AN EXTENSION OF ALEXANDROV’S THEOREM ON SECOND DERIVATIVES OF CONVEX FUNCTIONS

J.H.G. FU

Abstract. If \( f \) is a function of \( n \) variables that is locally uniformly approximable by a sequence of smooth functions satisfying local \( L^1 \) bounds on the determinants of the minors of the Hessian, then \( f \) admits a second order Taylor expansion almost everywhere. This extends a classical theorem of A.D. Alexandrov, covering the special case in which \( f \) is locally convex.

1. Introduction

A.D. Alexandrov [2] proved that if \( f : U \to \mathbb{R} \) is a locally convex function on a domain \( U \subset \mathbb{R}^n \) then \( f \) “admits a second derivative almost everywhere”, in the sense that for a.e. \( x \in U \) there is a second order Taylor expansion for \( f \) at \( x \). In the present article we extend this conclusion to a much larger class of functions. Let us say that \( f \) is twice differentiable at \( x \) if there is a quadratic polynomial \( Q_x : \mathbb{R}^n \to \mathbb{R} \) such that

\[
\lim_{y \to x} \frac{f(y) - Q_x(y)}{|y - x|^2} = 0.
\]

Theorem 1.1. Let \( U \subset \mathbb{R}^n \) be open. Suppose \( f : U \to \mathbb{R} \) may be expressed as the locally uniform limit of a sequence \( f_1, f_2, \ldots \) of smooth functions such that the absolute integrals of all minors of the Hessians of the \( f_k \) are locally bounded, i.e.

\[
\int_K \left| \det \left( \frac{\partial^2 f_k}{\partial x_i \partial x_j} \right)_{i \in I, j \in J} \right| \leq C(K), \quad k = 1, 2, \ldots
\]

whenever \( K \subset \subset U \) and \( I, J \subset \{1, \ldots, n\} \) have the same cardinality. Then \( f \) is twice differentiable at a.e. \( x \in U \).

(If \( n = 1 \) then \( f \) is approximable in this way iff it is expressible as the difference of two convex functions: but this is not the case for \( n \geq 2 \), as we show below.)

Alexandrov’s theorem is a special case: if \( f \) is convex then such a sequence \( f_k \) may be obtained by convolution with an approximate identity. But Theorem 1.1 is significantly stronger: even though a convex function \( f \) may fail to have first derivatives in the usual sense on a dense set, it has been observed (cf. [1]) that such \( f \) admits a multiple-valued differential everywhere, whose graph transforms into the graph of a Lipschitz function under the canonical linear change of variable \((x, y) \mapsto (x + y, x - y)\) of \( T^* \mathbb{R}^n \simeq \mathbb{R}^n \times \mathbb{R}^n \). From this point of view, Alexandrov’s theorem appears as a consequence of Rademacher’s theorem on the almost everywhere differentiability of Lipschitz functions. On the other hand, there exist functions \( f \) satisfying the hypothesis of Theorem 1.1 such that the graph of the
differential of $f$ is dense in $T^*\mathbb{R}^n$. In fact, such functions include the Sobolev space $W^{2,n}_{\text{loc}}(U)$ of functions with second distributional derivatives in $L^n_{\text{loc}}(U)$: convolution with an approximate identity yields a sequence of smooth functions with bounded local $W^{2,n}$ norms converging to $f$ in $W^{2,n}_{\text{loc}}$. By Hölder’s inequality, the local bounds imply that the resulting sequence satisfies the bounds (3), and the Sobolev embedding theorem implies that the convergence to $f$ is locally uniform. Meanwhile, Hutchinson-Meier [9] observed that if $n \geq 2$ then

$$f_{HM}(x) := x_1 \sin \log \log |x|^{-1}$$

belongs to $W^{2,n}(U)$ for sufficiently small neighborhoods $U$ of $0$, while $\nabla f_{HM}$ oscillates infinitely often between $(\pm 1,0,\ldots,0) + o(x)$ as $x \to 0$ along the $x_1$ axis. This implies first of all that $f_{HM}$ cannot be expressed as a difference of two convex functions. Furthermore, by cutting off, translating and multiplying by suitable constants, it is easy to construct a $W^{2,n}(\mathbb{R}^n)$-convergent sum whose differential has a graph that is dense in $T^*\mathbb{R}^n$.

