ON THE DISTRIBUTIONAL HESSIAN OF THE DISTANCE FUNCTION

CARLO MANTEGAZZA, GIOVANNI MASCELLANI, AND GENNADY URALTSEV

ABSTRACT. We describe the precise structure of the distributional Hessian of the distance function from a point of a Riemannian manifold. In doing this we also discuss some geometrical properties of the cutlocus of a point and we compare some different weak notions of Hessian and Laplacian.

1. INTRODUCTION

Let (M, g) be an n–dimensional, smooth, complete Riemannian manifold, for any point $p \in M$ we define $d_p : M \to \mathbb{R}$ to be the distance function from p.

Such distance functions and their relatives, the Busemann functions, enters in several arguments of differential geometry. It is easy to see that, apart from the obvious singularity at the point p, with some few exceptions such distance function is not smooth in $M \setminus \{p\}$ (for instance, when the manifold M is compact), it is anyway 1–Lipschitz and differentiable with a unit gradient almost everywhere (by Rademacher's theorem).

In this note we are concerned with the precise description of the distributional Hessian of d_p, having in mind the following *Laplacian and Hessian comparison theorems* (see [25], for instance).

Theorem 1.1. *If (M, g) satisfies $\mathrm{Ric} \geq (n-1)K$ then, considering polar coordinates around the points $p \in M$ and P in the simply connected, n–dimensional space S^K of constant curvature $K \in \mathbb{R}$, we have*

$$\Delta d_p(r) \leq \Delta^K d_P^K(r) \,.$$

If the sectional curvature of (M, g) is greater or equal to K, then

$$\mathrm{Hess}\, d_p(r) \leq \mathrm{Hess}^K d_P^K(r) \,.$$

Here $\Delta^K d_P^K(r)$ and $\mathrm{Hess}^K d_P^K(r)$ denote respectively the Laplacian and the Hessian of the distance function $d_P^K(\,\cdot\,) = d^K(P, \,\cdot\,)$ in S^K, at distance r from P.

It is often stated by several authors that these inequalities actually hold on the whole manifold (M, g), in some weak sense, that is, in sense of distributions, of viscosity, of barriers. Such conclusion can simplify and sometimes is actually necessary in global arguments involving this comparison theorem, more in general, one often would like to use (weak or strong) maximum principle for the Laplacian in situations where the functions involved are not smooth, for instance, in the proof of the "splitting" theorem (first proved by Cheeger and Gromoll [9]) by Eschenburg and Heintze [13], but also of Topogonov theorem and the "soul" theorem (Cheeger, Gromoll and Meyer [10, 17]).

To be precise, we give the respective definitions of these notions.

Definition 1.2. Let A be a smooth, symmetric $(0, 2)$–tensor field on a Riemannian manifold (M, g).

- We say that the function $f : M \to \mathbb{R}$ satisfies $\mathrm{Hess}\, f \leq A$ in *distributional sense* if for every smooth vector field V with compact support there holds $\int_M f \, \nabla^2_{ji}(V^i V^j) \, d\mathrm{Vol} \leq \int_M A_{ij} V^i V^j \, d\mathrm{Vol}$.
- For a continuous function $f : M \to \mathbb{R}$, we say that $\mathrm{Hess}\, f \leq A$ at the point $p \in M$ in *barrier sense* if for every $\varepsilon > 0$ there exists a neighborhood U_ε of the point p and a C^2–function $h_\varepsilon : U_\varepsilon \to \mathbb{R}$ such that $h_\varepsilon(p) = f(p)$, $h_\varepsilon \geq f$ in U_ε and $\mathrm{Hess}\, h_\varepsilon(p) \leq A(p) + \varepsilon g(p)$ as $(0, 2)$–tensor fields (such a function h_ε is called an *upper barrier*).

- For a continuous function $f : M \to \mathbb{R}$, we say that $\operatorname{Hess} f \leq A$ at the point $p \in M$ in *viscosity sense* if for every C^2–function h from a neighborhood U of the point p such that $h(p) = f(p)$ and $h \leq f$ in U, we have $\operatorname{Hess} h(p) \leq A(p)$.

The weak notions of the inequality $\Delta f \leq \alpha$, for some smooth function $\alpha : M \to \mathbb{R}$, are defined analogously.

- We say that the function $f : M \to \mathbb{R}$ satisfies $\Delta f \leq \alpha$ in *distributional sense* if for every smooth, nonnegative function $\varphi : M \to \mathbb{R}$ with compact support there holds $\int_M f \, \Delta\varphi \, d\mathrm{Vol} \leq \int_M \alpha\varphi \, d\mathrm{Vol}$.
- For a continuous function $f : M \to \mathbb{R}$, we say that $\Delta f \leq \alpha$ at the point $p \in M$ in *barrier sense* if for every $\varepsilon > 0$ there exists a neighborhood U_ε of the point p and a C^2–function $h_\varepsilon : U_\varepsilon \to \mathbb{R}$ such that $h_\varepsilon(p) = f(p)$, $h_\varepsilon \geq f$ in U_ε and $\Delta h_\varepsilon(p) \leq \alpha(p) + \varepsilon$.
- For a continuous function $f : M \to \mathbb{R}$, we say that $\Delta f \leq \alpha$ at the point $p \in M$ in *viscosity sense* if for every C^2–function h from a neighborhood U of the point p such that $h(p) = f(p)$ and $h \leq f$ in U, we have $\Delta h(p) \leq \alpha(p)$.

In this definition and in the following of this paper we will use the Einstein summation convention on repeated indices. In particular, with the notation $\nabla^2_{ij}(V^i V^j)$ we mean $\nabla^2_{ij}(V \otimes V)^{ij}$; i.e., the function obtained by contracting twice the second covariant derivative of the tensor product $V \otimes V$.

The notion of inequality "in barrier sense" was defined by Calabi [8] (for the Laplacian) back in 1958 (he used the terminology "weak sense" rather than "barrier sense") who also proved the relative global "weak" Laplacian comparison theorem (see also the book of Petersen [25, Section 9.3]).

