ON THE DISTRIBUTIONAL HESSIAN
OF THE DISTANCE FUNCTION

CARLO MANTEGAZZA, GIOVANNI MASCELLANI AND GENNADY URALTSEV
ON THE DISTRIBUTIONAL HESSIAN OF THE DISTANCE FUNCTION

CARLO MANTEGAZZA, GIOVANNI MASCELLANI AND GENNADY URALTSEV

We describe the precise structure of the distributional Hessian of the distance function from a point of a Riemannian manifold. At the same time we discuss some geometrical properties of the cut locus of a point, and compare some different weak notions of the 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, come into several arguments in differential geometry. With few exceptions they are not smooth in \(M \setminus \{p\}\) (and are obviously singular at \(p\)), but it is easy to see that they are 1-Lipschitz and so (by Rademacher’s theorem) differentiable almost everywhere, with unit gradient.

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 [Petersen 1998], for instance):

**Theorem 1.1.** If \((M, g)\) satisfies \(\text{Ric} \geq (n - 1)K\) 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

\[
\Delta d_p(r) \leq \Delta^K d^K_p(r).
\]

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

\[
\text{Hess} \, d_p(r) \leq \text{Hess}^K \, d^K_p(r).
\]

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

It is often stated that these inequalities actually hold on the whole manifold \((M, g)\) in some weak sense, say in the sense of distributions, or viscosity, or barriers. Such

**MSC2010:** primary 58E10; secondary 35D99.

**Keywords:** distance function, cut locus.
results can simplify, and sometimes are necessary for, global arguments involving this comparison theorem. More generally, one often would like to use the (weak or strong) maximum principle for the Laplacian in situations where the functions involved are not smooth, for instance in Eschenburg and Heintze’s proof [1984] of the splitting theorem (first proved in [Cheeger and Gromoll 1971]), or proofs of the Toponogov theorem and the soul theorem [Cheeger and Gromoll 1972; Gromoll and Meyer 1969].

To be precise, we give 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 a function $f : M \to \mathbb{R}$ satisfies $\text{Hess} f \leq A$ in the *distributional sense* if for every smooth vector field $V$ with compact support we have $\int_M f \nabla^2_{ij}(V^i V^j) d\text{Vol} \leq \int_M A_{ij} V^i V^j d\text{Vol}$.

- For a continuous function $f : M \to \mathbb{R}$, we say that $\text{Hess} f \leq A$ at the point $p \in M$ in the *barrier sense* if for every $\varepsilon > 0$ there exists a neighborhood $U_\varepsilon$ 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 $U_\varepsilon$ and $\text{Hess} h_\varepsilon(p) \leq A(p) + \varepsilon g(p)$ as $(0, 2)$-tensor fields. (Such a function $h_\varepsilon$ is called an upper barrier.)

- For a continuous function $f : M \to \mathbb{R}$, we say that $\text{Hess} f \leq A$ at the point $p \in M$ in the *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 $\text{Hess} h(p) \leq A(p)$.

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 a function $f : M \to \mathbb{R}$ satisfies $\Delta f \leq \alpha$ in the *distributional sense* if for every smooth, nonnegative function $\varphi : M \to \mathbb{R}$ with compact support we have $\int_M f \Delta \varphi d\text{Vol} \leq \int_M \alpha \varphi d\text{Vol}$.

- For a continuous function $f : M \to \mathbb{R}$, we say that $\Delta f \leq \alpha$ at the point $p \in M$ in the *barrier sense* if for every $\varepsilon > 0$ there exists a neighborhood $U_\varepsilon$ 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 $U_\varepsilon$ 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 the *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 the rest of this paper we have used the Einstein summation convention on repeated indices. In particular, by $\nabla^2_{ij}(V^i V^j)$ we mean $\nabla^2_{ij}(V \otimes V)^{ij}$, the function obtained by contracting twice the second covariant derivative of the tensor product $V \otimes V$.*
The notion of inequalities in the barrier sense was defined by Calabi [1958] for the Laplacian (he used the terminology “weak sense” rather than “barrier sense”). He also proved the relative global “weak” Laplacian comparison theorem (see [Petersen 1998, Section 9.3]).

