$$\Sigma_{d_K} \doteq \{x \mid \nexists \nabla d_K(x)\}$$

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.

In the example in §3 in [16] (see 1.9 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. The above theorem 4.4 states that if K is C^2, then Γ is rectifiable, so by (1.6) and 4.14.(iv), we obtain that $\overline{\Sigma}_{d_K}$ is rectifiable; so again 4.4 closes the gap between the counterexample and the rectifiability result.

Regarding theorem A in [11] (see 1.10 here), we can note that, by dividing the focal points according to their rank, if K is a smooth submanifold, then theorem 4.4 finds smooth functions $\lambda_{i,h}$ that cover the times t at which there are focal points of rank $n-i$.

§5.ii.1 Improvements

We want to show an improvement, based on theorem 4.4, 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 (the fiber is $\{-1, 1\}$ if K is orientable, and it is $\{0\}$ if K is not orientable); 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.

Let $s \in \mathbb{N}$. By theorem 4.4, we know that, if $K \in C^{2+s}$, 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 C.6 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^{2+s} 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 $M^{1/(s+1)}(E) = 0$.

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 4.7

We will now prove some results needed for 4.7 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 3.14, 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 (B.4); by well known properties of linear systems (see e.g. [10, pg. 46]), for any $z, q \in O$

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

Another way to prove this result, is by noting that the form

$$\omega \wedge \cdots \wedge \omega = n!\, dx_1 \wedge dp_1 \wedge \cdots \wedge dx_n \wedge dp_n$$

is preserved by the action of the flow.

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 3.15). We have proved that the rank in (6.1.$\star$) is n, for any $t \geq 0$.*

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

To continue, we will better clarify the

Remark 6.2 (double use of time:) *in the following, we will consider the flow $\Phi = (X, P)$ as a map from $(t, y) \in \mathbb{R}\times O$ into M; to this purpose, we will use the combined variable $\overline{x}\doteq(t, y) \in \mathbb{R}^+ \times O$ for convenience. In this first case, it is better to think of t as an obvious choice of local coordinate for "$\mathbb{R}$" in the manifold $\mathbb{R} \times O$ The differential $\frac{\partial X}{\partial \overline{x}}$ maps the tangent $T_{\overline{x}}(\mathbb{R} \times O) = \mathbb{R} \times T_y O$ into $T_{X(\overline{x})}M$. We will then derive some quantity, based on $\frac{\partial X}{\partial \overline{x}}$, again wrt to time. In this second case, it is better to think of t as the time at which we are studying the flowing of Φ.*

§6.i Proof of 4.7, case $i = 1$

First, we prove the special case $i = 1$, to show a proof based more on analytical methods: this proof follows the similar proof done in [4], where the model case was the Cauchy type equation (1.2).

Proposition 6.3 *We have that*

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

is symmetric

We delay the proof of this result to section §6.ii.1.

Lemma 6.4 *Suppose $n > 1$. Let* rank $\frac{\partial X}{\partial \overline{x}}(\overline{x}) = n - 1$ *for some $\overline{x}$ and let ν be a generator of* $\mathrm{Ker}\frac{\partial X}{\partial \overline{x}}(\overline{x})$*. Then $\frac{\partial P}{\partial \overline{x}}(\overline{x})\nu$ generates the orthogonal complement of* $\mathrm{Im}\frac{\partial X}{\partial \overline{x}}(\overline{x})$ *and*

$$\left(\frac{\partial X}{\partial \overline{x}}\right)^+ (\overline{x}) = c\nu\nu^t \frac{\partial P}{\partial \overline{x}}^t (\overline{x}) \tag{6.4.$\star$}$$

for some $c \in \mathbb{R}, c \neq 0$. Here A^+ is the adjoint of a matrix A, A^t is the transpose of a matrix A; vectors are considered to be column vectors.

The result is based on the previous proposition; see [4, Lemma 4.14].

