

A GEOMETRIC APPROACH TO SECOND-ORDER DIFFERENTIABILITY OF CONVEX FUNCTIONS

ABSTRACT. We show a new, elementary and geometric proof of the classical Alexandrov theorem about the second order differentiability of convex functions. We also show new proofs of recent results about Lusin approximation of convex functions and convex bodies by $C^{1,1}$ convex functions and convex bodies.

1. INTRODUCTION

The aim of this paper is to provide a new, elementary and geometric proof of the following classical theorem of Alexandrov [1]. For a history of the theorem and a list of known proofs, see [9].

Theorem 1.1. *If $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is convex, then it is differentiable a.e. and at almost every point where f is differentiable, there is a symmetric matrix denoted by $D^2f(x)$ such that*

$$\lim_{y \rightarrow x} \frac{f(y) - f(x) - Df(x)(y - x) - \frac{1}{2}(y - x)^T D^2f(x)(y - x)}{|y - x|^2} = 0. \quad (1)$$

Our proof is so simple and geometric in nature that its concept can be described in just a few sentences. All notation used in the Introduction will be explained in Section 2.

Consider the set $W(\delta)$, the union of all closed balls of radius $\delta > 0$ contained in the epigraph of f . The set $W(\delta)$ is convex and it is the epigraph of a convex function g . Clearly, $g \geq f$. Using elementary and geometric arguments we show that if $\delta > 0$ is sufficiently small, then the set $W(\delta)$ touches the graph of f along a set that is large in the sense of measure. More precisely, for every $R > 0$ and every $\varepsilon > 0$, there is $\delta > 0$ such that

$$|\{x \in B^n(0, R) : f(x) \neq g(x)\}| < \varepsilon.$$

Since the convex set $W(\delta)$ is the union of balls of fixed radius, it is well known and easy to prove that the boundary of $W(\delta)$ is of class $C_{\text{loc}}^{1,1}$, and hence $g \in C_{\text{loc}}^{1,1}$, i.e., the gradient of g is locally Lipschitz continuous. Since $g \in C_{\text{loc}}^{1,1}$, it follows from the Rademacher theorem that g is twice differentiable almost everywhere in the classical sense. Now it remains to observe that at almost all points x such that $f(x) = g(x)$, f is twice differentiable in the sense of (1). Namely, this is true whenever x is a density point of the set $\{f = g\}$ and g is twice differentiable at x . In that case (1) is satisfied with $D^2f(x) := D^2g(x)$.

There is also a second version of the Alexandrov theorem which says that the subdifferential ∂f is differentiable a.e.

Theorem 1.2. *If $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is convex, then for all $x \in \mathbb{R}^n$ where f is twice differentiable as in (1), we have*

$$\lim_{y \rightarrow x} \sup_{\sigma_y \in \partial f(y)} \frac{|\sigma_y - Df(x) - D^2f(x)(y - x)|}{|y - x|} = 0. \quad (2)$$

2020 *Mathematics Subject Classification.* 26B25, 28A75, 41A30, 52A20, 52A27, 53C45.

Key words and phrases. Convex function, convex body, Alexandrov theorem, Lusin property, Lipschitz gradient.

The usual way to prove Theorem 1.1 is to show Theorem 1.2 first and conclude Theorem 1.1 from it. In our approach we will prove Theorem 1.1 directly and we will conclude Theorem 1.2 as a corollary.

The argument described above also leads to new and elementary proofs of the following recent results [5, Theorem 1.4, Corollary 1.7, Theorem 1.12 and Corollary 1.13].

Theorem 1.3. *Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a convex function. Then, for every measurable set $A \subset \mathbb{R}^n$ of finite Lebesgue measure, and for every $\varepsilon > 0$, there exists a convex function $g \in C^{1,1}(\mathbb{R}^n)$ such that*

$$|\{x \in A : f(x) \neq g(x)\}| < \varepsilon.$$

Theorem 1.4. *Let K be a convex body in $\mathbb{R}^n$. Then for every $\varepsilon > 0$, there is a convex body $W \subset K$, such that $\partial W \in C^{1,1}$ and*

$$\mathcal{H}^{n-1}(\partial K \Delta \partial W) < \varepsilon.$$

In fact, there is $\delta_o > 0$ such that for every $\delta \in (0, \delta_o)$, the set W defined as the union of all closed balls of radius δ that are contained in K satisfies the claim of the theorem.