The notion of a viscosity solution (which is connected to the definition of inequality “in the viscosity sense”; see the Appendix) was introduced by Crandall and Lions [1983, 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 the maximum principle.

From the definitions it is easy to see that the barrier sense implies the viscosity sense; moreover, by [Ishii 1995], if \( f : M \to \mathbb{R} \) satisfies \( \Delta f \leq \alpha \) in the viscosity sense it also satisfies \( \Delta f \leq \alpha \) as distributions, and vice versa. In the Appendix 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 \( M \setminus (\{p\} \cup \text{Cut}_p) \), where \( \text{Cut}_p \) is the cut locus of the point \( p \), which we are now going to define and state some general properties of (we keep [Gallot et al. 1990; Sakai 1996] as general references). It is anyway well known that \( \text{Cut}_p \) is a closed set of zero (canonical) measure. Hence, in the open set \( M \setminus (\{p\} \cup \text{Cut}_p) \) 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 with what happens on \( \text{Cut}_p \) (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 \( U_p = \{ v \in T_p M \mid g_p(v, v) = 1 \} \) 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 \( \sigma_v \in \mathbb{R}^+ \) (or possibly equal to \( +\infty \)) be the maximal time such that \( \gamma_v([0, \sigma_v]) \) is minimal between any pair of its points. This defines a map \( \sigma : U_p \to \mathbb{R}^+ \cup \{+\infty\} \), and the point \( \gamma_v(\sigma_v) \) (when \( \sigma_v < +\infty \)) is called the cut point of the geodesic \( \gamma_v \).

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

There are two reasons why a geodesic can cease to be minimal:

**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 conditions is satisfied:

1. Another minimal geodesic from \( p \) arrives at the cut point \( q = \gamma_v(\sigma_v) \).
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 any interval larger that $[0, \sigma_v]$.

It is well known that the subset of points $q \in \text{Cut}_p$ where more than one minimal geodesic from $p$ arrive coincides with Sing, the singular set of the distance function $d_p$ in $M \setminus \{p\}$. We also define Conj, the set of points $q = \gamma_v(\sigma_v) \in \text{Cut}_p$ with $d_{exp} p$ not invertible at $\sigma_v \in T_p M$; we call Conj the locus of optimal conjugate points. See [Gallot et al. 1990; Sakai 1996].

2. The structure of the distributional Hessian of the distance function

The following properties of the function $d_p$ and the cut locus of $p \in M$ are proved in [Mantegazza and Mennucci 2003, Section 3] (see also the wonderful [Li and Nirenberg 2005] for other fine properties, notably the local Lipschitz continuity of the function $\sigma : U_p \rightarrow \mathbb{R}^+ \cup \{+\infty\}$ in Theorem 1.1 there).

Given an open set $\Omega \subset \mathbb{R}^n$, we say that a continuous function $u : \Omega \rightarrow \mathbb{R}$ is locally semiconcave if for any open convex set $K \subset \Omega$ with compact closure in $\Omega$, the function $u|_K$ is the sum of a $C^2$ function and a concave function.

A continuous function $u : M \rightarrow \mathbb{R}$ is called locally semiconcave if for any local chart $\psi : \mathbb{R}^n \rightarrow U \subset M$, the function $u \circ \psi$ is locally semiconcave in $\mathbb{R}^n$ according to the above definition.

**Proposition 2.1** [Mantegazza and Mennucci 2003, Proposition 3.4]. The function $d_p$ is locally semiconcave in $M \setminus \{p\}$.

This fact, which follows from recognizing $d_p$ as a viscosity solution of the eikonal equation $|\nabla u| = 1$ (see [Mantegazza and Mennucci 2003]), has some significant consequences; we need some definitions for the precise statements.