We have, for any (t, y), (we omit for simplicity the dependence of H on $(X(t,z,q), P(t,z,q))$)

$$\begin{aligned}
\frac{\partial}{\partial t} \det \frac{\partial X}{\partial \overline{x}}(\overline{x}) &= \operatorname{trace}\left(\frac{\partial^2 X}{\partial t \overline{x}}(\overline{x}) \frac{\partial X}{\partial \overline{x}}^+(\overline{x})\right) \\
&= \operatorname{trace}\left(\frac{\partial^2 H}{\partial px}\frac{\partial X}{\partial \overline{x}}(\overline{x}) \frac{\partial X}{\partial \overline{x}}^+(\overline{x})\right) + \operatorname{trace}\left(\frac{\partial^2 H}{\partial pp}\frac{\partial P}{\partial \overline{x}}(\overline{x}) \frac{\partial X}{\partial \overline{x}}^+(\overline{x})\right) \\
&= \det \frac{\partial X}{\partial \overline{x}}(\overline{x}) \operatorname{trace}\left(\frac{\partial^2 H}{\partial px}\right) + \operatorname{trace}\left(\frac{\partial^2 H}{\partial pp}\frac{\partial P}{\partial \overline{x}}(\overline{x}) \frac{\partial X}{\partial \overline{x}}^+(\overline{x})\right).
\end{aligned}$$

(as in equation *(4.8)*, [4]).

Now let us suppose that $\operatorname{rank}\frac{\partial X}{\partial \overline{x}}(\overline{x}) = n-1$. Then, if ν is the generator of $\operatorname{Ker}\frac{\partial X}{\partial \overline{x}}(\overline{x})$, we have by Lemma 6.4,

$$\frac{\partial}{\partial t} \det \frac{\partial X}{\partial \overline{x}}(\overline{x}) \quad = \quad c\left(\frac{\partial P}{\partial \overline{x}}(\overline{x})\nu\right)^t \frac{\partial^2 H}{\partial pp}\left(\frac{\partial P}{\partial \overline{x}}(\overline{x})\nu\right) \neq 0,$$

since $\frac{\partial^2 H}{\partial pp}$ is positive definite and

$$\frac{\partial P}{\partial \overline{x}}(\overline{x})\nu \neq 0 \quad .$$

On the other hand, if $\operatorname{rank}\frac{\partial X}{\partial \overline{x}}(\overline{x}) < n-1$, then $\frac{\partial X}{\partial \overline{x}}^+(\overline{x}) = 0$ and thus the derivative of $\det \frac{\partial X}{\partial \overline{x}}$ is zero.

As a consequence of the implicit function theorem, we conclude that the set $G^{(1)}$ defined in (4.3) is locally a graph, that is for any $(t_0, y_0) \in G^{(1)}$, there exist a neighborhood B of y_0 in O, $\varepsilon > 0$, and a map $\varphi : B \to]t_0 - \varepsilon, t_0 + \varepsilon[$ of class C^{R-1}, such that

$$\det \tfrac{\partial X}{\partial t,y}(t,y) = 0 \quad \text{if and only if } t = \varphi(z) \tag{6.5}$$

for any $(t, y) \in]t_0 - \varepsilon, t_0 + \varepsilon[\times B$. If we choose B small enough, we may also assume that $\frac{\partial X}{\partial t,y}(\varphi(z), z)$ has rank $n-1$ for any $z \in B$, by lower semi-continuity of the rank. This concludes the proof of the special case $i = 1$ for 4.7.

§6.ii Proof of 4.7, general case

Now, we prove the general case of 4.7, and show a proof based more on geometrical methods.

We define $\alpha = e_1 \wedge \cdots \wedge e_n$, as in 4.7; moreover we choose it such that the vectors $e_{n-i+1}(y') \dots e_n(y')$ span the kernel of $\frac{\partial}{\partial \overline{x}}X(t,y)$ (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). We write $\overline{x} = (t, y)$, and $DX^{(s',y')}$ for the differential $\frac{\partial X}{\partial \overline{x}}$ computed in s', y'.

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

$$X^{(s,y)^*}\alpha \doteq DX^{(s,y)}e_1 \wedge \cdots \wedge DX^{(s,y)}e_n \tag{6.6}$$