A *convex body* is a compact convex set $K \subset \mathbb{R}^n$ with non-empty interior. Notation $A \Delta B$ stands for the symmetric difference of the sets A and B , that is, $A \Delta B := (A \setminus B) \cup (B \setminus A)$, and $\mathcal{H}^s$ denotes the s -dimensional Hausdorff measure.

Theorem 1.5. *Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a convex function, and assume that $f \notin C_{\text{loc}}^{1,1}(\mathbb{R}^n)$. Then the following conditions are equivalent:*

(1) *For every $\varepsilon > 0$ there exists a convex function $g \in C_{\text{loc}}^{1,1}(\mathbb{R}^n)$ such that*

$$|\{x \in \mathbb{R}^n : f(x) \neq g(x)\}| < \varepsilon.$$

(2) *The function f is essentially coercive.*

Moreover, if f is essentially coercive, we can find g satisfying $g \geq f$.

We call a convex function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ *essentially coercive* if there exists a linear function $\ell : \mathbb{R}^n \rightarrow \mathbb{R}$ such that $\lim_{|x| \rightarrow \infty} (f(x) - \ell(x)) = \infty$; this is equivalent to saying that the epigraph of f does not contain lines, see [7, Theorem 1.11]. Here and in what follows by a line we mean a set isometric to $\mathbb{R}$ so half-line is not a line.

Theorem 1.6. *Let S be a convex hypersurface of $\mathbb{R}^n$, and assume that S is not of class $C_{\text{loc}}^{1,1}$. Then the following assertions are equivalent:*

(1) *For every $\varepsilon > 0$ there exists a convex hypersurface S_ε of $\mathbb{R}^n$ of class $C_{\text{loc}}^{1,1}$ such that*

$$\mathcal{H}^{n-1}(S \Delta S_\varepsilon) < \varepsilon.$$

(2) *S does not contain any line.*

We call the boundary ∂W of a closed convex set W with nonempty interior (not necessarily bounded) a *convex hypersurface*, and we say that it is of class $C_{\text{loc}}^{1,1}$ if it is locally a graph of a $C^{1,1}$ function (if the set W is unbounded, we will say that W is an *unbounded convex body*).

Remark 1.7. It follows from the proof that if $S = \partial W$, where W is a (possibly unbounded) convex body that contains no lines, then there exists a (possibly unbounded) convex body $W_\varepsilon \subset W$ such that $S_\varepsilon := \partial W_\varepsilon \in C_{\text{loc}}^{1,1}$ and $\mathcal{H}^{n-1}(S \Delta S_\varepsilon) < \varepsilon$.

From Theorem 1.6 we will also deduce the following new generalization of Theorem 1.5 for convex functions defined on arbitrary open convex subsets of $\mathbb{R}^n$.

Theorem 1.8. *Let $U \subset \mathbb{R}^n$ be open and convex, and let $f : U \rightarrow \mathbb{R}$ be a convex function, such that $f \notin C_{\text{loc}}^{1,1}(U)$. Then, the following statements are equivalent:*

- (1) *For every $\varepsilon > 0$ there exists a convex function $g \in C_{\text{loc}}^{1,1}(U)$ such that*

$$|\{x \in U : f(x) \neq g(x)\}| < \varepsilon. \quad (3)$$
- (2) *The graph of f does not contain any line of $\mathbb{R}^{n+1}$.*

Moreover, if the graph of f contains no lines, we can find g satisfying $g \geq f$.

Remark 1.9. It was recently proved in [4] that if $f : U \rightarrow \mathbb{R}$ is locally strongly convex, then there is a locally strongly convex function $g \in C^2(U)$ that satisfies (3) (and other estimates). The proof is however, much more difficult.

The original proofs of Theorems 1.3, 1.4 and 1.5 used the Whitney extension theorem for convex functions [6, 8, 3], and the Alexandrov Theorem 1.1. Our proofs presented here are elementary and avoid these tools. As explained above, the proofs are based on a simple geometric idea that is also used in our proof of Alexandrov's theorem.

We will prove Theorem 1.4 first and we will use it as a main tool in the proofs of Theorems 1.1 and 1.3. Indeed, the brief description of the proof of Theorem 1.1 presented above is based on the approximation of the epigraph of f by the convex set $W(\delta)$ of class $C_{\text{loc}}^{1,1}$ and this is strictly related to Theorem 1.4.