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

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

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** [Alberti et al. 1992, Proposition 2.1]. Let the function $u : M \rightarrow \mathbb{R}$ be semiconcave. Then the superdifferential $\partial^+ u$ is not empty at each point; moreover, $\partial^+ u$ is upper semicontinuous, that is,

$$q_k \rightarrow q, \quad \nu_k \rightarrow \nu, \quad \nu_k \in \partial^+ u(q_k) \implies \nu \in \partial^+ u(q).$$  

In particular, if the differential $du$ exists at every point of $M$, then $u \in C^1(M)$.
Proposition 2.3 [Alberti et al. 1992, Remark 3.6]. The set $\text{Ext}(\partial^+ d_p(q))$ 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 $q$. In symbols,

$$\mathcal{G}(q) = \{ t \mapsto \exp_q(-vt) \mid v \in \text{Ext}(\partial^+ d_p(q)) \},$$

where $t \in [0, 1]$.

We now deal with the structure of the cut locus of $p \in M$. Let $\mathcal{H}^{n-1}$ denote the $(n-1)$-dimensional Hausdorff measure on $(M, g)$ (see [Federer 1969; Simon 1983]).

Definition 2.4. We say that a subset $S \subset M$ is $C^r$-rectifiable, for $r \geq 1$, 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 $\mathcal{H}^{n-1}$-measure zero. (See the references just cited for a complete discussion of the notion of rectifiability.)

Proposition 2.5 [Mantegazza and Mennucci 2003, Theorem 4.10]. The cut locus of $p \in M$ is $C^\infty$-rectifiable. Hence, its Hausdorff dimension is at most $n-1$. Moreover, for any compact subset $K$ of $M$, the measure $\mathcal{H}^{n-1}(\text{Cut}_p \cap K)$ is finite [Li and Nirenberg 2005, Corollary 1.3].

To explain the following consequence of such rectifiability, we need to briefly introduce the theory of functions with bounded variation; see [Ambrosio et al. 2000; Braides 1998; Federer 1969; Simon 1983] for details. We say that a function $u : \mathbb{R}^n \to \mathbb{R}^m$ is a function with locally bounded variation, denoted $u \in \text{BV}_{\text{loc}}$, if its distributional derivative $Du$ is a Radon measure. This 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 $\text{BV}_{\text{loc}}$; in view of Proposition 2.1, this implies that the vector field $\nabla d_p$ belongs to $\text{BV}_{\text{loc}}$ in the open set $M \setminus \{p\}$. Then we define the subspace of $\text{BV}_{\text{loc}}$ of functions (or vector fields, as before) with locally special bounded variation, called $\text{SBV}_{\text{loc}}$ (see [Ambrosio 1989a; 1989b; 1990; Ambrosio et al. 2000; Braides 1998]).

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 into three mutually singular measures

$$Du = \tilde{Du} + Ju + Cu,$$

where the first term is the part absolutely continuous with respect to the Lebesgue measure $\mathcal{L}^n$, $Ju$ is a measure concentrated on an $(n-1)$-rectifiable set and $Cu$, called the Cantor part, vanishes on subsets of Hausdorff dimension $n-1$. 
The space $SBV_{loc}$ is defined as the class of functions $u \in BV_{loc}$ such that $Cu = 0$; that is, the Cantor part of the distributional derivative of $u$ is zero. Again, by means of local charts, this notion is easily generalized to Riemannian manifolds.

**Proposition 2.6** [Mantegazza and Mennucci 2003, Corollary 4.13]. The $(\mathcal{H}^{n-1}$-almost everywhere defined) measurable unit vector field $\nabla d_p$ belongs to the space $SBV_{loc}(M \setminus \{p\})$ of vector fields with locally special bounded variation.

An immediate consequence of this proposition is that the $(0, 2)$-tensor field valued distribution $Hess d_p$ is actually a Radon measure with an absolutely continuous part, with respect to the canonical volume measure $Vol$ of $(M, g)$, concentrated in $M \setminus ((\{p\} \cup \text{Cut}_p)$, where $d_p$ is a smooth function. Hence in this set $Hess d_p$ coincides with the standard Hessian $\widetilde{Hess} d_p$ times the volume measure $Vol$. When the dimension of $M$ is at least two, the singular part of the measure $Hess d_p$ does not “see” the singular point $p$; hence, it is concentrated on $\text{Cut}_p$ and absolutely continuous with respect to the Hausdorff measure $\mathcal{H}^{n-1}$ restricted to $\text{Cut}_p$.