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

$$\frac{\partial}{\partial s}X^{(s,y)^*}\alpha = \sum_j DX^{(s,y)}e_1 \wedge \cdots \wedge DX^{(s,y)}e_{j-1} \wedge \frac{\partial}{\partial s}(DX^{(s,y)}e_j) \wedge \qquad (6.7)$$

$$\wedge DX^{(s,y)}e_{j+1} \wedge \cdots \wedge DX^{(s,y)}e_n =$$

$$= \sum_j DX^{(s,y)}e_1 \wedge \cdots \wedge DX^{(s,y)}e_{j-1} \wedge \left(\frac{\partial^2 H}{\partial px}\frac{\partial X}{\partial \overline{x}} + \frac{\partial^2 H}{\partial pp}\frac{\partial P}{\partial \overline{x}}\right) e_j \wedge$$

$$\wedge DX^{(s,y)}e_{j+1} \wedge \cdots \wedge DX^{(s,y)}e_n$$

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

$$i!\, DX^{(s,y)}e_1 \wedge \cdots \wedge DX^{(s,y)}e_{n-i} \wedge \frac{\partial}{\partial s}(DX^{(s,y)}e_{n-i+1}) \wedge \ldots \qquad (6.8)$$

$$\cdots \wedge \frac{\partial}{\partial s}(DX^{(s,y)}e_n)$$

which may be nonzero at (s', y'). We now study this term at point (s', y'); we write (s, y) for (s', y'), for sake of simplicity.

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

$$i!\, DX^{(s,y)}e_1 \wedge \cdots \wedge DX^{(s,y)}e_{n-i} \wedge (\frac{\partial^2 H}{\partial pp}DP^{(s,y)}e_{n-i+1}) \wedge \ldots$$

$$\cdots \wedge (\frac{\partial^2 H}{\partial pp}DP^{(s,y)}e_n)$$

The rank of $DX^{(s,y)}e_1 \wedge \cdots \wedge DX^{(s,y)}e_{n-i}$ is $n - i$, by our choice of $e_1 \ldots e_n$; and similarly, the rank of

$$(\frac{\partial^2 H}{\partial pp}DP^{(s,y)}e_{n-i+1}) \wedge \cdots \wedge (\frac{\partial^2 H}{\partial pp}DP^{(s,y)}e_n)$$

is i: otherwise, there would be a non–null vector v in the kernel V of $DX^{(s,y)}$ such that $DP^{(s,y)}v = 0$: but this would contradict the lemma 6.1, since then

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

Similarly by the above lemma 6.4, 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 $\frac{\partial^2 H}{\partial pp} DP^{(s,y)} v = DX^{(s,y)} w$, then,

$$(DP^{(s,y)} v) \cdot (\frac{\partial^2 H}{\partial pp} DP^{(s,y)} v) = (DP^{(s,y)} v) \cdot (DX^{(s,y)} w)$$

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 6.4, the RHS is

$$\sum_{i,j,k} \frac{\partial X_h}{\partial \overline{x}_i}(\overline{x}) w_i \frac{\partial P_h}{\partial \overline{x}_k}(\overline{x}) v_k = \sum_{i,j,k} \frac{\partial X_h}{\partial \overline{x}_i}(\overline{x}) v_i \frac{\partial P_h}{\partial \overline{x}_k}(\overline{x}) w_k = 0 \tag{6.9}$$

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

Then the above form (6.8) is non degenerate at (s', y').

This concludes the proof of 4.7.

§6.ii.1 Proof of 6.3

We prove now that the representation of $\frac{\partial P}{\partial \overline{x}}^t \frac{\partial X}{\partial \overline{x}}$ in local coordinates is a symmetric matrix.

Lemma 6.10 *The representation in local coordinates of*

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

is symmetric.

Proof. In this proof, A^t is the transpose of a matrix A; vectors are considered to be column vectors. By prop. [12, 3.1.10.ii]), for any $\nu, \nu' \in T_{(x,p)} T^* M$, we have

$$\Phi_t^* \omega = \omega$$

ie

$$\omega(D\Phi_t(x,p)\nu',\ D\Phi_t(x,p)\nu) = \omega(\nu', \nu)$$

This can be written in local coordinates as

$$\nu' \cdot \frac{\partial \Phi_t}{\partial(x,p)}^t J \frac{\partial \Phi_t}{\partial(x,p)} \nu = \nu' \cdot J\nu \tag{6.10.$\star$}$$

where J was defined in (A.1).