Except Section 7, our exposition is elementary and self-contained. We have made an effort to make it accessible to anyone with basic knowledge of real analysis, and no knowledge in convex analysis is required.

The paper is structured as follows. In Section 2 we fix notation and recall basic definitions and facts needed to understand the paper. All results mentioned in this section are well known. In Section 3 we prove Theorem 1.4 and then we use it to prove Corollary 3.10 which is a version of Theorem 1.3. This corollary will play a central role in the proofs of Theorems 1.1, 1.3 and 1.5. Theorems 1.1 and 1.3 are proved in Sections 4 and 6 respectively. In Section 5 we prove Theorem 1.2 as a direct consequence of Theorem 1.1. This proof is independent of all other sections of the paper and it can be read independently. In Section 7 we present the proofs of Theorems 1.5, 1.6, and 1.8.

We made an effort to make different parts of the paper as independent as possible. Section 3 is needed in Sections 4, 6 and 7, but the content in Sections 4, 6 and 7 of the paper stands alone and is not dependent on one another. Similarly Section 5 is independent of any other part of the paper.

2. PRELIMINARIES

In this brief section we will explain notation and basic facts needed in the paper. This section will also clarify necessary prerequisites. By no means the definitions and facts presented here are detailed. The reader may find missing details in standard textbooks.

Balls in $\mathbb{R}^n$ are denoted by $B(x, r)$ or $B^n(x, r)$. The unit sphere in $\mathbb{R}^n$ that is centered at the origin is denoted by $\mathbb{S}^{n-1}$. The interior of a set A is denoted by $\text{int } A$. An interval in $\mathbb{R}^n$ with

endpoints $x, y \in \mathbb{R}^n$ is denoted by $[x, y]$. The scalar product of vectors $u, v \in \mathbb{R}^n$ is denoted by $\langle u, v \rangle$.

The Lebesgue measure of $A \subset \mathbb{R}^n$ is denoted by $|A|$. We say that $x \in \mathbb{R}^n$ is a *density point* of a measurable set $A \subset \mathbb{R}^n$ if $\frac{|A \cap B(x, r)|}{|B(x, r)|} \rightarrow 1$ as $r \rightarrow 0^+$. It follows from the Lebesgue differentiation theorem that almost all points $x \in A$ are density points of A .

The Hausdorff measure is denoted by $\mathcal{H}^s$. It follows from the definition that if f is L -Lipschitz, then $\mathcal{H}^s(f(A)) \leq L^s \mathcal{H}^s(A)$. If $A \subset \mathbb{R}^n$, $\lambda > 0$ and $\lambda A = \{\lambda x : x \in A\}$ is the dilation of A by the factor λ , then $\mathcal{H}^s(\lambda A) = \lambda^s \mathcal{H}^s(A)$. $\mathcal{H}^n$ coincides with the Lebesgue measure in $\mathbb{R}^n$.

We say that $f \in C^{1,1}(U)$ ($f \in C_{\text{loc}}^{1,1}(U)$), if $U \subset \mathbb{R}^n$ is open, $f \in C^1(U)$, and Df is Lipschitz (locally Lipschitz) continuous on U . If $f \in C^{1,1}(B^n(0, R))$, then it follows that

$$|f(y) - f(x) - Df(x)(y - x)| \leq M|y - x|^2 \quad \text{for all } x, y \in B^n(0, R), \quad (4)$$

where M is the Lipschitz constant of Df . Indeed, we can write $f(y) - f(x) = Df(\xi)(y - x)$ for some $\xi \in [x, y]$ and (4) follows. This inequality implies that if $f \in C_{\text{loc}}^{1,1}(U)$, where $U \subset \mathbb{R}^n$ is open, then

$$f(y) = f(x) + Df(x)(y - x) + O(|y - x|^2) \quad \text{for all } x, y \in U. \quad (5)$$

We say that the boundary of a bounded domain $U \subset \mathbb{R}^n$ is of class $C^{1,1}$ if it is locally a graph of a $C^{1,1}$ function.

We use notation $\nabla f(x)$ for the gradient vector while $Df(x)$ is the linear derivative. With this notation we have $Df(x)v = \langle \nabla f(x), v \rangle$.