As a final introductory remark, we note that this class of functions is a subset of the class of “Monge-Ampère functions” introduced by Jerrard in [10, 11] (extending the class defined in [5]). It is natural to conjecture that Theorem 1.1 applies to this nominally larger class as well. However, this is just one of many perplexing questions about Monge-Ampère functions: for example, we do not even know that such functions are necessarily continuous, or even locally bounded. On the other hand, it seems highly plausible that every Monge-Ampère function satisfies the hypotheses of Theorem 1.1.

Acknowledgement. I would like to thank Bob Jerrard for stimulating conversations on the topics discussed here.

2. BASIC FACTS

2.1. SOME MEASURE THEORY. Recall that if $\mu$ is a Radon measure on $U \subset \mathbb{R}^n$ then its upper density at $x \in U$ is

$$\Theta^*(\mu, x) := \limsup_{r \downarrow 0} \frac{\mu(B(x, r))}{\omega_n r^n},$$

and its density is

$$\Theta(\mu, x) := \lim_{r \downarrow 0} \frac{\mu(B(x, r))}{\omega_n r^n},$$

if the limit exists. In fact the limit exists for a.e. $x \in U$ with respect to Lebesgue measure, and defines a Lebesgue-integrable function of $x$, whose integrals yield the absolutely continuous part of $\mu$ with respect to the Lebesgue decomposition into absolutely continuous and singular parts (cf. [4], Thm. 3.22).

2.2. ABSOLUTE HESSIAN DETERMINANT MEASURES. From now on we fix an open set $U \subset \mathbb{R}^n$, a function $f: U \to \mathbb{R}$, and an approximating sequence $f_1, f_2, \cdots \to f$ as in the hypothesis of Theorem 1.1. In particular, $f$ is itself continuous. For $d = 0, \ldots, n$, and $k = 1, 2, \ldots$, we define the measures $\nu_{d,k}$ on $U$ by

$$\nu_{d,k}(S) := \sum_{I,J \subset \{1, \ldots, n\}, |I|=|J|=d} \int_S \left| \det \left( \frac{\partial^2 f_k}{\partial x^i \partial x^j} \right)_{i,j \in I} \right|.$$
Taking subsequences, we may assume that the each sequence \( \nu_{d,k}, k = 1, 2, \ldots \), converges weakly to a Radon measure \( \nu_d, d = 0, \ldots, n \). We will refer to the \( \nu_d \) as absolute Hessian determinant measures for \( f \).

The following lemma is obvious. For quadratic polynomials \( Q : \mathbb{R}^n \to \mathbb{R} \), put \( ||Q|| \) to be the maximum of the absolute values of the coefficients.

**Lemma 2.1.** Let \( Q \) be a quadratic polynomial. Then \( f + Q \) and the approximating sequence \( f_1 + Q, f_2 + Q, \ldots \) also satisfy the hypothesis of Theorem 1.1. Furthermore the resulting absolute Hessian determinant measures \( \tilde{\nu}_d \) for \( f + Q \) satisfy

\[
\tilde{\nu}_d \leq d! 8^n \sum_{i+j=d} ||Q||^i \nu_j, \quad d = 0, \ldots, n.
\]

### 2.3. An inequality from multivariable calculus

The key fact that makes the main theorem work is the extension of the following elementary classical inequality about \( C^2 \) functions to our larger class. Put \( \omega_n \) for the volume of the unit ball in \( \mathbb{R}^n \).

**Lemma 2.2.** If \( V \subset \subset U \) is open then

\[
\nu_n(V) \geq \omega_n (\sup_{\partial V} |f| - \sup_{\partial V} |f| - \sup_{\partial V} f - \sup_{\partial V} f - \epsilon) \geq \sup_{\partial V} f - \sup_{\partial V} f - |\lambda| \diam V
\]

\[
\sup_{\partial V} f - \sup_{\partial V} f - |\lambda| \diam V - 2\epsilon
\]