We need now to compose the injection $j : O \to T^* M$, with Φ_t; we write the derivative of j, as $B \doteq \frac{\partial j}{\partial y}$, so that in local coordinates

$$\frac{\partial X}{\partial y} = \frac{\partial X}{\partial x, p} B \qquad \frac{\partial P}{\partial y} = \frac{\partial P}{\partial x, p} B \ ;$$

so we multiply in (6.10.⋆) to obtain

$$B^t \left(\frac{\partial \Phi_t}{\partial(x,p)}\right)^t J \frac{\partial \Phi_t}{\partial(x,p)} B = \begin{pmatrix} \frac{\partial X}{\partial y} \\ \frac{\partial P}{\partial y} \end{pmatrix}^t J \begin{pmatrix} \frac{\partial X}{\partial y} \\ \frac{\partial P}{\partial y} \end{pmatrix} = B^t J B$$

With a direct computation we obtain

$$\begin{pmatrix} \frac{\partial X}{\partial y} \\ \frac{\partial P}{\partial y} \end{pmatrix}^t J \begin{pmatrix} \frac{\partial X}{\partial y} \\ \frac{\partial P}{\partial y} \end{pmatrix} = \begin{pmatrix} \frac{\partial X}{\partial y}^t & \frac{\partial P}{\partial y}^t \end{pmatrix} \begin{pmatrix} -\frac{\partial P}{\partial y} \\ \frac{\partial X}{\partial y} \end{pmatrix} =$$

$$= -\frac{\partial X}{\partial y}^t \frac{\partial P}{\partial y} + \frac{\partial P}{\partial y}^t \frac{\partial X}{\partial y}.$$

To see if $\frac{\partial P}{\partial y}^t \frac{\partial X}{\partial y} = \frac{\partial X}{\partial y}^t \frac{\partial P}{\partial y}$, *we need to prove that* $\omega(\nu',\nu) = 0$ *when* $\nu,\nu' \in TO$. *To compute the term* B^tJB *we choose two curves* $\gamma,\gamma' : \mathbb{R} \to T^*M$ *such that* $\gamma(t),\gamma'(t) \in O$*; in local coordinates we write* $\gamma(t) = (x(t),p(t))$, $\gamma'(t) = (x'(t),p'(t))$, *and we suppose for convenience that* $\dot x', \dot x$ *are constant;* $\omega(\nu',\nu)$ *becomes*

$$\dot\gamma' \cdot J\dot\gamma = \dot p \cdot \dot x' - \dot p' \cdot \dot x \quad ;$$

since $\dot x', \dot x \in T_xK$, *we have that, by the definition of* O *in (3.5),*

$$p \cdot \dot x' = p' \cdot \dot x' = du_0(x') \cdot \dot x' \ ,$$
$$p \cdot \dot x = p' \cdot \dot x = du_0(x) \cdot \dot x \ ;$$

deriving these equalities wrt to t we obtain

$$\dot p \cdot \dot x' = \dot x \cdot \frac{\partial^2 u_0}{\partial x x} \dot x' = \dot p' \cdot \dot x$$

which gives QED. [15]

□

Lemma 6.11

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

Proof. *by simple substitution of (B.4) we obtain*

$$\frac{\partial P}{\partial t}^t (t,y) \frac{\partial X}{\partial y}(t,y) = \sum_j \frac{\partial P_j}{\partial t}(t,y) \frac{\partial X_j}{\partial y}(t,y) = \sum_i \frac{\partial H}{\partial x_i} \frac{\partial X_i}{\partial y}(t,y)$$

---

(15) Note that by direct computation on (6.10.⋆) we get that

$$\frac{\partial P}{\partial x}^t \frac{\partial X}{\partial x} = \frac{\partial X}{\partial x}^t \frac{\partial P}{\partial x}$$
$$\frac{\partial P}{\partial p}^t \frac{\partial X}{\partial p} = \frac{\partial X}{\partial p}^t \frac{\partial P}{\partial p}$$
$$-\frac{\partial P}{\partial x}^t \frac{\partial X}{\partial p} + \frac{\partial X}{\partial x}^t \frac{\partial P}{\partial p} = I$$
$$-\frac{\partial P}{\partial p}^t \frac{\partial X}{\partial x} + \frac{\partial X}{\partial p}^t \frac{\partial P}{\partial x} = -I$$

which is in itself a nice result.

and

$$\frac{\partial P}{\partial y}^t(t,y)\frac{\partial X}{\partial t}(t,y) = \sum_j \frac{\partial P_j}{\partial y}(t,y)\frac{\partial X_j}{\partial t}(t,y) = -\sum_i \frac{\partial P_i}{\partial y}(t,y)\frac{\partial H}{\partial p_i}$$