If $W \subset \mathbb{R}^n$ is a closed convex set, then it is easy to see that for every $x \in \mathbb{R}^n$, there is a unique point denoted by $\pi_W(x)$ such that

$$\pi_W(x) \in W \quad \text{and} \quad |x - \pi_W(x)| = \text{dist}(x, W). \quad (6)$$

Clearly, if $x \notin W$, then $\pi_W(x) \in \partial W$. The next result is well known see e.g., [11, Proposition 3.1.3] or [16, Theorem 1.2.1].

Lemma 2.1. $\pi_W : \mathbb{R}^n \rightarrow W$ is 1-Lipschitz.

The *convex hull* of a set $A \subset \mathbb{R}^n$ (defined as the intersection of all convex sets containing A , or equivalently, as the set of all convex combinations of points of A) is denoted by $\text{co}(A)$. Every closed and convex set $W \subset \mathbb{R}^n$ is the intersection of all closed half-spaces that contain W . In fact, for every $x \in \partial W$ there is a half-space H_x such that $W \subset H_x$ and $x \in T_x \cap W$, where $T_x = \partial H_x$. The hyperplane T_x is called a *hyperplane supporting* W at x . Thus for every $x \in \partial W$, there is a hyperplane supporting W at x , but such a hyperplane is not necessarily unique. This implies that if $U \subset \mathbb{R}^n$ is open and convex and $f : U \rightarrow \mathbb{R}$ is convex, then for every $x \in U$ there is $v \in \mathbb{R}^n$ such that $f(y) \geq f(x) + \langle v, y - x \rangle$ for all $y \in U$. Indeed, on the right hand side we have an equation of the supporting hyperplane of the convex *epigraph* $\text{epi}(f) = \{(x, y) \in U \times \mathbb{R} : x \in U, y \geq f(x)\}$. The set of all such v is denoted by $\partial f(x)$ and called the *subdifferential* of f at x . Thus $\partial f(x) \neq \emptyset$ for any $x \in U$. If in addition f is differentiable at x_o , then $\partial f(x_o) = \{\nabla f(x_o)\}$ i.e., $f(y) \geq f(x_o) + Df(x_o)(y - x_o)$ meaning that the tangent hyperplane to the graph of f at x_o is the unique hyperplane supporting the epigraph of f at $(x_o, f(x_o))$. Convex functions are locally Lipschitz continuous and hence they are differentiable a.e. by the Rademacher theorem, so $\partial f(x) = \{\nabla f(x)\}$ for almost all $x \in U$. In fact we will prove the a.e. differentiability of convex functions directly and *without* any reference to the Rademacher theorem, see Corollary 3.6 and Remark 3.7, but we will need the Rademacher theorem in the proof

of Theorem 1.1, because we will need to know that the gradient of a convex function $g \in C^{1,1}$ is differentiable a.e.

3. PROOF OF THEOREM 1.4

We will precede the proof with auxiliary results.

For a convex body $K \subset \mathbb{R}^n$ and $r > 0$ we define the *inner parallel convex body* by

$$K_r := \{x \in K : \text{dist}(x, \partial K) \geq r\}.$$

Lemma 3.1. *K_r is convex for any $r > 0$.*

Proof. Let $x, y \in K_r$. We need to show that $[x, y] \subset K_r$. Clearly, $\bar{B}(x, r), \bar{B}(y, r) \subset K$ and for any $z \in [x, y]$, $\bar{B}(z, r) \subset \text{co}(\bar{B}(x, r) \cup \bar{B}(y, r)) \subset K$, so $\text{dist}(z, \partial K) \geq r$, $z \in K_r$, and hence $[x, y] \subset K_r$. $\square$

Let $r_o = \sup_{x \in K} \text{dist}(x, \partial K)$. Clearly $K_r = \emptyset$ for $r > r_o$. $K_{r_o} \neq \emptyset$, but it has empty interior. However, for $r \in (0, r_o)$, K_r has non-empty interior, so K_r is a convex body only for $r \in (0, r_o)$.

For a convex body K and $r > 0$ we also define

$$K(r) := \bigcup \{\bar{B}(x, r) : \bar{B}(x, r) \subset K\}. \quad (7)$$

It is easy to see that $K(r)$ is convex and compact (it can be empty). Moreover, if K contains a ball of radius r_o , then for any $r \in (0, r_o]$, $K(r)$ has non-empty interior and hence $K(r)$ is a convex body.