By the properties of rectifiable sets, at $\mathcal{H}^{n-1}$-almost every point $q \in \text{Cut}_p$, there exists an $(n-1)$-dimensional approximate tangent space $ap_{\mathcal{T}_q} \text{Cut}_p \subset \mathcal{T}_q M$ (in the sense of geometric measure theory; see [Federer 1969; Simon 1983] for details). To give an example, we say that a hyperplane $T \subset \mathbb{R}^n$ is the approximate tangent space to an $(n-1)$-dimensional rectifiable set $K \subset \mathbb{R}^n$ at the point $x_0$ if $\mathcal{H}^{n-1} \{ T \cap \rho(K - x_0) \}$ around the point $x_0$. With some technicalities, this notion can be extended also to Riemannian manifolds.

Moreover (see [Ambrosio et al. 2000]), at $\mathcal{H}^{n-1}$-almost every point $q \in \text{Cut}_p$, the field $\nabla d_p$ has two distinct approximate (in the sense of the Lebesgue differentiation theorem) limits “on the two sides” of $ap_{\mathcal{T}_q} \text{Cut}_p \subset \mathcal{T}_q M$, given by $\nabla d_p^+$ and $\nabla d_p^-$. We want to see now that exactly two distinct geodesics and no more arrive at $\mathcal{H}^{n-1}$-almost every point of $\text{Cut}_p$. We underline that a stronger form of this theorem was already obtained in [Ardoy and Guijarro 2011] and [Figalli et al. 2011], concluding that the set $\text{Cut}_p \setminus U$ (where $U$ is as in the following statement) has Hausdorff dimension not greater that $n-2$.

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

(i) The subset $\text{Cut}_p \cap U$ does not contain conjugate points; hence the set of optimal conjugate points has $\mathcal{H}^{n-1}$-measure zero.

(ii) Exactly two minimal geodesics from $p \in M$ arrive at every point of $\text{Cut}_p \cap U$.

(iii) Locally around every point of $\text{Cut}_p \cap U$ the set $\text{Cut}_p$ is a smooth $(n-1)$-dimensional hypersurface; hence $ap_{\mathcal{T}_q} \text{Cut}_p$ 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 \( \mathcal{H}^{n-1} \)-measure zero, then we will see that the points of Sing \( \setminus \) Conj where more than two geodesics arrive also form a closed subset of \( \mathcal{H}^{n-1} \)-measure zero. Claim (iii) then follows by the analysis in the proof of Proposition 4.7 in [Mantegazza and Mennucci 2003].

Recalling that \( U_p = \{ v \in T_p M \mid g_p(v, v) = 1 \} \) 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 \( \gamma_v \); 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 [Mantegazza and Mennucci 2003], in the open subset \( V \subset U_p \) where the rank of the differential of the map \( F : U_p \to M \) defined by \( F(v) = \exp_p(c_v v) \) is \( n - 1 \), the map \( c : U_p \to \mathbb{R}^+ \cup \{ +\infty \} \) is smooth; hence \( F(V) \) is locally a smooth hypersurface. Since, by Sard’s theorem, the image of \( U_p \setminus V \) is a closed set of \( \mathcal{H}^{n-1} \)-measure zero, we only have to deal with the images \( F(v) \) of unit vectors \( v \in V \) such that \( c_v = \sigma_v \) (see end of the introduction), that is, with \( F(V) \cap \text{Cut}_p \), which is a closed set.

We then consider the set

\[
D \subset (F(V) \cap \text{Cut}_p)
\]

of points \( q \) where \( \text{ap}T_q \text{Cut}_p \) exists and the density of the rectifiable set \( F(V) \cap \text{Cut}_p \) in the cut locus of the point \( p \) with respect to the Hausdorff measure \( \mathcal{H}^{n-1} \) is 1 (see [Federer 1969; Simon 1983]). It is well known that \( D \) and \( F(V) \cap \text{Cut}_p \) only differ by a set of \( \mathcal{H}^{n-1} \)-measure zero. If \( F(v) = q \in D \), then \( c_v = \sigma_v \) and, by the above density property, the hypersurface \( F(V) \) is “tangent” to \( \text{Cut}_p \) at the point \( q \); that is, \( T_q F(V) = \text{ap}T_q \text{Cut}_p \).