(where the argument of H is $(X(t,y),P(t,y))$) the thesis can be rewritten as

$$\frac{\partial H}{\partial x}\frac{\partial X}{\partial y}(t,y) + \frac{\partial H}{\partial p}\frac{\partial P}{\partial y}(t,y) = 0$$

the left hand side is equal to

$$\frac{d}{dy}(H(X(t,y),P(t,y)))$$

and $H(X(t,y),P(t,y)) = 0$ by the method of characteristics and by having $y \in O$: we obtain QED. □

Proposition 6.12 (a.k.a. 6.3) *We have that*

$$(DP)^t(DX)$$

is symmetric

Proof. an easy proof comes from the theory of symplectic geometry: consider the Lagrangian submanifold Λ' *of T^*M (introduced in 3.10), 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 [17].*

An alternative, direct proof follows.

$$(DP)^t(DX) = \begin{pmatrix} \frac{\partial P}{\partial t}^t \\ \frac{\partial P}{\partial y}^t \end{pmatrix}\begin{pmatrix} \frac{\partial X}{\partial t} & \frac{\partial X}{\partial y} \end{pmatrix} =$$

$$= \begin{pmatrix} \frac{\partial P}{\partial t}^t(t,y)\frac{\partial X}{\partial t}(t,y) & \frac{\partial P}{\partial t}^t(t,y)\frac{\partial X}{\partial y}(t,y) \\ \frac{\partial P}{\partial y}^t(t,y)\frac{\partial X}{\partial t}(t,y) & \frac{\partial P}{\partial y}^t(t,y)\frac{\partial X}{\partial y}(t,y) \end{pmatrix}$$

then, the term $\frac{\partial P}{\partial t}^t(t,y)\frac{\partial X}{\partial t}(t,y)$ is one dimensional, and we apply the two above lemmas □

A Symplectic manifolds

Following §3.1 in [12], we endow T^*M with a canonical symplectic structure so defined: let $\pi^*_M : T^*M \to M$ be the projection, such that $\pi^*_M(x,p) = x$; we define the 1–form ζ on TT^*M by

$$\zeta(\nu) = p \ \cdot \ T\pi^*_M \nu$$

whenever $\nu \in T_{(x,p)}T^*M$. If we choose local coordinates x_i for M and *canonical coordinates* p_i for T^*M, (such that in local coordinates in $\mathbb{R}^n$, $p \cdot v = \sum_i p_i v_i$), we have the representation $\nu = (\dot{x}, \dot{p}) \in T_{(x,p)}T^*M$ from which

$$\zeta(\nu) = p \cdot \dot{x}$$

or

$$\zeta = \sum_i p_i dx_i \quad .$$

We define the 2–form

$$\omega = -d\zeta \quad ;$$

then (ω, T^*M) is an *exact symplectic manifold* (see [12], prop 3.1.8, or [17], sec. 3.1 pg 88). In the canonical local coordinates of T^*M we can write

$$\omega_{(x,p)} = \sum_i dx_i \wedge dp_i$$

note that if

$$\nu = \begin{pmatrix} \dot{x} \\ \dot{p} \end{pmatrix}, \nu' = \begin{pmatrix} \dot{x}' \\ \dot{p}' \end{pmatrix} \in T_{(x,p)}T^*M$$

we have

$$\omega(\nu, \nu') = \sum_i \dot{p}'_i \dot{x}_i - \sum_i \dot{p}_i \dot{x}'_i \quad ;$$

by defining the matrix

$$J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \tag{A.1}$$

we obtain that in local coordinates we can write

$$\omega(\nu, \nu') = \nu \cdot J\nu' \quad .$$

ω induces a duality $\omega^\#$ between TT^*M and TTM, given by

$$\omega^\#(\nu) \cdot \nu' = \omega(\nu, \nu') \quad \forall \nu' \in T^*M \tag{A.2}$$