Thus if \( |\lambda| < \frac{\sup_{\partial V} f - \sup_{\partial V} f - 2\epsilon}{\diam V} \) then the left hand side is positive, and \( f_k - \lambda \) has interior local maximum in \( W \), at which \( df_k = \lambda \). Thus for \( k \geq k_0 \) let the image of \( W \) under \( \partial V \) includes the ball of radius \( \frac{\sup_{\partial V} f - \sup_{\partial V} f - 2\epsilon}{\diam V} \), and by the area formula and the properties of weak convergence of measures,

\[
\nu_n(V) \geq \nu_n(W) \geq \limsup_{k \to \infty} \int_W |\det D^2 f_k| \geq \omega_n \left( \frac{\sup_{\partial V} f - \sup_{\partial V} f - 2\epsilon}{\diam V} \right)^n
\]

by (7), and the lemma follows letting \( \epsilon \downarrow 0 \). \( \square \)

In fact we will need to apply Lemma 2.2 to affine subspaces as well.

**Corollary 2.3.** Given \( d < n \), let \( M \subset \mathbb{R}^n \) be a linear subspace of codimension \( d \). For each \( x \in M \), denote by \( P_x \) the affine \( d \)-plane orthogonal to \( M \) and passing through the point \( x \). If \( V \subset \subset U \) is open then

\[
\nu_d(V) \geq \omega_d \int_M \left| \frac{\sup_{V\cap P_x} |f| - \sup_{\partial(V \cap P_x)} |f| - \epsilon}{\diam(V \cap P_x)} \right|^d dx
\]
where the integrand is understood to be zero if \( V \cap P_x = \emptyset \).

**Proof.** Let \( W \subset \subset V \). By Lemma 2.2 for each approximating function \( f_k \)

\[
\omega_d \int_M \left| \frac{\sup_{W \cap P_x} |f_k| - \sup_{\partial(W \cap P_x)} |f_k|}{\text{diam}(W \cap P_x)} \right|^d \text{d}x \leq \int_M \int_{W \cap P_x} |\det D^2(f_k|P_x)| \text{d}x \leq \nu_{k,d}(W).
\]

By dominated convergence,

\[
\int_M \left| \frac{\sup_{W \cap P_x} |f| - \sup_{\partial(W \cap P_x)} |f|}{\text{diam}(W \cap P_x)} \right|^d \text{d}x = \lim_{k \to \infty} \int_M \left| \frac{\sup_{W \cap P_x} |f_k| - \sup_{\partial(W \cap P_x)} |f_k|}{\text{diam}(W \cap P_x)} \right|^d \text{d}x \leq \limsup_{k \to \infty} \nu_{k,d}(W) \leq \nu_d(V).
\]

Taking \( W = W_1, W_2, \ldots \) to be an exhaustion of \( V \), dominated convergence implies that the left hand side converges to the corresponding expression for \( V \). \( \square \)

3. **An extension of a special case of a theorem of Calderón and Zygmund**

We say that \( f \) admits an approximate second derivative at \( x \) if (1) holds with \( f \) replaced by the restriction \( f|E \), where \( E \) is some measurable set with density 1 at \( x \). We say that \( f \) admits a \( k \)th derivative in the \( L^1 \) sense at \( x \) if there exists a polynomial \( Q_x \) of degree \( k \) such that

\[
(9) \quad r^{-n} \int_{B_r(x)} |f(y) - Q_x(y)| \text{d}y = o(r^k)
\]
as \( r \downarrow 0 \). An easily proved equivalent formulation for first derivatives may be stated as follows.

**Lemma 3.1.** The function \( g \) is differentiable at 0 in the \( L^1 \) sense iff the one-parameter family of functions \( g_r(x) := rg(r^{-1}x) \) converges in \( L^1_{\text{loc}} \) to a linear function.

The following is also obvious.

**Lemma 3.2.** If \( f \) admits a second derivative in the \( L^1 \) sense at \( x \), then \( f \) admits an approximate second derivative at \( x \).

The next proposition generalizes a result of Calderón-Zygmund \[3\]. The original result of \[3\] (or rather the very special case of it that we have in mind) states that a function with distributional second derivatives in \( L^1_{\text{loc}} \) admits a second derivative in the \( L^1 \) sense a.e. By a straightforward adaptation of the argument of \[3\] we prove that this conclusion is true if the distributional second derivatives are only locally finite signed measures. We recall that the space \( \text{BV}_{\text{loc}}(U) \) of functions of locally bounded variation consists of all locally integrable functions whose distributional gradients are (vector) measures (cf. \[8\]).