The notion of viscosity solution (which is connected to the definition of inequality "in viscosity sense", see Appendix A) was introduced by Crandall and Lions [12, Definition 3.2] for partial differential *equations*, the above definition for the Hessian is a generalization to a very special *system* of PDEs.

The distributional notion is useful when integrations (by parts) are involved, the other two concepts when the arguments are based on maximum principle.

It is easy to see, by looking at the definitions, that "barrier sense" implies "viscosity sense", moreover, by the work [19], if $f : M \to \mathbb{R}$ satisfies $\Delta f \leq \alpha$ in viscosity sense it also satisfies $\Delta f \leq \alpha$ as distributions and viceversa. In the Appendix A we will discuss in detail the relations between these definitions.

In the next section we will describe the distributional structure of the Hessian (and hence of the Laplacian) of d_p which will imply the mentioned validity of the above inequalities on the whole manifold.

It is a standard fact that the function d_p is smooth in the set $M \setminus (\{p\} \cup \mathrm{Cut}_p)$, where Cut_p is the *cutlocus* of the point p, which we are now going to define along with stating its basic properties (we keep the books [16] and [26] as general references). It is anyway well known that Cut_p is a closed set of zero (canonical) measure. Hence, in the open set $M \setminus (\{p\} \cup \mathrm{Cut}_p)$ the Hessian and Laplacian of d_p are the usual ones (even seen as distributions or using other weak definitions) and all the analysis is concerned to what happens on Cut_p (the situation at the point p is straightforward as d_p is easily seen to behave as the function $\|x\|$ at the origin of $\mathbb{R}^n$).

We let $U_p = \{v \in T_p M \,|\, g_p(v, v) = 1\}$ to be the set of unit tangent vectors to M at p. Given $v \in U_p$ we consider the geodesic $\gamma_v(t) = \exp_p(tv)$ and we let $\sigma_v \in \mathbb{R}^+$ (possibly equal to $+\infty$) to be the maximal time such that $\gamma_v([0, \sigma_v])$ is minimal between any pair of its points. It is so defined a map $\sigma : U_p \to \mathbb{R}^+ \cup \{+\infty\}$ and the point $\gamma_v(\sigma_v)$ (when $\sigma_v < +\infty$) is called the *cutpoint* of the geodesic γ_v.

Definition 1.3. The set of all cutpoints $\gamma_v(\sigma_v)$ for $v \in U_p$ with $\sigma_v < +\infty$ is called the *cutlocus* of the point $p \in M$.

The reasons why a geodesic ceases to be minimal are explained in the following proposition.

Proposition 1.4. *If for a geodesic $\gamma_v(t)$ from the point $p \in M$ we have $\sigma_v < +\infty$, at least one of the following two (mutually non exclusive) conditions is satisfied:*

(1) *at the cutpoint $q = \gamma_v(\sigma_v v)$ there arrives at least another minimal geodesic from p,*
(2) *the differential $d\exp_p$ is not invertible at the point $\sigma_v v \in T_p M$.*

Conversely, if at least one of these conditions is satisfied the geodesic $\gamma_v(t)$ cannot be minimal on an interval larger that $[0, \sigma_v]$.

It is well known that the subset of points $q \in \mathrm{Cut}_p$ where more than a minimal geodesic from p arrive coincides with Sing, which is the singular set of the distance function d_p in $M \setminus \{p\}$. We also define the set Conj of the points $q = \gamma_v(\sigma_v) \in \mathrm{Cut}_p$ with $d\exp_p$ not invertible at $\sigma_v v \in T_p M$, we call Conj the *locus of optimal conjugate points* (see [16, 26]).

2. The Structure of the Distributional Hessian of the Distance Function

The following properties of the function d_p and of the cutlocus of $p \in M$ are proved in the paper [22], Section 3 (see also the wonderful work [20] for other fine properties, notably the local Lipschitzianity of the function $\sigma : U_p \to \mathbb{R}^+ \cup \{+\infty\}$, in Theorem 1.1).

Given an open set $\Omega \subset \mathbb{R}^n$, we say that a continuous function $u : \Omega \to \mathbb{R}$ is *locally semiconcave* if, for any open convex set $K \subset \Omega$ with compact closure in Ω, the function $u|_K$ is the sum of a C^2 function with a concave function.
A continuous function $u : M \to \mathbb{R}$ is called *locally semiconcave* if, for any local chart $\psi : \mathbb{R}^n \to U \subset M$, the function $u \circ \psi$ is locally semiconcave in $\mathbb{R}^n$ according to the above definition.

Proposition 2.1 (Proposition 3.4 in [22])**.** *The function d_p is locally semiconcave in $M \setminus \{p\}$.*

This fact, which follows by recognizing d_p as a viscosity solution of the *eikonal equation $|\nabla u| = 1$* (see [22]), has some relevant consequences, we need some definitions for the precise statements.

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

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

For any locally Lipschitz function u, the set $\partial^+ u(q)$ is a compact convex set, almost everywhere coinciding with the differential of the function u, by Rademacher's theorem.

Proposition 2.2 (Proposition 2.1 in [1])**.** *Let the function $u : M \to \mathbb{R}$ be semiconcave, then the superdifferential $\partial^+ u$ is not empty at each point, moreover, $\partial^+ v$ is upper semicontinuous, namely*

$$q_k \to q, \quad v_k \to v, \quad v_k \in \partial^+ u(q_k) \implies v \in \partial^+ u(q).$$

In particular, if the differential du exists at every point of M, then $u \in C^1(M)$.

Proposition 2.3 (Remark 3.6 in [1])**.** *The set $Ext(\partial^+ d_p(q)$ of extremal points of the (convex) superdifferential set of d_p at q is in one–to–one correspondence with the family $\mathcal{G}(q)$ of minimal geodesics from p to q. Precisely $\mathcal{G}(q)$ is described by*

$$\mathcal{G}(q) = \left\{ \exp_q(-vt), \ t \in [0, 1] \,\Big|\, \forall v \in Ext(\partial^+ d_p(q) \right\}.$$

We now deal with structure of the cutlocus of $p \in M$. Let $\mathcal{H}^{n-1}$ denote the $(n-1)$–dimensional Hausdorff measure on (M, g) (see [14, 27]).