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

\[
d F_v(w) = (dc[v](w)) \dot{\gamma}_v(c_v) + d\exp_p[c_v v](c_v w) = (dc[v](w)) \dot{\gamma}_v(c_v);
\]

thus, \( \dot{\gamma}_v(c_v) \) belongs to the tangent space \( d F(T_u U_p) \) to the hypersurface \( F(V) \) at the point \( q \), which coincides with \( \text{ap}T_q \text{Cut}_p \), 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”, meaning that exactly two minimal geodesics arrive at \( q \), both intersecting the cut locus transversally (the vectors \( \nabla d_p^+ \) and \( \nabla d_p^- \) do not belong to \( \text{ap}T_q M \)); hence the above minimal geodesic \( \gamma_v \) cannot coincide with any of these two.

We then conclude that \( \mathcal{H}^{n-1}(D) = 0 \), and the same for the set Conj.
Now suppose that $q \in \text{Cut}_p \setminus \text{Conj} \subset \text{Sing}$; by the analysis in the proof of Proposition 4.7 in [Mantegazza and Mennucci 2003] (and Lemma 4.8), a finite number $m \geq 2$ of distinct minimal geodesics arrive at the point $q$, and when $m > 2$ the cut locus of $p$ is given by the union of at least $m$ smooth hypersurfaces with Lipschitz boundary going through the point $q$. In particular, the above blow-up at $q$ cannot be a single hyperplane $apT_q\text{Cut}_p$. By the preceding discussion, the set of such points with $m > 2$ is then of $\mathcal{H}^{n-1}$-measure zero; moreover, by Propositions 2.2 and 2.3, the set of points in $\text{Cut}_p \setminus \text{Conj}$ with only two minimal geodesics is open, and we are done.

\begin{remark}
In the special two-dimensional and analytic case, more can be said: the number of optimal conjugate points is locally finite and the cut locus is a locally finite graph with smooth edges; see [Myers 1935; 1936]. We conjecture that in general the set of optimal conjugate points is an $(n-2)$-dimensional rectifiable set.
\end{remark}

By Theorem 2.7(iii), in the open set $U$ the two side limits $\nabla d_p^+$ and $\nabla d_p^-$ of the gradient field $\nabla d_p$ are actually smooth and classical limits; moreover, there is a locally defined smoothly varying unit normal vector $v_q \in T_qM$ orthogonal to $T_q\text{Cut}_p$, with the convention that $g_q(v_q, v)$ is positive for every vector $v \in T_qM$ belonging to the half-space corresponding to the side associated to $\nabla d_p^+$. Hence, since $\mathcal{H}^{n-1}(\text{Cut}_p \setminus U) = 0$, we have a precise description of the singular jump part as follows:

$$
J\nabla d_p = -((\nabla d_p^+ - \nabla d_p^-) \otimes v) \mathcal{H}^{n-1} \setminus \text{Cut}_p,
$$

and, noticing that the jump in the gradient of $d_p$ in $U$ must be orthogonal to the tangent space $T_q\text{Cut}_p$, and thus parallel to the unit normal vector $v_q \in T_qM$, we conclude

$$
J\nabla d_p = -(v \otimes v) \left| \nabla d_p^+ - \nabla d_p^- \right|_g \mathcal{H}^{n-1} \setminus \text{Cut}_p.
$$

Notice that the singular part of the distributional Hessian of $d_p$ is a rank-1 symmetric $(0, 2)$-tensor field.