**Proposition 3.3.** If the distributional gradient of \( f \in L^1(U) \) lies in \( \text{BV}_{\text{loc}}(U) \), then \( f \) admits a second derivative in the \( L^1 \) sense a.e. in \( U \).

**Lemma 3.4.** If \( g \in \text{BV}_{\text{loc}}(U) \) then \( g \) is differentiable in the \( L^1 \) sense at a.e. \( x \in U \).
Proof. By [4], Thm. 3.22, in the Lebesgue decomposition of the distributional gradient $\nabla g$ into its singular and absolutely continuous parts with respect to Lebesgue measure, the singular part has density zero at a.e. $x \in U$. Let $x = 0$ be such a point, and assume further that $0$ is a Lebesgue point for $g$ and for the absolutely continuous part of $\nabla g$. We may assume without loss of generality that $g(0) = 0$.

By the BV compactness theorem (cf. [8], Theorem 1.9), for some sequence $r_i \downarrow 0$ the dilates $G_i(x) := r_ig(r_i^{-1}x)$ converge in $L^1_{\text{loc}}(\mathbb{R}^n)$ to some function $\lambda \in L^1_{\text{loc}}(\mathbb{R}^n)$. Furthermore the distributional gradients $\nabla G_i$ converge weakly to the constant $\nabla g(0)$. Therefore $\lambda$ is the linear function $\lambda(x) := \sum_{i=1}^n \frac{\partial g}{\partial x_i}(0)x_i$.

Since this conclusion is independent of the sequence $r_i$, in fact $rg(r^{-1}x) \to \lambda(x)$ in $L^1_{\text{loc}}$, which by Lemma 3.4 is equivalent to the stated conclusion. $\square$

Proof of Prop. 3.3. By Lemma 3.4 it is enough to show that there exists an $L^1$ quadratic Taylor approximation for $f$ at 0 provided $\nabla f$ is differentiable in the $L^1$ sense at 0. We may assume also that $f(0) = \nabla f(0) = D^2f(0) = 0$, where $D^2f(0)$ is the $L^1$ derivative of $\nabla f$ at 0. Put

$$G(\rho) := \int_{B(0, \rho)} \frac{|\nabla f(x)|}{|x|^{n-1}} \, dx, \quad F(\rho) := \int_{B(0, \rho)} |\nabla f(x)| \, dx.$$ 

Then $F, G$ are both absolutely continuous on $[0, \infty)$, with $G'(\rho) = \frac{F'(\rho)}{\rho^{n-1}}$, and $F(\rho) = o(\rho^{n+1})$ by Lemma 3.4. Integrating by parts, it follows that $G(\rho) = o(\rho^2)$ as $\rho \downarrow 0$.

On the other hand,

$$\int_{B(0, \rho)} |f| \leq C \int_0^\rho r^{n-1} \, dr \int_{S^{n-1}} \, dv \int_0^r |Df(sv)| \, ds \leq C \int_0^\rho r^{n-1} \, dr \int_{S^{n-1}} \, dv \int_0^\rho |Df(sv)| \, ds = C \rho^n \int_{S^{n-1}} \, dv \int_0^\rho |Df(sv)| \, ds = C \rho^n G(\rho) = o(\rho^{n+2}),$$

which gives the result. $\square$

4. Proof of Theorem 1.1

To prove Theorem 1.1 we use induction on the dimension $n$. For $n = 1$, the result follows at once from Aleksandrov’s theorem and the parenthetical remark following the statement of the Theorem 1.1. Alternatively, it is also easy to prove by direct integration that $f$ is twice differentiable at every point at which $f'$ is differentiable in the $L^1$ sense. Supposing this to be true at 0, we may assume that the $L^1$ second derivative at 0 is 0, so the relation (9) yields

$$f(x) = \int_0^x f'(s) \, ds = o(x^2),$$

i.e. (9) holds at 0 with $Q_0 = 0$.

4.1. Setting up the inductive step. Put $H$ for the space of all hyperplanes in $\mathbb{R}^n$, not necessarily passing through the origin, and $I := \{(x, P) \in \mathbb{R}^n \times H : x \in P\} \simeq \mathbb{R}^n \times \mathbb{R}^{n-1}$. Thus $I$ admits the double fibration

$$(10) \quad \mathbb{R}^n \overset{p}{\leftarrow} I \overset{q}{\rightarrow} H.$$
Furthermore $I$ admits a measure $\mu$, invariant with respect to euclidean motions, expressible in local coordinates for either bundle structure as the product of invariant measures on the base and the fiber.