Definition 2.4. We say that a subset $S \subset M$ is C^r–*rectifiable*, with $r \geq 1$, if it can be covered by a countable family of embedded C^r–submanifolds of dimension $(n-1)$, with the exception of a set of $\mathcal{H}^{n-1}$–zero measure (see [14, 27] for a complete discussion of the notion of rectifiability).

Proposition 2.5 (Theorem 4.10 in [22])**.** *The cutlocus of $p \in M$ is C^∞–rectifiable. Hence, its Hausdorff dimension is at most $n-1$. Moreover, for any compact subset K of M the measure $\mathcal{H}^{n-1}(\mathrm{Cut}_p \cap K)$ is finite (Corollary 1.3 in [20]).*

To explain the following consequence of such rectifiability, we need to introduce briefly the theory of functions with *bounded variation*, see [5, 7, 14, 27] for details. We say that a function $u : \mathbb{R}^n \to \mathbb{R}^m$ is a function with *locally bounded variation*, that is, $u \in BV_{\mathrm{loc}}$, if its distributional derivative Du is a Radon measure. Such notion can be easily extended to maps between manifolds using smooth local charts.
A standard result says that the derivative of a locally semiconcave function stays in BV_{loc}, in view of Proposition 2.1 this implies that the vector field ∇d_p belongs to BV_{loc} in the open set $M \setminus \{p\}$.
Then, we define the subspace of BV_{loc} of functions (or vector fields, as before) with locally *special bounded variation*, called SBV_{loc} (see [2, 3, 4, 5, 7]).
The Radon measure representing the distributional derivative Du of a function $u : \mathbb{R}^n \to \mathbb{R}^m$ with locally bounded variation can be always uniquely separated in three mutually singular measures

$$Du = \widetilde{Du} + \mathrm{J}u + \mathrm{C}u$$

where the first term is the part absolutely continuous with respect to the Lebesgue measure $\mathcal{L}^n$, $\mathrm{J}u$ is a measure concentrated on an $(n-1)$–rectifiable set and $\mathrm{C}u$ (called the *Cantor part*) is a measure which does not charge the subsets of Hausdorff dimension $(n-1)$.
The space SBV_{loc} is defined as the class of functions $u \in BV_{\mathrm{loc}}$ such that $\mathrm{C}u = 0$, that is, the Cantor part of the distributional derivative of u is zero. Again, by means of the local charts, this notion is easily generalized to Riemannian manifolds.

Proposition 2.6 (Corollary 4.13 in [22]). *The ($\mathcal{H}^{n-1}$–almost everywhere defined) measurable unit vector field ∇d_p belongs to the space $SBV_{\mathrm{loc}}(M \setminus \{p\})$ of vector fields with locally special bounded variation.*

The immediate consequence of this proposition is that the $(0, 2)$–tensor field valued distribution $\mathrm{Hess}\, d_p$ is actually a Radon measure with an absolutely continuous part, with respect to the canonical volume measure Vol of (M, g), concentrated in $M \setminus (\{p\} \cup \mathrm{Cut}_p)$ where d_p is a smooth function, hence in this set $\mathrm{Hess}\, d_p$ coincides with the standard Hessian $\widetilde{\mathrm{Hess}}\, d_p$ times the volume measure Vol. When the dimension of M is at least two, the singular part of the measure $\mathrm{Hess}\, d_p$ does not "see" the singular point p, hence, it is concentrated on Cut_p, absolutely continuous with respect to the Hausdorff measure $\mathcal{H}^{n-1}$, restricted to Cut_p.
By the properties of rectifiable sets, at $\mathcal{H}^{n-1}$–almost every point $q \in \mathrm{Cut}_p$ there exists an $(n-1)$–dimensional *approximate tangent space* $\mathrm{ap}T_q\mathrm{Cut}_p \subset T_qM$ (in the sense of geometric measure theory, see [14, 27] for details). To give an example, we say that an hyperplane $T \subset \mathbb{R}^n$ is the approximate tangent space to an $(n-1)$–dimensional rectifiable set $K \in \mathbb{R}^n$ at the point x_0, if $\mathcal{H}^{n-1} \llcorner T$ is the limit in the sense of Radon measures, as $\rho \to +\infty$, of the blow–up measures $\mathcal{H}^{n-1} \llcorner \rho(K - x_0)$ around the point x_0. With some technicalities, this notion can be extended also to Riemannian manifolds.
Moreover, see [5], at $\mathcal{H}^{n-1}$–almost every point $q \in \mathrm{Cut}_p$, the field ∇d_p has two distinct *approximate* (in the sense of Lebesgue differentiation theorem) limits "on the two sides" of $\mathrm{ap}T_q\mathrm{Cut}_p \subset T_qM$, given by ∇d_p^+ and ∇d_p^-.
We want to see now that at $\mathcal{H}^{n-1}$–almost every point of Cut_p there arrive exactly two distinct geodesics and no more. We underline that a stronger form of this theorem was already obtained in [6] and [15], concluding that the set $\mathrm{Cut}_p \setminus U$ (where U is like in the following statement) actually has Hausdorff dimension not greater that $n-2$.

Theorem 2.7. *There is an open set $U \subset M$ such that $\mathcal{H}^{n-1}(\mathrm{Cut}_p \setminus U) = 0$ and*

- *the subset $\mathrm{Cut}_p \cap U$ does not contain conjugate points, hence the set of optimal conjugate points has $\mathcal{H}^{n-1}$–zero measure;*
- *at every point of $\mathrm{Cut}_p \cap U$ there arrive exactly two minimal geodesics from $p \in M$;*
- *locally around every point of $\mathrm{Cut}_p \cap U$ the set Cut_p is a smooth $(n-1)$–dimensional hypersurface, hence, $\mathrm{ap}T_q\mathrm{Cut}_p$ is actually the classical tangent space to a hypersurface.*

Proof. First we show that the set of optimal conjugate points Conj is a closed subset of $\mathcal{H}^{n-1}$–zero measure, then we will see that the points of $\mathrm{Sing} \setminus \mathrm{Conj}$ where there arrive more than two geodesics is also a closed subset of $\mathcal{H}^{n-1}$–zero measure. The third point then follows by the analysis in the proof of Proposition 4.7 in [22].