\begin{remark}
This description of the jump part of the singular measure follows directly from the structure theorem for BV functions (see [Ambrosio et al. 2000]), even if we didn’t know from Theorem 2.7 that the cut locus is $\mathcal{H}^{n-1}$-almost everywhere smooth.
\end{remark}

\begin{theorem}
If $n \geq 2$, the distributional Hessian of the distance from a point $p \in M$ is given by the Radon measure

$$
\text{Hess} d_p = \widetilde{\text{Hess}} d_p \text{Vol} - (v \otimes v) \left| \nabla d_p^+ - \nabla d_p^- \right|_g \mathcal{H}^{n-1} \setminus \text{Cut}_p,
$$

where $\widetilde{\text{Hess}} d_p$ is the standard Hessian of $d_p$, where it exists ($\mathcal{H}^{n-1}$-almost everywhere on $M$), and $\nabla d_p^+, \nabla d_p^-$, $v$ are defined above.
\end{theorem}
Corollary 2.11. If \( n \geq 2 \), the distributional Laplacian of \( d_p \) is the Radon measure
\[
\Delta d_p = \tilde{\Delta} d_p \, \text{Vol} - |\nabla d_p^+ - \nabla d_p^-| g \, \mathcal{H}^{n-1} \mathcal{L} \text{Cut}_p,
\]
where \( \tilde{\Delta} d_p \) is the standard Laplacian of \( d_p \), where it exists.

Corollary 2.12. We have
\[
\Delta d_p \leq \tilde{\Delta} d_p \, \text{Vol}
\]
and
\[
\text{Hess } d_p \leq \tilde{\text{Hess}} d_p \, \text{Vol},
\]
as \((0, 2)\)-tensor fields. Hence the Hessian and Laplacian inequalities in Theorem 1.1 hold in the sense of distributions. Moreover,
\[
\Delta d_p \geq \tilde{\Delta} d_p \, \text{Vol} - 2 \mathcal{H}^{n-1} \mathcal{L} \text{Cut}_p
\]
and
\[
\text{Hess } d_p \geq \tilde{\text{Hess}} d_p \, \text{Vol} - 2g \mathcal{H}^{n-1} \mathcal{L} \text{Cut}_p,
\]
as \((0, 2)\)-tensor fields.

Remark 2.13. From their definition, it is easy to see that the same inequalities hold also for the Busemann functions; see for instance [Petersen 1998, Subsection 9.3.4] (in Section 9.3 of the same book, it is shown that the above Laplacian comparison holds on all of \( M \) in the barrier sense, while an analogous result for the Hessian can be found in Section 11.2). We stress 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 the barrier and viscosity senses.

Remark 2.14. Several of the conclusions of this paper also hold for the distance function from a closed subset of \( M \) with boundary of class at least \( C^3 \); see [Mantergazza and Mennucci 2003] for details.

Appendix: 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 \( \text{Hess } f \leq A \) at the point \( p \in M \) in the barrier sense, for every \( \varepsilon > 0 \) there exists a neighborhood \( U_\varepsilon \) 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 \( U_\varepsilon \) and \( \text{Hess } h_\varepsilon(p) \leq A(p) + \varepsilon g(p) \); 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
Hess $h(p) \leq \text{Hess} \, h_{\varepsilon}(p) \leq A(p) + \varepsilon g(p)$ for every $\varepsilon > 0$, hence $\text{Hess} \, h(p) \leq A(p)$. This shows that $\text{Hess} \, f \leq A$ at the point $p \in M$ also in the \textit{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 the viscosity sense but not in the barrier sense.

The same argument clearly also applies to the two definitions of $\Delta f \leq \alpha$ for a smooth function $\alpha: M \to \mathbb{R}$.

Nonetheless, the notions of viscosity sense and distributional sense coincide:

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

In order to show the proposition, we recall the definitions of \textit{viscosity} (sub/super) solutions 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 $\Omega$ 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 \textit{monotonicity condition}

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

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 to be a \textit{viscosity subsolution} of the above PDE if for every point $x \in \Omega$ and $\varphi \in C^2(\Omega)$ such that $f(x) - \varphi(x) = \sup_{\Omega} (f - \varphi)$, we have $F(x, \varphi, \nabla \varphi, \nabla^2 \varphi) \leq 0$ (see [Crandall et al. 1992; Ishii 1995]). Analogously, $f \in C^0(\Omega)$ is a \textit{viscosity supersolution} if for every point $x \in \Omega$ and $\varphi \in C^2(\Omega)$ such that $f(x) - \varphi(x) = \inf_{\Omega} (f - \varphi)$, we have $F(x, \varphi, \nabla \varphi, \nabla^2 \varphi) \geq 0$. If $f \in C^0(\Omega)$ is both a viscosity subsolution and supersolution, it is then a \textit{viscosity solution} of $F(x, f, \nabla f, \nabla^2 f) = 0$ in $\Omega$.