**Lemma 4.1.** Put $E$ for the set of pairs $(x, P) \in I$ such that $f|P$ is twice differentiable at $x$. Then $E$ is a Borel subset of $I$.

**Proof.** For $i, j, k \in \mathbb{N}$, define

$$E_{ijk} := \{(x, P) \in I : \text{there exists a quadratic polynomial } Q, \|Q\| \leq k, \text{ such that } |f(x') - Q(x')| \leq \frac{1}{i}|x' - x|^2 \text{ whenever } x' \in P \text{ and } |x' - x| < \frac{1}{j}\}.$$

We claim first that each $E_{ijk}$ is a closed subset of $I$. Suppose $E_{ijk} \ni (x_m, P_m) \to (x_0, P_0)$ and $x'_0 \in P_0, |x'_0 - x_0| < \frac{1}{j}$. Let $Q_m$ be a quadratic polynomial as in the definition of $E_{ijk}$ for $(x_m, P_m)$. We may assume that the sequence of the $Q_m$ converges to a quadratic polynomial $Q_0$ with $\|Q_0\| \leq k$. Clearly there exist points $x'_m \in P_m$ with $|x'_m - x_m| < \frac{1}{j}$ and $x'_m \to x'_0$ (e.g., for large $m$ we may take $x'_m$ to be the orthogonal projection of $x'_0$ to $P_m$) so

$$|f(x'_0) - Q_0(x'_0)| = \lim |f(x'_m) - Q_m(x'_m)| \leq \frac{1}{i} \lim |x'_m - x_m|^2 = \frac{1}{i} |x'_0 - x_0|^2.$$

Thus $(x_0, P_0) \in E_{ijk}$.

Next we claim that

$$E = \bigcup_{i,j,k} E_{ijk} =: F,$$

which is sufficient to establish the desired conclusion.

That $E \subset F$ is obvious. To prove that $F \subset E$, suppose $(x, P) \in F$. Since $\|Q(x - x)\| \leq (1 + 2|x| + |x|^2) \|Q\|$, we may suppose that $x = 0$. Fix $k$, and suppose that $(0, P) \in \bigcap_i \bigcup_j E_{ijk}$. Then for each $i$ there is $j_i$ such that $(0, P) \in E_{ijk}$; let $Q_i$ be the quadratic polynomial in the corresponding defining condition. Clearly the restrictions $Q_i|_P$ all agree to first order at 0, and $\|(Q_i - Q_i)(v)\| \leq \frac{2|v|^2}{\min_i \langle x', x' \rangle}$ for $v \in P$. Altering the $Q_i$ off of $P$ if necessary, we may assume that this is true for all $v \in \mathbb{R}^n$. Thus the $Q_i$ converge to a quadratic polynomial $Q_0$, and for $x' \in P, |x'| < \frac{1}{j}$,

$$|f(x') - Q_0(x')| \leq |f(x') - Q_i(x')| + |Q_i(x') - Q_0(x')|$$

$$\leq \frac{1}{i} |x'|^2 + \frac{2}{j} |x'|^2 = \frac{3}{i} |x'|^2.$$

It follows that $f|P$ is twice differentiable at 0, with second order Taylor expansion $Q_0$, which establishes the claim. \hfill \Box

Assume the conclusion of Theorem 3.1 is true for dimension $n - 1$, and let $x_0 \in \mathbb{R}^n$ be a point at which $f$ is approximately twice differentiable—this is true for a.e. $x_0 \in U$ by Lemma 3.2 and Prop. 3.3. We may assume further that $f|P$ is twice differentiable at $x_0$ for almost every hyperplane $P$ passing through $x_0$: it is clear that for a.e. $P$, the hypotheses of Theorem 3.1 hold for some subsequence $f_k|P \to f|P$, hence $f|P$ is twice differentiable a.e. in $P$ by induction. Thus by Lemma 4.1 and Fubini’s theorem, the set $\{(x, P) : f|P$ is twice differentiable at $x)\}$ has full measure in $I$. Applying Fubini’s theorem again, we find that for a.e. $x \in \mathbb{R}^n$, the set $\{P : f|P$ is twice differentiable at $x)\}$ has full measure in the space of hyperplanes passing through $x$. We will show that if $x_0$ is such a point, and if $f$ does not admit a second order Taylor expansion at $x_0$, then at least one of the
absolute Hessian determinant measures \( \nu_0, \ldots, \nu_n \) has upper density \( \Theta^*(\nu_i, x_0) = \infty \) at \( x_0 \). This will conclude the proof.