B Hamiltonian flow and characteristics

Suppose that we are given a $z \in M$ and a cotangent vector $q \in T^*_z M$; we denote the *hamiltonian flow* to be the function

$$\Phi_t : T^*M \to T^*M \tag{B.1}$$

such that Φ_0 is the identity, and that solves

$$\forall \nu \in TT^*M, \qquad DH(\Phi_t)\cdot\nu = \omega\left(\frac{\partial \Phi_t}{\partial t},\ \nu\right) \tag{B.2}$$

or equivalently

$$\frac{\partial \Phi_t}{\partial t}(x,p) = \omega^{\#}(DH(\Phi_t(x,p)))$$

Using the canonical coordinates, we see that

$$(X(t,z,q),P(t,z,q)) \doteq \Phi_t(z,q) \tag{B.3}$$

is the solution of the *system of characteristics*

$$\begin{cases} \dot X_i(s) = \frac{\partial H}{\partial p_i}(X(s),P(s)) \\ \dot P_i(s) = -\frac{\partial H}{\partial x_i}(X(s),P(s)) \end{cases} \tag{B.4}$$

with initial conditions

$$\begin{cases} X(0) = z \\ P(0) = q \ . \end{cases} \tag{B.5}$$

We complete the system by defining $U(t,z,q)$ as the solution to the ODE

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

that is,

$$U(t,z,q) \doteq \int_0^t P(s,z,q)\cdot \dot X(s,z,q)ds \quad ;$$

U is the integral of the 1–form ζ on the flow lines.

The system (X,P,U) is used classically in the "Cauchy's method of characteristics" to solve (locally) PDEs like (1.1) and (1.2); in particular, we have that

$$H(X(t,x,p),P(t,x,p))$$

is constant in t, (and equal to $H(x,p)$); if $H \in C^{(R+1,\theta)}$ in a neighborhood of $\{H=0\}$, then $\Phi(t,x,p)\in C^{(R,\theta)}$ for all (x,p) with $H(x,p)=0$; moreover, if u is a regular solution, and $p=du(x)$ then $P(t,x,p)=du(X(t,x,p))$; we refer the reader to [10, § VI.7] for details. We will also use this lemma;

Lemma B.6 (a version of Gauss lemma) *If $t=t(\tau)$, $x=x(\tau)$, $p=p(\tau)$ then*

$$\frac{dU(t(\tau),x(\tau),p(\tau))}{d\tau} - P(t(\tau),x(\tau),p(\tau))\cdot\frac{dX(t(\tau),x(\tau),p(\tau))}{d\tau} = t\frac{d}{d\tau}H(x(\tau),p(\tau)) \tag{B.6.$\star$}$$

(whenever the problem is regular enough to define the above derivatives, that is, $\Phi\in C^1$)

Indeed, as in thm 9.1 in §VI.9 in [10], if we derive wrt t the LHS in the formula, we obtain

$$\frac{d}{d\tau}H\big(X(t(\tau),x(\tau),p(\tau)),P(t(\tau),x(\tau),p(\tau))\big) = \frac{d}{d\tau}H(x(\tau),p(\tau)) \tag{B.6.$\star\star$}$$

Eq. (B.6.⋆) can be rephrased as saying that

$$\begin{aligned}\frac{\partial U}{\partial x}(t,x,p) - P(t,x,p)\cdot\frac{\partial X}{\partial x}(t,x,p) &= t\frac{\partial}{\partial x}H(x,p)\\ \frac{\partial U}{\partial p}(t,x,p) - P(t,x,p)\cdot\frac{\partial X}{\partial p}(t,x,p) &= t\frac{\partial}{\partial p}H(x,p)\\ \frac{\partial U}{\partial t}(t,x,p) - P(t,x,p)\cdot\frac{\partial X}{\partial t}(t,x,p) &= 0\end{aligned}$$

Then, let $Z = \{(x,p) \mid H(x,p) = 0\}$; consider a n–dimensional submanifold B of Z, and suppose that X is a diffeomorphism between B and an open subset A of M; if we define a function g on A by

$$g(x) \doteq U(X^{-1}(x))$$

by the above equation (B.6.⋆), we obtain that

$$dg(x) = P(X^{-1}(x))$$

and then

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

If $H(x,p) = |p|^2 - 1$ is based on a Riemannian norm $|p|$ on T^*M (see §3.v) this lemma reduces to the classical Gauss lemma (applied to the cogeodesic flow): indeed, in this case $p \mapsto \frac{\partial}{\partial p}H(x,p)$ is the duality between T^*_xM and T_xM induced by the norm; if we keep t, x fixed wrt τ, and we choose $p = p(\tau)$ s.t. $|p(\tau)| = 1$, since $U(x, p(\tau)) = 2t$, we obtain that $\dot X \cdot \frac{\partial}{\partial \tau}X = 0$.

C 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

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

§C.i 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

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

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 [23], or cor. 3.11 in [22])

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

  (whereas the Hausdorff dimension of B is zero).