It is easy to see that the functions $f \in C^0(\Omega)$ such that $\Delta f \leq \alpha$ in the \textit{viscosity sense} at any point of $\Omega$, 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 $F(x, r, p, X) = -\text{trace} \, X + \alpha(x)$).

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

$$\text{Hess}^M_{ij} f(x) = \frac{\partial^2 f(x)}{\partial x^i \partial x^j} - \Gamma^k_{ij}(x) \frac{\partial f}{\partial x^k}$$

and

$$\Delta^M f(x) = g^{ij}(x) \text{Hess}^M_{ij} f(x),$$

where $\Gamma^k_{ij}$ are the Christoffel symbols.
Analogously to the case of $\mathbb{R}^n$, taking 

$$F(x, r, p, X) = -g^{ij}(x)X_{ij} + g^{ij}(x)\Gamma^k_{ij}(x)p_k + \alpha(x)$$

(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 the 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

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

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

$$\int_{\Omega} (fL^*\varphi + \alpha \varphi)\,dx \geq 0$$

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

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

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 [1995] showed 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 [Lions 1983]):

**Theorem A.2** [Ishii 1995, Theorem 1]. If $f \in C^0(\Omega)$ is a viscosity subsolution of the equation $Lf + \alpha = 0$, then it is a distribution subsolution of the same equation.

**Theorem A.3** [Ishii 1995, Theorem 2]. 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 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 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 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 vice versa. Moreover, notice that setting \( L = -\sqrt{g} \Delta^M \), we have that \( L^* = L \); that is, \( L \) is a self-adjoint operator. \( L \) also satisfies the hypotheses of Ishii’s theorems, since the matrices \( g_{ij} \) and \( g^{ij} \) are smooth and positive definite in \( \Omega \). See [Horn and Johnson 1994, 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, for each \( \varphi \in C^\infty_c(\Omega) \), \( f \) satisfies

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

hence,

\[
\int_{M} -f \Delta^M \varphi \, d\text{Vol} = \int_{\Omega} -f \sqrt{g} \Delta^M \varphi \, dx \geq -\int_{\Omega} \alpha \sqrt{g} \varphi \, dx = -\int_{M} \alpha \varphi \, d\text{Vol}.
\]

This shows that then \( f \) satisfies \( \Delta^M f \leq \alpha \) in the 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 the viscosity and distributional senses 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 \( \text{Hess} f \leq A \) into a family of scalar inequalities; indeed, if everything is smooth, such an inequality is satisfied if and only if for every compactly supported, smooth vector field \( W \) we have \( W^i W^j \text{Hess}_{ij} f \leq A_{ij} W^i W^j \). The only price to pay is that we lose the constant coefficients of the Hessians, hence making the linear operator \( L^W \), acting on \( f \in C^2(\Omega) \) as \( L^W f = -W^i W^j \text{Hess}_{ij} f \), only degenerate elliptic. Notice that Ishii’s condition in Theorem A.3 is satisfied for every smooth vector field \( W \) such that \( \|W\| \in C^1_\text{c}(\Omega) \), 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 \( \|W\| \in C^1_\text{c}(\Omega) \), we have that \( f \) is a viscosity supersolution of the equation \(-W^i W^j \text{Hess}_{ij} f + A_{ij} W^i W^j = 0\), then the function \( f \) satisfies \( \text{Hess} f \leq A \) in the viscosity sense in all of \( \Omega \).
Vice versa, if $f \in C^0(\Omega)$ satisfies $\text{Hess } f \leq A$ in the viscosity sense in $\Omega$, then $f$ is a viscosity supersolution of the equation $-V^i V^j \text{Hess}_{ij} f + A_{ij} V^i V^j = 0$ 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 $\varphi$ 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 $\|W\| = \varphi \in C^1_c(\Omega)$. By the hypothesis of the first statement, the function $f$ is then a viscosity supersolution of the equation $-W^i W^j \text{Hess}_{ij} f + A_{ij} W^i W^j = 0$, which implies that $-W^i_x W^j_x \text{Hess}_{ij} h(x) + A_{ij}(x) W^i_x W^j_x \geq 0$. Since this holds for every point $x \in \Omega$ and unit vector $W_x$, we conclude that $\text{Hess } h(x) \leq A(x)$ as $(0,2)$-tensor fields, and hence $\text{Hess } f \leq A$ in the viscosity sense in $\Omega$.