We may assume that \( x_0 = 0 \). We may assume also that the second order approximate Taylor expansion \( Q \) of \( f \) at 0 is identically zero: otherwise, in light of Lemma 2.1 we may replace \( f \) by \( f - Q \). With these assumptions we have

**Lemma 4.2.** For a.e. hyperplane \( P \) passing through 0, the second order Taylor expansion \( Q^P \) of \( f|_P \) at 0 is zero.

**Proof.** Let \( E \subset U \) be the subset of density 1 at 0 in the definition of the approximate second derivative at the beginning of section 3. Since \( f \) is continuous, we may assume that \( E \) is closed.

We claim that a.e. line \( l \) through 0 meets \( E \) at points lying arbitrarily close to 0, i.e. \( l \cap B(0, r) \cap E \neq \emptyset \) for all \( r > 0 \). In other words, putting \( A_r := \{ l : l \cap B(0, r) \cap E \neq \emptyset \} \) for \( r > 0 \) and \( A_0 := \cap_{r > 0} A_r \), where \( A_r \downarrow A_0 \) as \( r \downarrow 0 \), we claim that \( A_0 \) has full measure in the space of lines through 0. If this were not the case then there would exist some \( r_0 > 0 \) such that \( A_{r_0} \) does not have full measure, from which it follows that for \( r < r_0 \)

\[
\frac{\text{vol}(B(0, r) \cap E)}{\omega_n r^n} \leq \frac{\text{vol}(B(0, r_0) \cap \bigcup_{l \in A_{r_0}} l)}{\omega_n r_0^n} < 1,
\]

which is a contradiction.

Now, if \( l \in A_0 \) then \( \liminf_{v \to 0, v \perp l} \frac{|f'(v)|}{|v|} = 0 \). Hence if \( v \in l \subset P \), and \( f|_P \) is twice differentiable at 0, then \( Q^P(v) = 0 \). But, by basic integral geometry, a.e. \( P \) through 0 has the property that a.e. line \( l \subset P \) through 0 belongs to \( A_0 \). Therefore \( Q^P \equiv 0 \) on a dense subset of \( P \), and therefore \( Q^P \equiv 0 \) by continuity. \( \square \)

**4.2. The main lemma.** For \( \epsilon > 0 \) we consider the open sets

\[
V_\epsilon := \{ x : |f(x)| > \epsilon |x|^2 \}
\]

The approximate twice differentiability hypothesis means that each \( V_\epsilon \) has density 0 at 0. The statement that \( f \) is twice differentiable at 0, with zero 2nd order Taylor expansion there, is equivalent to the statement that for each \( \epsilon > 0 \) the set \( V_\epsilon \cap B(x, r) = \emptyset \) for all sufficiently small \( r > 0 \).

**Lemma 4.3.** Suppose the upper densities \( \Theta^*(\nu_i, 0) < \infty, i = 0, \ldots, n - 1 \). Then for fixed \( \epsilon > 0 \), the supremum of the diameters of the components of \( B(0, r) \cap V_\epsilon \) is \( o(r) \) as \( r \downarrow 0 \).

**Proof of Lemma 4.3.** Put \( I_0 \) for the space of hyperplanes through 0. By Lemma 4.2 for a.e. \( P \in I_0 \) there exists \( r = r(P) > 0 \) such that

\[
P \cap V_\epsilon \cap B(0, r) = \emptyset.
\]

For \( r > 0 \), put

\[
\Pi_r := \{ P : P \in I_0, P \cap V_\epsilon \cap B(0, r) \neq \emptyset \}.
\]

Thus the family of subsets \( \Pi_r \subset I_0 \) is decreasing as \( r \downarrow 0 \), with

\[
\lim_{r \downarrow 0} \mu_0(\Pi_r) = \mu_0(\bigcap_{r > 0} \Pi_r) = 0,
\]

where \( \mu_0 \) is the invariant probability measure on \( I_0 \).
Now consider the spherical open sets
\begin{equation}
W_{r,\epsilon} := \{ v \in S^{n-1} : \text{there exists } r' \in (0, r) \text{ such that } r'v \in V_\epsilon \},
\end{equation}
i.e. \( W_{r,\epsilon} \) is the spherical projection of \( B(0, r) \cap V_\epsilon \). Thus \( P \in \Pi_r \iff P \cap W_{r,\epsilon} \neq \emptyset \).