Recalling that $U_p = \{v \in T_p M \,|\, g_p(v,v) = 1\}$ is the set of unit tangent vectors to M at p, we define the function $c : U_p \to \mathbb{R}^+ \cup \{+\infty\}$ such that the point $\gamma_v(c_v)$ is the first conjugate point (if it exists) along the geodesic γ_v, that is, the differential $d\exp_p$ is not invertible at the point $c_v v \in T_p M$. By Lemma 4.11 and the proof of Proposition 4.9 in [22], in the open subset $V \subset U_p$ where the rank of the differential of the map $F : U_p \to M$, defined as $F(v) = \exp_p(c_v v)$ is $n-1$, the map $c : U_p \to \mathbb{R}^+ \cup \{+\infty\}$ is smooth hence $F(V)$ is locally a smooth hypersurface. As, by Sard's theorem, the image of $U_p \setminus V$ is a closed set of $\mathcal{H}^{n-1}$–zero measure, we only have to deal with the images $F(v)$ of the unit vectors $v \in V$ with $c_v = \sigma_v$ (see at the end of the introduction), that is, with $F(V) \cap \mathrm{Cut}_p$, which is a closed set.

We then consider the set $D \subset (F(V) \cap \mathrm{Cut}_p)$ of the points q where $\mathrm{ap}T_q\mathrm{Cut}_p$ exists and the *density* of the rectifiable set $F(V) \cap \mathrm{Cut}_p$ in the cutlocus of the point p, with respect to the Hausdorff measure $\mathcal{H}^{n-1}$, is one (see [14, 27]). It is well known that D and $F(V) \cap \mathrm{Cut}_p$ only differ by a set of $\mathcal{H}^{n-1}$–zero measure. If $F(v) = q \in D$ then, $c_v = \sigma_v$ and, by the above density property, the hypersurface $F(V)$ is "tangent" to Cut_p at the point q, that is, $T_q F(V) = \mathrm{ap}T_q\mathrm{Cut}_p$.

We claim now that the minimal geodesic γ_v is tangent to the hypersurface $F(V)$, hence to the cutlocus, at the point q. Indeed, as $d\exp_p$ is not invertible at $c_v v \in T_p M$, by Gauss lemma there exists a vector $w \in T_v U_p$ such that $d\exp_p[c_v v](w) = 0$, hence

$$dF_v(w) = (dc[v](w))\dot{\gamma}_v(c_v) + d\exp_p[c_v v](c_v w) = (dc[v](w))\dot{\gamma}_v(c_v)\,,$$

thus, $\dot{\gamma}_v(c_v)$ belongs to the tangent space $dF(T_v U_p)$ to the hypersurface $F(V)$ at the point q, which coincides with $\mathrm{ap}T_q\mathrm{Cut}_p$, as we claimed.

By the properties of SBV functions described before, at $\mathcal{H}^{n-1}$–almost every point $q \in D$, the blow–up of the function d_p is a "roof", that is, there arrive exactly two minimal geodesics both intersecting transversally the cutlocus at q (the vectors ∇d_p^+ and ∇d_p^- do not belong to $\mathrm{ap}T_q M$), hence the above minimal geodesic γ_v cannot coincide with any of these two.

We then conclude that $\mathcal{H}^{n-1}(D) = 0$ and the same for the set Conj.

Suppose now that $q \in \mathrm{Cut}_p \setminus \mathrm{Conj} \subset \mathrm{Sing}$, by the analysis in the proof of Proposition 4.7 in [22] (and Lemma 4.8), at the point q there arrive a *finite* number $m \geq 2$ of distinct minimal geodesics and when $m > 2$ the cutlocus of p is given by the union of at least m smooth hypersurfaces with Lipschitz boundary passing at the point q, in particular the above blow–up at q cannot be a single hyperplane $\mathrm{ap}T_q\mathrm{Cut}_p$. By the above discussion, such points with $m > 2$ are then of $\mathcal{H}^{n-1}$–zero measure, moreover, by Propositions 2.2 and 2.3 the set of points in $\mathrm{Cut}_p \setminus \mathrm{Conj}$ with only two minimal geodesics is open and we are done. $\qquad\square$

Remark 2.8. In the special two–dimensional and analytic case, it can be said something more, that is, the number of optimal conjugate points is locally finite and the cutlocus is a locally finite graph with smooth edges, see the classical papers by Myers [23, 24]. We conjecture that, in general, the set of optimal conjugate points is an $(n-2)$–dimensional rectifiable set.

By the third point if this theorem, in the open set U the two side limits ∇d_p^+ and ∇d_p^- of the gradient field ∇d_p are actually smooth and classical limits, moreover it is locally defined a smooth unit normal vector $\nu_q \in T_q M$ orthogonal to $T_q \mathrm{Cut}_p$, with the convention that $g_q(\nu_q, v)$ is positive for every vector $v \in T_q M$ belonging to the halfspace corresponding to the side associated to ∇d_p^+. Hence, since $\mathcal{H}^{n-1}(\mathrm{Cut}_p \setminus U) = 0$, we have a precise description of the singular "jump" part as follows,

$$\mathrm{J}\nabla d_p = -\Big((\nabla d_p^+ - \nabla d_p^-) \otimes \nu \Big) \mathcal{H}^{n-1} \llcorner \mathrm{Cut}_p$$

and, noticing that the "jump" of the gradient of d_p in U must be orthogonal to the tangent space $T_q \mathrm{Cut}_p$, thus parallel to the unit normal vector $\nu_q \in T_q M$, hence we conclude

$$\mathrm{J}\nabla d_p = -(\nu \otimes \nu) |\nabla d_p^+ - \nabla d_p^-|_g \, \mathcal{H}^{n-1} \llcorner \mathrm{Cut}_p\,.$$

Notice that the singular part of the distributional Hessian of d_p is a rank one symmetric $(0,2)$–tensor field.