Lemma 3.2. *If a convex body K contains a ball of radius r_o , then for all $r \in (0, r_o)$, K_r is a convex body, and*

$$\mathcal{H}^{n-1}(\partial K_r) \leq \mathcal{H}^{n-1}(\partial K \cap \partial K(r)). \quad (8)$$

Proof. Clearly, for $r \in (0, r_o)$, K_r has non-empty interior, so it is a convex body by Lemma 3.1. Observe that (see (6))

$$\pi_{K_r}(\partial K \cap \partial K(r)) = \partial K_r. \quad (9)$$

Indeed, if $z \in \partial K_r$, then there is $x \in \partial K$, such that $|x - z| = r$. Therefore, $x \in \bar{B}(z, r) \subset K$, and hence $x \in K(r)$. Thus, $x \in \partial K \cap \partial K(r)$, $|x - z| = r \geq \text{dist}(x, K_r)$, and hence $z = \pi_{K_r}(x)$. Now, (8) follows from (9) and the fact that π_{K_r} is 1-Lipschitz (Lemma 2.1). $\square$

The next beautiful result is due to McMullen [15]. While it can be concluded from Alexandrov's theorem, we present here a direct and surprisingly elementary proof which is a small modification of McMullen's argument. In fact, Lemma 3.3 will play an important role in our proof of Alexandrov's theorem.

Lemma 3.3. *If $K \subset \mathbb{R}^n$ is a convex body, then $\lim_{r \rightarrow 0^+} \mathcal{H}^{n-1}(\partial K \setminus \partial K(r)) = 0$.*

Remark 3.4. Lemma 3.3 has the following geometric interpretation: for almost all $x \in \partial K$, there is a closed ball $\bar{B} \subset K$ touching the boundary of K at x , i.e., $x \in \bar{B}$.

Proof. Without loss of generality we may assume that $\bar{B}(0, r_o) \subset K$. If $r \in (0, r_o)$, then 0 belongs to the interior of K_r . For $\lambda > 0$ we define

$$\lambda K_r := \{\lambda z : z \in K_r\},$$

that is, λK_r is a dilation of K_r . For $r \in (0, r_o)$, let

$$\lambda(r) := \inf\{\lambda > 0 : K \subset \lambda K_r\}.$$

Clearly, $K \subset \lambda(r)K_r$. It is easy to see that the function $r \mapsto \lambda(r)$ is non-decreasing and $\lambda(r) \rightarrow 1$ as $r \rightarrow 0^+$. Indeed, for any $\varepsilon > 0$, $(1 + \varepsilon)^{-1}K \subset \text{int } K$, and hence $\delta := \text{dist}((1 + \varepsilon)^{-1}K, \partial K) > 0$, so for all $r \in (0, \delta]$

$$(1 + \varepsilon)^{-1}K \subset K_r, \quad \text{i.e.,} \quad K \subset (1 + \varepsilon)K_r.$$

In other words $1 \leq \lambda(r) \leq 1 + \varepsilon$ for all $0 < r \leq \delta$ proving that $\lambda(r) \rightarrow 1$ as $r \rightarrow 0^+$.

It is easy to see that $\pi_K(\partial(\lambda(r)K_r)) = \partial K$ (see (6)). Indeed, if $x \in \partial K$ and $\nu(x)$ is the outer unit normal vector to a supporting hyperplane of K at x , then there is $t \geq 0$ such that $z := x + t\nu(x) \in \partial(\lambda(r)K_r)$ and it easily follows that $\pi_K(z) = x$. Since π_K is 1-Lipschitz and it maps $\partial(\lambda(r)K_r)$ onto ∂K , we have that

$$\begin{aligned} \mathcal{H}^{n-1}(\partial K) &\leq \mathcal{H}^{n-1}(\partial(\lambda(r)K_r)) = \lambda(r)^{n-1} \mathcal{H}^{n-1}(\partial K_r) \leq \lambda(r)^{n-1} \mathcal{H}^{n-1}(\partial K \cap \partial K(r)) \\ &\leq \lambda(r)^{n-1} \mathcal{H}^{n-1}(\partial K) \rightarrow \mathcal{H}^{n-1}(\partial K) \quad \text{as } r \rightarrow 0^+. \end{aligned}$$