It now follows from \cite{12} and elementary integral geometry (cf. \cite{13}) that the supremum of the spherical diameters of the components of \( W_{r,\epsilon} \) converges to 0 as \( r \downarrow 0 \) (\( \epsilon \) fixed).

To complete the proof of the lemma, it is now enough to show that the “radial diameter” \( \sup_C |x| - \inf_C |x| = o(r) \) for any component \( C \) of \( B(0, r) \cap V_\epsilon \).

Let \( \delta < \frac{\pi}{6} \) be given, and take \( r_0 > 0 \) small enough that
\begin{enumerate}[(1)]
  \item each component of \( W_{2r_0/\sqrt{3}, \epsilon} \subset S^{n-1} \) is included in a spherical ball of radius \( \delta \), and
  \item \( \frac{\omega_{n-1}(B(0, 2r_0/\sqrt{3})))}{\omega_n(2r_0/\sqrt{3})} < \Theta^*(\nu_{n-1}, 0) + 1. \)
\end{enumerate}

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure1.png}
\caption{Proof of Lemma 4.3}
\end{figure}

Let \( C \) be any component of \( V_\epsilon \cap B(0, r_0) \), and \( C' \) the component of \( V_{\epsilon/2} \cap B(0, 2r_0/\sqrt{3}) \) that includes \( C \). Assumption (1) above guarantees that there is a spherical ball of radius \( \delta \) that includes the spherical projection of \( C' \). Let \( v_0 \) be its center and put for \( t \in \mathbb{R} \)
\[ P_t := \{ x : \langle x, v_0 \rangle = t \}. \]
Finally, let \((a, b) \coloneqq \{ t \in (0, r_0) : P_t \cap C \neq \emptyset \}\). Thus for \( t \in (a, b) \)
\begin{equation}
\sup_{C \cap P_t} \left| f \right| = \sup_{C \cap P_t} \left| f \right| > \epsilon t^2.
\end{equation}

By assumption [1],
\begin{equation}
\sup_C \left| x \right| - \inf_C \left| x \right| < (b - a) \sec \delta < (b - a) \frac{\sqrt{3}}{2}
\end{equation}
and, if \( x \in C' \cap P_t \),
\begin{equation}
t \leq \left| x \right| < t \sec \delta < \frac{2t}{\sqrt{3}}.
\end{equation}
In particular, if \( P_t \cap B(0, r_0) \neq \emptyset \) (i.e. if \( t < r_0 \)) then \( C' \cap P_t \subset \subset B(0, 2r_0/\sqrt{3}) \), and in particular
\begin{equation}
|f(x)| = \frac{\epsilon}{2} |x|^2 < \frac{2 \epsilon}{3} t^2
\end{equation}
for \( x \in \partial C' \cap P_t \). Thus we may apply Corollary 2.3 to obtain
\begin{equation}
(\Theta^*(\nu_{n-1}, 0) + 1) \omega_n \left( \frac{2r_0}{\sqrt{3}} \right)^n > \nu_{n-1} \left( \frac{2r_0}{\sqrt{3}} \right)^n
\end{equation}
\begin{equation}
\geq \int_a^b \frac{\sup_{C \cap P_t} f - \sup_{\partial C \cap P_t} f}{\text{diam}(C \cap P_t)} \left| f \right|^{n-1} dt
\end{equation}
\begin{equation}
\geq \int_a^b \left( \frac{\epsilon}{2} t^2 - \frac{2 \epsilon}{3} t^2 \right)^{n-1} \frac{n-1}{2t \tan \delta} \left( \frac{|x|}{2} \right)^n dt
\end{equation}
\begin{equation}
= \epsilon \cot \delta \left( \frac{b^n - a^n}{n} \right) \frac{n}{n}
\end{equation}
\begin{equation}
\geq \left( \frac{\epsilon}{6} \right) \cot \delta \left( b^n - a^n \right).\]
Therefore, by [17],
\begin{equation}
\sup_C \left| x \right| - \inf_C \left| x \right| \leq \frac{2}{\sqrt{3}}(b - a) < \frac{C}{\epsilon} (\Theta^*(\nu_{n-1}, 0) + 1)^\frac{n}{2} r_0 \tan \delta.
\end{equation}
Since \( \delta > 0 \) can be taken arbitrarily small (and \( \epsilon \) is fixed), this completes the proof of the lemma.