Remark 2.9. This description of the "jump" part of the singular measure is actually a direct consequence of the structure theorem of BV functions (see [5]), even without knowing, by Theorem 2.7, that the cutlocus is $\mathcal{H}^{n-1}$–almost everywhere smooth.

Theorem 2.10. *If $n \geq 2$, the distributional Hessian of the distance from a point $p \in M$ is given by the Radon measure*

$$\mathrm{Hess}\, d_p = \widetilde{\mathrm{Hess}}\, d_p \, \mathrm{Vol} - (\nu \otimes \nu)\, |\nabla d_p^+ - \nabla d_p^-|_g \, \mathcal{H}^{n-1} \llcorner \mathrm{Cut}_p \,,$$

where $\widetilde{\mathrm{Hess}}\, d_p$ is the standard Hessian of d_p, where it exists ($\mathcal{H}^{n-1}$–almost everywhere on M), and ∇d_p^+, ∇d_p^-, ν are defined above.

Corollary 2.11. *If $n \geq 2$, the distributional Laplacian of d_p is the Radon measure*

$$\Delta d_p = \widetilde{\Delta} d_p \, \mathrm{Vol} - |\nabla d_p^+ - \nabla d_p^-|_g \, \mathcal{H}^{n-1} \llcorner \mathrm{Cut}_p \,,$$

where $\widetilde{\Delta} d_p$ is the standard Laplacian of d_p, where it exists.

Corollary 2.12. *There hold*

$$\Delta d_p \leq \widetilde{\Delta} d_p \, \mathrm{Vol}$$

and

$$\mathrm{Hess}\, d_p \leq \widetilde{\mathrm{Hess}}\, d_p \, \mathrm{Vol} \,,$$

as $(0,2)$–tensor fields.
As a consequence, the Hessian and Laplacian inequalities in Theorem 1.1 hold in the sense of distributions. Moreover, we have

$$\Delta d_p \geq \widetilde{\Delta} d_p \, \mathrm{Vol} - 2\mathcal{H}^{n-1} \llcorner \mathrm{Cut}_p \,.$$

and

$$\mathrm{Hess}\, d_p \geq \widetilde{\mathrm{Hess}}\, d_p \, \mathrm{Vol} - 2(\nu \otimes \nu)\, \mathcal{H}^{n-1} \llcorner \mathrm{Cut}_p \geq \widetilde{\mathrm{Hess}}\, d_p \, \mathrm{Vol} - 2g \, \mathcal{H}^{n-1} \llcorner \mathrm{Cut}_p \,,$$

as $(0,2)$–tensor fields.

Remark 2.13. By their definition, it is easy to see that the same inequalities hold also for the Busemann functions, see for instance [25, Subsection 9.3.4] (in Section 9.3 of the same book it is shown that the above Laplacian comparison holds on the whole M in barrier sense while the analogous result for the Hessian can be found in Section 11.2). We underline here that Propositions 2.1, 2.2 and 2.3 about the semiconcavity and the structure of the superdifferential of the distance function d_p can also be used to show that the above inequalities hold in barrier/viscosity sense.

Remark 2.14. Several of the conclusions of this paper holds also for the distance function from a closed *subset* of M with boundary of class C^3 at least, see [22] for details.

APPENDIX A. WEAK DEFINITIONS OF SUB/SUPERSOLUTIONS OF PDEs

Let (M, g) be a smooth, complete, Riemannian manifold and let A be a smooth $(0, 2)$ tensor field.

If $f : M \to \mathbb{R}$ satisfies $\mathrm{Hess}\, f \leq A$ at the point $p \in M$ in barrier sense, for every $\varepsilon > 0$ there exists a neighborhood U_ε of the point p and a C^2–function $h_\varepsilon : U_\varepsilon \to \mathbb{R}$ such that $h_\varepsilon(p) = f(p)$, $h_\varepsilon \geq f$ in U_ε and $\mathrm{Hess}\, h_\varepsilon(p) \leq A(p) + \varepsilon g(p)$, hence, every C^2–function h from a neighborhood U of the point p such that $h(p) = f(p)$ and $h \leq f$ in U satisfies $h(p) = h_\varepsilon(p)$ and $h \leq h_\varepsilon$ in $U \cap U_\varepsilon$. It is then easy to see that it must be $\mathrm{Hess}\, h(p) \leq \mathrm{Hess}\, h_\varepsilon(p) \leq A(p) + \varepsilon g(p)$, for every $\varepsilon > 0$, hence $\mathrm{Hess}\, h(p) \leq A(p)$. This shows that $\mathrm{Hess}\, f \leq A$ at the point $p \in M$ also in *viscosity sense*.
The converse is not true, indeed, it is straightforward to check that the function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^2 \sin(1/x)$ when $x \neq 0$ and $f(0) = 0$ satisfies $f''(0) \leq 0$ in viscosity sense but not in barrier sense.
The same argument clearly also applies to the two definitions of $\Delta f \leq \alpha$, for a smooth function $\alpha : M \to \mathbb{R}$.

We see now that instead the definitions of viscosity and distributional sense coincide.

Proposition A.1. *If $f : M \to \mathbb{R}$ satisfies $\mathrm{Hess}\, f \leq A$ in viscosity sense, it also satisfies $\mathrm{Hess}\, f \leq A$ in distributional sense and viceversa. The same holds for $\Delta f \leq \alpha$.*

In order to show the proposition, we recall the definitions of *viscosity* (sub/super) solution to a second order PDE. Take a continuous map $F : \Omega \times \mathbb{R} \times \mathbb{R}^n \times S^n \to \mathbb{R}$, where Ω is an open subset of $\mathbb{R}^n$ and S^n denotes the space of real $n \times n$ symmetric matrices; also suppose that F satisfies the *monotonicity condition*

$$X \geq Y \qquad \Longrightarrow \qquad F(x, r, p, X) \leq F(x, r, p, Y) \,,$$

for every $(x, r, p) \in \Omega \times \mathbb{R} \times \mathbb{R}^n$, where $X \geq Y$ means that the difference matrix $X - Y$ is nonnegative definite. We consider then the second order PDE given by $F(x, f, \nabla f, \nabla^2 f) = 0$.