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 $-V^i_x V^j_x \text{Hess}_{ij} h(x) + A_{ij}(x) V^i_x V^j_x \geq 0$, hence the thesis. □

Suppose now that $f \in C^0(\Omega)$ satisfies $\text{Hess } f \leq A$ in the viscosity sense on the whole $\Omega$; hence, by this lemma, for every compactly supported, smooth vector field $V$, the function $f$ is a viscosity supersolution of the equation $-V^i V^j \text{Hess}_{ij} f + A_{ij} V^i V^j = 0$. 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 1 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 $\text{Hess } f \leq A$ in the distributional sense.

Conversely, if $f \in C^0(\Omega)$ satisfies $\text{Hess } f \leq A$ in the distributional sense, then for every compactly supported, smooth vector field $W$ with $\|W\| \in C^1_c(\Omega)$ and every smooth, nonnegative function $\varphi \in C^\infty_c(\Omega)$, we define the smooth, nonnegative functions $\varphi_n = \varphi + \psi/n$, where $\psi$ 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 $\text{Hess } f \leq A$ in the 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 as $n \to \infty$ 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 $-W^i W^j \text{Hess}_{ij} f + A_{ij} W^i W^j = 0$.

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 $\text{Hess} f \leq A$ in the viscosity sense.

Summarizing, we have the following sharp relations among the weak notions of the partial differential inequalities $\text{Hess} f \leq A$ and $1 f \leq \alpha$:

\[
\text{barrier sense } \implies \text{viscosity sense } \iff \text{distributional sense.}
\]

**Acknowledgements**

We thank Giovanni Alberti, Luigi Ambrosio, Davide Lombardo and Federico Poloni for several valuable suggestions. Carlo Mantegazza is partially supported by the Italian project FIRB–IDEAS “Analysis and Beyond.”

**References**


Received April 23, 2013. Revised January 15, 2014.
Hermitian categories, extension of scalars and systems of sesquilinear forms

EVA BAYER-FLUCKIGER, URIYA A. FIRST and DANIEL A. MOLDOVAN

Realizations of the three-point Lie algebra $\mathfrak{sl}(2, \mathbb{R}) \oplus (\Omega_{\mathbb{R}}/d\mathbb{R})$

BEN COX and ELIZABETH JURISICH

Multi-bump bound state solutions for the quasilinear Schrödinger equation with critical frequency

YUXIA GUO and ZHONGWEI TANG

On stable solutions of the biharmonic problem with polynomial growth

HATEM HAJLAOU, ABDELAZIZ HARRABI and DONG YE

Valuative multiplier ideals

ZHENGYU HU

Quasiconformal conjugacy classes of parabolic isometries of complex hyperbolic space

YOUNGJU KIM

On the distributional Hessian of the distance function

CARLO MANTEGAZZA, GIOVANNI MASCELLANI and GENNADY URALTSEV

Noether's problem for abelian extensions of cyclic $p$-groups

IVO M. MICHAILOV

Legendrian $\theta$-graphs

DANIELLE O’DONNOL and ELENA PAVELESCU

A class of Neumann problems arising in conformal geometry

WEIMIN SHENG and LI-XIA YUAN

Ryshkov domains of reductive algebraic groups

TAKAOKO WATANABE