Therefore, $\mathcal{H}^{n-1}(\partial K \cap \partial K(r)) \rightarrow \mathcal{H}^{n-1}(\partial K)$, as $r \rightarrow 0^+$. This completes the proof of Lemma 3.3. $\square$

Lemma 3.5. *Let $f, g : B^n(0, R) \rightarrow \mathbb{R}$ be convex functions. If $g \in C^{1,1}$, $f \leq g$ and $f(x) = g(x)$ for some $x \in B^n(0, R)$, then f is differentiable at x , $Df(x) = Dg(x)$ and*

$$f(y) = f(x) + Df(x)(y - x) + O(|y - x|^2). \quad (10)$$

Proof. If $v \in \partial f(x)$, then clearly, $v \in \partial g(x)$ and hence $v = \nabla g(x)$. Therefore, the result follows from the estimate

$$f(x) + \langle \nabla g(x), y - x \rangle \leq f(y) \leq g(y) = f(x) + \langle \nabla g(x), y - x \rangle + O(|y - x|^2),$$

where in the last equality we used (5) and the fact that $g(x) = f(x)$. $\square$

Corollary 3.6. *If $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is convex, then it is differentiable a.e. Moreover*

$$f(y) = f(x) + Df(x)(y - x) + O(|y - x|^2) \quad \text{for almost all } x \in \mathbb{R}^n. \quad (11)$$

Proof. Since the boundary of a ball is parameterized by a smooth convex function, Lemma 3.5 implies (11) whenever there is a ball in the epigraph of f that touches the graph of f at $(x, f(x))$ and it follows from Lemma 3.3 that it is true for almost all x . $\square$

Remark 3.7. Note that the proof of Corollary 3.6 does not use Rademacher's theorem. Moreover, the estimate (11), is stronger than the a.e. differentiability of f that would follow from an application of Rademacher's theorem. We will not need Corollary 3.6 in this paper.

The following result, was proven in a more general form in the unpublished work [14, Theorem 1, p. 32]. It is also mentioned without any proof or reference in [13]. Although a detailed proof can be found in [10, Proposition 2.4.3], the origin of the result is not referenced in this work.

Lemma 3.8. *A convex body W has $C^{1,1}$ boundary if and only if there is $r > 0$ such that $W = W(r)$.*

Remark 3.9. In other words a convex body W has boundary of class $C^{1,1}$ if and only if there is $r > 0$ such that W is the union of closed balls of radius r .

We will only prove the implication from right to left, that is we will prove that if $W = W(r)$, then ∂W is of class $C^{1,1}$. This is the only implication that we need in the proof of Theorem 1.4. For the proof of the implication from left to right, see [10, Proposition 2.4.3]. We will present two proofs. The first proof is sketched only and it uses the implicit function theorem. The second one is detailed and it does not use the implicit function theorem.

First proof. It is very elementary and easy to prove that if $K \subset \mathbb{R}^n$ is compact, then the function $d_K^2(x) = \text{dist}(x, K)^2$ is differentiable and $\nabla d_K^2(x) = 2(x - \pi_K(x))$, see [11, p. 181]. Since the function π_K is Lipschitz by Lemma 2.1, we have that $d_K^2 \in C^{1,1}$ and all points in $\mathbb{R}^n \setminus K$ are regular so for $t > 0$, $\{x : \text{dist}(x, K) = t\} = (d_K^2)^{-1}(t^2)$ is a $C^{1,1}$ -submanifold of $\mathbb{R}^n$ and hence it is locally a graph of a $C^{1,1}$ function by the implicit function theorem. It remains to observe that $\partial W(r) = \{x \in \mathbb{R}^n : \text{dist}(x, W_r) = r\}$. $\square$

Second proof. Thus, we assume that for each $p \in \partial W$ there is $h(p) \in W$ such that $p \in \bar{B}(h(p), r) \subset W$. It follows that the hyperplane T_p tangent to the ball $\bar{B}(h(p), r)$ at p is the unique hyperplane supporting W at p .

Note that $\text{dist}(h(p), \partial W) = r$ implies that $h(p) \in W_r$, so $|p - h(p)| = r = \text{dist}(p, W_r)$, and hence $h(p) = \pi_{W_r}(p)$ (see (6)).