A continuous function $f : \Omega \to \mathbb{R}$ is said a *viscosity subsolution* of the above PDE if for every point $x \in \Omega$ and $\varphi \in C^2(\Omega)$ such that $f(x) - \varphi(x) = \sup_\Omega (f - \varphi)$, there holds $F(x, \varphi, \nabla \varphi, \nabla^2 \varphi) \leq 0$

(see [11, 19]). Analogously, $f \in C^0(\Omega)$ is a *viscosity supersolution* if for every point $x \in \Omega$ and $\varphi \in C^2(\Omega)$ such that $f(x) - \varphi(x) = \inf_\Omega (f - \varphi)$, there holds $F(x, \varphi, \nabla \varphi, \nabla^2 \varphi) \geq 0$. If $f \in C^0(\Omega)$ is both a viscosity subsolution and supersolution, it is then a *viscosity solution* of $F(x, f, \nabla f, \nabla^2 f) = 0$ in Ω.

It is easy to see that the functions $f \in C^0(\Omega)$ such that $\Delta f \leq \alpha$ in *viscosity sense* at any point of Ω, as in Definition 1.2, coincide with the viscosity supersolutions of the equation $-\Delta f + \alpha = 0$ at the same point (here the function F is given by $F(x, r, p, X) = -\operatorname{trace} X + \alpha(x)$).

In the case of a Riemannian manifold (M, g), one works in local charts and the operators we are interested in become

$$\operatorname{Hess}_{ij}^M f(x) = \frac{\partial^2 f(x)}{\partial x^i \partial x^j} - \Gamma_{ij}^k(x) \frac{\partial f}{\partial x^k}$$
$$\Delta^M f(x) = g^{ij}(x) \operatorname{Hess}_{ij}^M f(x) \,,$$

where Γ_{ij}^k are the Christoffel symbols.

Analogously to the case of $\mathbb{R}^n$, taking $F(x, r, p, X) = -g^{ij}(x) X_{ij} + g^{ij}(x) \Gamma_{ij}^k(x) p_k + \alpha(x)$ (which is a smooth function independent of the variable r), we see that, according to Definition 1.2, f satisfies $\Delta^M f \leq \alpha$ in *viscosity sense* at any point of M if and only if it is a viscosity supersolution of the equation $F(x, f, \nabla f, \nabla^2 f) = 0$ at the same point.

Getting back to $\mathbb{R}^n$, given a linear, degenerate elliptic operator L with smooth coefficients, that is, defined by

$$Lf(x) = -a^{ij}(x) \nabla_{ij}^2 f(x) + b^k(x) \nabla_k f(x) + c(x) f(x) \,,$$

and a smooth function $\alpha : \Omega \to \mathbb{R}$, we say that $f \in C^0(\Omega)$ is a distributional supersolution of the equation $Lf + \alpha = 0$ when

$$\int_\Omega \Big(f L^* \varphi + \alpha \varphi \Big) dx \geq 0$$

for every nonnegative, smooth function $\varphi \in C_c^\infty(\Omega)$. Here L^* is the formal adjoint operator of L:

$$L^* \varphi(x) = -\nabla_{ji}^2 (a^{ij} \varphi)(x) - \nabla_k (b^k \varphi)(x) + c(x) \varphi(x).$$

Under the hypothesis that the matrix of coefficients (a_{ij}) (which is nonnegative definite) has a "square root" matrix belonging to $C^1(\Omega, S^n)$, Ishii showed in paper [19] the equivalence of the class of continuous viscosity subsolutions and the class of continuous distributional subsolutions of the equation $Lf + \alpha = 0$. More precisely, he proved the following two theorems (see also [21]).

Theorem A.2 (Theorem 1 in [19]). *If $f \in C^0(\Omega)$ is a viscosity subsolution of the equation $Lf + \alpha = 0$, then then it is a distribution subsolution of the same equation.*

Theorem A.3 (Theorem 2 in [19]). *Assume that the "square root" of the matrix of coefficients (a_{ij}) belongs to $C^1(\Omega)$. If $f \in C^0(\Omega)$ is a distributional subsolution of the equation $Lf + \alpha = 0$, then then it is a viscosity subsolution of the same equation.*

As the PDE is linear, a function $f \in C^0(\Omega)$ is a viscosity (distributional) supersolution of the equation $Lf + \alpha = 0$ if and only if the function $-f$ is a viscosity (distributional) subsolution of $L(-f) - \alpha = 0$, in the above theorems every occurrence of the term "subsolution" can replaced with "supersolution" (and actually also with "solution").

For simplicity, we will work in a single coordinate chart of M mapping onto $\Omega \subseteq \mathbb{R}^n$, while the general situation can be dealt with by means of standard partition of unity arguments.

Consider $f \in C^0(M)$ which is a viscosity supersolution of $-\Delta^M f + \alpha = 0$. It is a straightforward computation to check that this happens if and only if f is a viscosity supersolution of $-\sqrt{g}\Delta^M f + \alpha\sqrt{g} = 0$, where $\sqrt{g} = \sqrt{\det g_{ij}}$ is the density of Riemannian volume of (M, g), and viceversa. Moreover, notice that setting $L = -\sqrt{g}\Delta^M$ we have that $L^* = L$, that is, L is a self–adjoint operator; it also satisfies the hypotheses of Ishii's theorems, being the matrices g_{ij} and g^{ij} smooth and positive definite in Ω (see [18, Chapter 6], in particular Example 6.2.14, for instance).