4.3. Completion of the proof Theorem 1.1. Suppose \( x_1, x_2, \ldots \in V_i \) with \( x_i \to 0 \). Let \( C_i \) be the component of \( V_i \cap B(0, 2|x_i|) \) containing \( x_i \), and \( C'_i \) the component of \( V'_i \cap B(0, 2|x_i|) \) that includes \( C_i \). By Lemma 4.3 the diameter of \( C'_i \) is \( o(|x_i|) \) as \( i \to \infty \), hence \( C'_i \subset \subset B(0, 2|x_i|) \) for \( i \) sufficiently large. Applying Lemma 2.2
\begin{equation}
\frac{\nu_n(B(0, 2|x_i|))}{|x_i|^n} \geq \frac{\nu_n(C'_i)}{|x_i|^n}
\end{equation}
\begin{equation}
\geq \frac{\sup_{C'_i} |f| - \sup_{\partial C'_i} |f|}{|x_i| o(|x_i|)}
\end{equation}
\begin{equation}
\geq \left( \frac{\epsilon}{2} \right) \left( \frac{|x_i|}{o(|x_i|)} \right)^n
\end{equation}
\begin{equation}
= \frac{1}{o(1) \to \infty}
\end{equation}
as $i \to \infty$ since $\epsilon$ is fixed. Therefore the upper density of $\nu_i$ at 0 is infinite.

4.4. A concluding remark. In fact our proof yields a slightly more general statement than Theorem 1.1: besides continuity, the only properties of $f$ that we have used are the existence of the absolute Hessian determinant measures, the transformation law of Lemma 2.1 and their role in Lemmas 2.2 and 2.3.

References

[1] Alberti, G., Ambrosio, L. A geometrical approach to monotone functions in $\mathbb{R}^n$. Math. Z. 230 (1999), pp. 259–316.
[2] Aleksandrov, A.D. Almost everywhere existence of the second differential of a convex function and some properties of convex surfaces connected with it. Leningrad Univ. Ann. (Math. ser.) 6 (1939), 335 (in Russian).
[3] Calderon, A.P. and Zygmund, A. Local properties of elliptic partial differential equations. Studia Math. 20 (1961), 171–225
[4] Folland, G.B. Real analysis. Wiley, New York, 1999.
[5] Fu, J. H. G.: Monge-Ampère functions I. Indiana Univ. Math. J. 38 (1989), 745–771
[6] Fu, J. H. G.: Monge-Ampère functions II. Indiana Univ. Math. J. 38 (1989), 773–789
[7] Gilbarg, D. and Trudinger, N. Elliptic Partial Differential Equations of Second Order. Springer, New York, 1983
[8] Giusti, E. Minimal Surfaces and Functions of Bounded Variation. Birkhäuser, Boston, 1984
[9] Hutchinson, J.E. and Meier, M. A remark on the nonuniqueness of tangent cones. Proc. Amer. Math. Soc. 97 (1986), pp. 184–185.
[10] Jerrard, R.L.: Some remarks on Monge-Ampère functions. Singularities in PDE and the Calculus of Variations, Stanley Alama, Lia Bronsard, and Peter J. Sternberg, eds. - AMS — CRM, 2008
[11] Jerrard, R.L.: Some rigidity results related to Monge-Ampère functions. To appear in Canadian Journal of Mathematics
[12] Santaló, L. Integral geometry and geometric probability, 2nd ed. Cambridge University Press, Cambridge 2004

E-mail address: fu@math.uga.edu

Department of Mathematics, University of Georgia, Athens, GA 30602, USA