The inner unit normal vector to T_p is given by

$$\nu(p) = \frac{h(p) - p}{r} = \frac{\pi_{W_r}(p) - p}{r}$$

and Lemma 2.1 implies that the function $\nu : \partial W \rightarrow \mathbb{S}^{n-1}$ is Lipschitz continuous:

$$|\nu(p) - \nu(q)| \leq \frac{|\pi_{W_r}(p) - \pi_{W_r}(q)| + |p - q|}{r} \leq \frac{2}{r}|p - q|.$$

This in turn, implies that the boundary ∂W is of class $C^{1,1}$. Indeed, choose any point $p_o \in \partial W$ and choose a Euclidean coordinate system $(x_1, \dots, x_n) = (x', x_n)$ such that $p_o = 0$ and $T_{p_o} = \{x_n = 0\}$. Then ∂W in a neighborhood $U = B^{n-1}(0, \frac{r}{2})$ of $p_o = 0$ is a graph of a function $x_n = f(x')$ i.e., $p(x') := (x', f(x')) \in \partial W$. Since for $x' \in U$, the graph of f lies above $T_{p(x')}$ and below $\bar{B}(h(p(x')), r)$, it follows from geometric considerations (as in the proof of Corollary 3.6) that f is differentiable at x' and $T_{p(x')}$ is the tangent hyperplane to the graph of f at $p(x')$. Note also that $|\nabla f| \leq M$ on U for some $M > 0$, because the tangent hyperplane to the graph of f cannot intersect with $B(h(0), r)$. It remains to show that ∇f is Lipschitz continuous in U .

The inner unit normal vector in terms of ∇f is given by

$$\nu(p(x')) = \frac{(-\nabla f(x'), 1)}{\sqrt{1 + |\nabla f(x')|^2}}, \quad \text{so} \quad \pi(\nu(p(x'))) = \frac{-\nabla f(x')}{\sqrt{1 + |\nabla f(x')|^2}},$$

where $\pi : \mathbb{R}^n \rightarrow \mathbb{R}^{n-1}$, $\pi(x', x_n) = x'$ is the orthogonal projection. Since $\Psi(\Phi(z)) = z$ for all $z \in \mathbb{R}^{n-1}$, where $\Psi(z) = -z/\sqrt{1 + |z|^2}$ and $\Phi(z) = -z/\sqrt{1 + |z|^2}$, it follows that

$$\nabla f(x') = \Psi\left(\frac{-\nabla f(x')}{\sqrt{1 + |\nabla f(x')|^2}}\right) = \Psi(\pi(\nu(x', f(x')))) \quad \text{for } x' \in U.$$

This proves Lipschitz continuity of ∇f in U , as a composition of Lipschitz functions. The only issue could be the Lipschitz continuity of Ψ : it is a smooth function defined for $|z| < 1$, but it is

unbounded. However, this does not cause any problems here, because

$$\left| \frac{-\nabla f(x')}{\sqrt{1 + |\nabla f(x')|^2}} \right| \leq \frac{M}{\sqrt{1 + M^2}} < 1.$$

□

Proof of Theorem 1.4. Let $K \subset \mathbb{R}^n$ be a convex body. According to Lemma 3.3, for every $\varepsilon > 0$ there is $\delta_o > 0$ such that for any $\delta \in (0, \delta_o)$, $\mathcal{H}^{n-1}(\partial K \setminus \partial K(\delta)) < \varepsilon/2$. Since $K(\delta) \subset K$, it is easy to see that $\pi_{K(\delta)}(\partial K) = \partial K(\delta)$, so $\pi_{K(\delta)}(\partial K \setminus \partial K(\delta)) = \partial K(\delta) \setminus \partial K$ and Lemma 2.1 yields $\mathcal{H}^{n-1}(\partial K(\delta) \setminus \partial K) \leq \mathcal{H}^{n-1}(\partial K \setminus \partial K(\delta))$. Therefore, $\mathcal{H}^{n-1}(\partial K \triangle \partial K(\delta)) < \varepsilon$. Since the boundary of $K(\delta)$ is of class $C^{1,1}$ by Lemma 3.8, $W := K(\delta)$ satisfies the claim of the theorem. □

The next result is a direct consequence of Theorem 1.4 and it is a version of Theorem 1.3. We will use Corollary 3.10 in the proofs of Theorems 1.1, 1.3, and 1.5.