Then, in local coordinates, Ishii's theorems guarantee that f is a distributional supersolution of the same equation, that is, f satisfies, for each $\varphi \in C_c^\infty(\Omega)$,

$$\int_\Omega f L^* \varphi \, dx \geq -\int_\Omega \alpha\sqrt{g}\varphi \, dx \, ,$$

hence,

$$\int_M -f\Delta^M \varphi \, d\mathrm{Vol} = \int_\Omega -f\sqrt{g}\Delta^M \varphi \, dx \geq -\int_\Omega \alpha\sqrt{g}\varphi \, dx = -\int_M \alpha\varphi \, d\mathrm{Vol} \, .$$

This shows that then f satisfies $\Delta^M f \leq \alpha$ in distributional sense, as in Definition 1.2.

Following these steps in reverse order, one gets the converse. Hence, the notions of $\Delta^M \leq \alpha$ in viscosity and distributional sense coincide.

Now we turn our attention to the Hessian inequality; it is not covered directly by Ishii's theorems, which are peculiar to PDEs and do not deal with *systems* (like the general theory of viscosity solutions). For simplicity, we discuss the case of an open set $\Omega \subset \mathbb{R}^n$ (with its canonical flat metric), since all the arguments can be extended to any Riemannian manifold (M, g) by localization and introduction of the first–order correction given by Christoffel symbols, as above.

The idea is to transform the matrix inequality Hess $f \leq A$ into a family of scalar inequalities; indeed, if everything is smooth, such inequality is satisfied if and only if for every compactly–supported, smooth vector field W we have $W^i W^j \mathrm{Hess}_{ij} f \leq A_{ij} W^i W^j$. The only price to pay is that we lose the constant coefficients of the Hessian, hence making the linear operator L^W, acting on $f \in C^2(\Omega)$ as $L^W f = -W^i W^j \mathrm{Hess}_{ij} f$, only *degenerate* elliptic. Notice that Ishii's condition in Theorem A.3 is satisfied for every smooth vector field W such that $\|W\| \in C_c^1(\Omega)$, but not by any arbitrary smooth vector field. This has the collateral effect of making the proof of the Hessian case in Proposition A.1 slightly asymmetric.

Lemma A.4. *Let $f \in C^0(\Omega)$. If for every compactly–supported, smooth vector field W with $\|W\| \in C_c^1(\Omega)$, we have that f is a viscosity supersolution of the equation $-W^i W^j \mathrm{Hess}_{ij} f + A_{ij} W^i W^j = 0$, then the function f satisfies Hess $f \leq A$ in viscosity sense in the whole Ω.*

Viceversa, if $f \in C^0(\Omega)$ satisfies Hess $f \leq A$ in viscosity sense in Ω, then f is a viscosity supersolution of the equation $-V^i V^j \mathrm{Hess}_{ij} f + A_{ij} V^i V^j = 0$ for every compactly–supported, smooth vector field V.

Proof. Let us take a point $x \in \Omega$ and a C^2–function h in a neighborhood U of the point x such that $h(x) = f(x)$ and $h \leq f$. Choosing a unit vector W_x and a smooth, nonnegative function φ, which is 1 at x and zero outside a small ball inside U, we consider the smooth vector field $W(y) = W_x \varphi^2(y)$, for every $y \in \Omega$, which clearly satisfies $\|W\| = \varphi \in C_c^1(\Omega)$. By the hypothesis of the first statement, the function f is then a viscosity supersolution of the equation $-W^i W^j \mathrm{Hess}_{ij} f + A_{ij} W^i W^j = 0$ which implies that $-W_x^i W_x^j \mathrm{Hess}_{ij} h(x) + A_{ij}(x)W_x^i W_x^j \geq 0$. Since this holds for every point $x \in \Omega$ and unit vector W_x, we conclude that Hess $h(x) \leq A(x)$ as $(0, 2)$–tensor fields, hence Hess $f \leq A$ in viscosity sense in Ω.

The argument to show the second statement is analogous: given a compactly–supported, smooth vector field V, a point $x \in \Omega$ and a function h as above, the hypothesis implies that $-V_x^i V_x^j \mathrm{Hess}_{ij} h(x) + A_{ij}(x)V_x^i V_x^j \geq 0$, hence the thesis. $\square$

Suppose now that $f \in C^0(\Omega)$ satisfies Hess $f \leq A$ in viscosity sense on the whole Ω; hence, by this lemma, for every compactly–supported, smooth vector field V, the function f is a viscosity supersolution of the equation $-V^i V^j \mathrm{Hess}_{ij} f + A_{ij} V^i V^j = 0$. By Theorem A.2 and the subsequent

discussion, it is then a distributional supersolution of the same equation, that is,

$$\int_{\Omega} \left[-f \nabla^2_{ji}(V^i V^j \varphi) + A_{ij} V^i V^j \varphi \right] dx \geq 0$$

for every nonnegative, smooth function $\varphi \in C^{\infty}_c(\Omega)$.

Considering a nonnegative, smooth function $\varphi \in C^{\infty}_c(\Omega)$ such that it is one on the support of the vector field V we conclude

$$\int_{\Omega} f \nabla^2_{ji}(V^i V^j) \, dx \leq \int_{\Omega} A_{ij} V^i V^j \, dx \,,$$

which means that Hess $f \leq A$ in distributional sense.

Conversely, if $f \in C^0(\Omega)$ satisfies Hess $f \leq A$ in distributional sense, then for every compactly–supported, smooth vector field W with $\|W\| \in C^1_c(\Omega)$ and smooth, nonnegative function $\varphi \in C^{\infty}_c(\Omega)$, we define the smooth, nonnegative functions $\varphi_n = \varphi + \psi/n$, where ψ is a smooth, nonnegative and compactly–supported function such that $\psi \equiv 1$ on the support of W. It follows that the vector field $V = W\sqrt{\varphi_n}$ is smooth, hence, applying the definition of Hess $f \leq A$ in distributional sense, we get

$$\int_{\Omega} \left[-f \nabla^2_{ji}(W^i W^j \varphi_n) + A_{ij} W^i W^j \varphi_n \right] dx \geq 0 \,.$$

As $\varphi_n \to \varphi$ in $C^{\infty}_c(\Omega)$ and f is continuous, we can pass to the limit in $n \to \infty$ and conclude that

$$\int_{\Omega} \left[-f \nabla^2_{ji}(W^i W^j \varphi) + A_{ij} W^i W^j \varphi \right] dx \geq 0 \,,$$

for every nonnegative, smooth function $\varphi \in C^{\infty}_c(\Omega)$ and every compactly–supported, smooth vector field W with $\|W\| \in C^1_c(\Omega)$. That is, for any vector field W as above, we have that f is a distributional supersolution of the equation $-W^i W^j \, \mathrm{Hess}_{ij} f + A_{ij} W^i W^j = 0$.