By (C.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 $\dim^{\mathcal{H}}(A)$.

§C.ii Yomdin's Sard–type statement

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

[16] indeed, $\mathcal{M}^s(\cdot)$ is finitely subaddittive, but it is not additive: 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

Proposition C.3 *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*

$$\|D^k f(x) - D^k f(y)\| \le \omega(|x-y|) \quad \forall x, y \in A \ ; \qquad (C.3.\star)$$

let for convenience

$$R(r) = R(r, f|A, k) = \frac{1}{k!} r^k \omega(r, f|A, k) \qquad (C.3.\star\star)$$

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

$$|f(x) - p_y(x)| \le R(|x-y|) = \frac{1}{k!}|x-y|^k \omega(|x-y|) \ . \qquad (C.3.\diamond)$$

Regarding $\omega(r)$, we note that: $\omega(r)$ is well defined, and is bounded; $\lim_{r\to 0}\omega(r) = 0$; for all $r > 0$ we have $\omega(r) > 0$, iff for some $r > 0$ we have $\omega(r) > 0$, iff f is not a polynomial of degree k. The function f is of class $C^{k,\theta}$ iff $\omega(r) \le cr^\theta$ (by definition); the least constant c is called $\|D^k f\|_\theta$.

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 [24]) on the approximating polynomials, so that he proves, in section 4 in [23], that

Theorem C.4 *for t small, for $\theta \in [0,1]$, if $f \in C^{k,\theta}$ with $\theta > 0$ then*

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

else if $\theta = 0$,

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

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

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

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 [8].

§C.iii 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 C.5 *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*

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

and

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

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 $\mathcal{H}^\alpha$.

Similarly it can be shown that the class $C^{k,\theta}$ is well defined, indipendently 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; 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 C.6 *let $A \subset E$ be such that $\overline{A}$ is compact and $\overline{A} \subset E$; then*

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

§C.iv Acknowledgement

The author wishes to thank 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 topic, that were at the origin of this appendix section.

Contents

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] M. G. Crandall, H. Ishii, and P. L. Lions. User's guide to viscosity solutions of second order partial differential equations. *Bull. Am. Math. Soc.*, 27/1:1–67, 1992.

[7] Lawrence C. Evans. *Partial Differential Equations*, volume 19 of *Graduate Studies in Mathematics*. A.M.S.

[8] H. Federer. *Geometric measure theory*. Springer Verlag, Berlin, 1969.

[9] W. H. Fleming and H. M. Soner. *Controlled Markov processes and viscosity solutions*. Springer Verlag, Berlin, 1993.

[10] P. Hartman. *Ordinary Differential Equations*. Wiley, New York, 1964.

[11] 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.

[12] Wilhelm Klingenberg. *Riemannian Geometry*. W. de Gruyter, Berlin, 1982.

[13] S. N. Kružhkov. The cauchy problem in the large for certain non-linear first order differential equations. *Soviet Math. Dockl.*, 1:474–475, 1960.

[14] S. N. Kružhkov. The cauchy problem in the large for non-linear equations and for certain quasilinear systems of the first order with several variables. *Soviet Math. Dockl.*, 1:493–496, 1964. also in ≪Dokl. Akad. Nauk **155** (1964), 743–746≫.

[15] P. L. Lions. *Generalized Solutions of Hamilton–Jacobi Equations*. Pitman, Boston, 1982.

[16] C. Mantegazza and A. C. Mennucci. Hamilton–Jacobi equations and distance functions on Riemannian manifolds. *Applied Math. and Optim.*, 2002. DOI 10.1007/s00245-002-0736-4.

[17] Dusa McDuff and Dietmar Salomon. *Introduction to Symplectic Topology*. Oxford Mathematical Monographs. Oxford University Press - Clarendon Press - Oxford, 1995.

[18] A. C. G. Mennucci. Regularity and variationality of solutions to Hamilton-Jacobi equations. part ii: variationality, existence, uniqueness. in preparation, 2003.

[19] 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.

[20] *Advances in System Theory*, Boston, October 1999. Kluwer Academic Publishers.

[21] L. Simon. *Lectures on Geometric Measure Theory*, volume 3 of *Proc. Center for Mathematical Analysis*. Australian National University, Canberra, 1983.

[22] Y. Yomdin. β-spreads of sets in metric spaces and critical values of smooth functions.

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

[24] 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.