By Theorem A.3 and the subsequent discussion, it is then a viscosity supersolution of the same equation and, by Lemma A.4, we conclude that the function f satisfies Hess $f \leq A$ in viscosity sense.

Summarizing, we have the following sharp relations among the weak notions of the partial differential inequalities Hess $f \leq A$ and $\Delta f \leq \alpha$,

$$\text{barrier sense} \quad \implies \quad \text{viscosity sense} \quad \Longleftrightarrow \quad \text{distributional sense.}$$

REFERENCES

1. G. Alberti, L. Ambrosio, and P. Cannarsa, *On the singularities of convex functions*, Manuscripta Math. **76** (1992), 421–435.
2. L. Ambrosio, *A compactness theorem for a new class of functions of bounded variation*, Boll. Un. Mat. Ital. **3-B** (1989), 857–881.
3. _____, *Variational problems in SBV*, Acta Applicandae Mathematicae **17** (1989), 1–40.
4. _____, *Existence theory for a new class of variational problems*, Arch. Rat. Mech. Anal. **111** (1990), 291–322.
5. L. Ambrosio, N. Fusco, and D. Pallara, *Functions of bounded variation and free discontinuity problems*, Oxford University Press, 2000.
6. P. A. Ardoy and L. Guijarro, *Cut and singular loci up to codimension 3*, Ann. Inst. Fourier (Grenoble) **61** (2011), no. 4, 1655–1681 (2012).
7. A. Braides, *Approximation of free–discontinuity problems*, Springer–Verlag, Berlin, 1998.
8. E. Calabi, *An extension of E. Hopf's maximum principle with an application to Riemannian geometry*, Duke Math. J. **25** (1958), 45–56.
9. J. Cheeger and D. Gromoll, *The splitting theorem for manifolds of nonnegative Ricci curvature*, J. Diff. Geom. **6** (1971/72), 119–128.
10. _____, *On the structure of complete manifolds of nonnegative curvature*, Ann. of Math. (2) **96** (1972), 413–443.
11. M. G. Crandall, H. Ishii, and P.-L. Lions, *User's guide to viscosity solutions of second order partial differential equations*, Bull. Amer. Math. Soc. **27/1** (1992), 1–67.
12. M. G. Crandall and P.-L. Lions, *Viscosity solutions of Hamilton–Jacobi equations*, Trans. Amer. Math. Soc. **277** (1983), 1–43.
13. J. Eschenburg and E. Heintze, *An elementary proof of the Cheeger–Gromoll splitting theorem*, Ann. Global Anal. Geom. **2** (1984), no. 2, 141–151.
14. H. Federer, *Geometric measure theory*, Springer–Verlag, 1969.
15. A. Figalli, L. Rifford, and C. Villani, *Necessary and sufficient conditions for continuity of optimal transport maps on Riemannian manifolds*, Tohoku Math. J. **63** (2011), no. 4, 855–876.
16. S. Gallot, D. Hulin, and J. Lafontaine, *Riemannian geometry*, Springer–Verlag, 1990.
17. D. Gromoll and W. Meyer, *On complete open manifolds of positive curvature*, Ann. of Math. (2) **90** (1969), 75–90.

18. R. A. Horn and C. R. Johnson, *Topics in matrix analysis*, Cambridge University Press, Cambridge, 1994.
19. H. Ishii, *On the equivalence of two notions of weak solutions, viscosity solutions and distribution solutions*, Funkcial. Ekvac. **38** (1995), no. 1, 101–120.
20. Y. Li and L. Nirenberg, *The distance function to the boundary, Finsler geometry, and the singular set of viscosity solutions of some Hamilton–Jacobi equations*, Comm. Pure Appl. Math. **58** (2005), no. 1, 85–146.
21. P.-L. Lions, *Optimal control of diffusion processes and Hamilton–Jacobi–Bellman equations. II. Viscosity solutions and uniqueness*, Comm. Partial Differential Equations **8** (1983), no. 11, 1229–1276.
22. C. Mantegazza and A. C. Mennucci, *Hamilton–Jacobi equations and distance functions on Riemannian manifolds*, Appl. Math. Opt. **47** (2003), no. 1, 1–25.
23. S. B. Myers, *Connections between differential geometry and topology*, Duke Math. J. **1** (1935), 376–391.
24. _____, *Connections between differential geometry and topology II*, Duke Math. J. **2** (1936), 95–102.
25. P. Petersen, *Riemannian geometry*, Springer–Verlag, 1998.
26. T. Sakai, *Riemannian geometry*, AMS, 1996.
27. L. Simon, *Lectures on geometric measure theory*, Proc. Center Math. Anal., vol. 3, Australian National University, Canberra, 1983.

(Carlo Mantegazza) Scuola Normale Superiore di Pisa, Piazza dei Cavalieri 7, Pisa, Italy, 56126
E-mail address, C. Mantegazza: `c.mantegazza@sns.it`

(Giovanni Mascellani) Scuola Normale Superiore di Pisa, Piazza dei Cavalieri 7, Pisa, Italy, 56126
E-mail address, G. Mascellani: `giovanni.mascellani@sns.it`

(Gennady Uraltsev) Scuola Normale Superiore di Pisa, Piazza dei Cavalieri 7, Pisa, Italy, 56126
E-mail address, G. Uraltsev: `gennady.uraltsev@sns.it`