Corollary 3.10. *Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a convex function. Then for every $R > 0$ and $\varepsilon > 0$, there is a convex function $g \in C^{1,1}(B^n(0, R))$ such that $g \geq f$ and*

$$|\{x \in B^n(0, R) : f(x) \neq g(x)\}| < \varepsilon. \quad (12)$$

Proof. Let $M := \sup_{\bar{B}^n(0, 2R)} f(x)$ and define

$$W := \{(x, y) \in \bar{B}^n(0, 2R) \times \mathbb{R} : f(x) \leq y \leq M + 2R\}.$$

That is, W is an $(n + 1)$ -dimensional convex body bounded by the graph of f , the cylinder $\partial B^n(0, 2R) \times \mathbb{R}$ and the hyperplane $y = M + 2R$. According to Lemma 3.3, there is $\delta < R$ such that

$$\mathcal{H}^n(\partial W \setminus \partial W(\delta)) < \varepsilon.$$

Since $W(\delta)$ is the union of closed balls of radius $\delta < R$ that are contained in W , it follows that

$$\bar{B}^n(0, 2R) \times \{M + R\} \subset W(\delta),$$

i.e., the intersection of $W(\delta)$ with the hyperplane $y = M + R$ is an n -dimensional closed ball of radius $2R$. Thus, if $\pi : \mathbb{R}^{n+1} \rightarrow \mathbb{R}^n$ is the orthogonal projection, $\pi(W(\delta)) = \bar{B}^n(0, 2R)$, and hence for $x \in \bar{B}^n(0, 2R)$, we can define

$$g(x) := \inf\{y : (x, y) \in W(\delta)\}.$$

That is, the function $g : \bar{B}^n(0, 2R) \rightarrow \mathbb{R}$ parametrizes the bottom part of the boundary of $W(\delta)$. According to Lemma 3.8, the boundary of $W(\delta)$ is of class $C^{1,1}$ so $g \in C_{\text{loc}}^{1,1}(B^n(0, 2R))$ and hence g is a convex function in $C^{1,1}(B^n(0, R))$. Since $W(\delta)$ is contained in W and hence in the epigraph of f , it follows that $g \geq f$.

Observe that

$$\{x \in B^n(0, R) : f(x) \neq g(x)\} \subset \pi(\partial W \setminus \partial W(\delta))$$

and hence

$$|\{x \in B^n(0, R) : f(x) \neq g(x)\}| \leq |\pi(\partial W \setminus \partial W(\delta))| \leq \mathcal{H}^n(\partial W \setminus \partial W(\delta)) < \varepsilon,$$

because the orthogonal projection does not increase the Hausdorff measure and $\mathcal{H}^n$ coincides with the Lebesgue measure in $\mathbb{R}^n$. □

4. PROOF OF THEOREM 1.1

Lemma 4.1. *Suppose that $f, g : B^n(0, R) \rightarrow \mathbb{R}$ are convex, $f \leq g$ and $g \in C^{1,1}(B^n(0, R))$. Then for almost all $x_o \in \{f = g\}$ we have*

$$f(x) = f(x_o) + Df(x_o)(x - x_o) + \frac{1}{2}(x - x_o)^T D^2g(x_o)(x - x_o) + o(|x - x_o|^2). \quad (13)$$

Remark 4.2. Note that $D^2g(x_o)$ in (13) is not a typo. Also, we do not need the assumption that g is convex or $C^{1,1}$. With a small modification, the proof works under the assumption that $f \leq g \in C^1$ and Dg is differentiable at x_o .

Proof. It follows from Lemma 3.5 that f is differentiable at every point of the set $\{f = g\}$ and that $Df = Dg$ in $\{f = g\}$. Since Dg is Lipschitz continuous, Dg is differentiable a.e. by Rademacher's theorem. Therefore, it suffices to prove the result whenever $x_o \in \{f = g\}$ is a density point of that set and Dg is differentiable at x_o .

To simplify notation, without loss of generality, we may assume that $x_o = 0$, and we need to prove that

$$f(x) - f(0) - Df(0)x - \frac{1}{2}x^T D^2g(0)x = o(|x|^2).$$

Since $f(0) = g(0)$ and $Df(0) = Dg(0)$, the left hand side equals

$$(f(x) - g(x)) + \left(g(x) - g(0) - Dg(0)x - \frac{1}{2}x^T D^2g(0)x \right) = (f(x) - g(x)) + o(|x|^2).$$

We used here the fact that g is twice differentiable at 0 (Taylor's theorem with the Peano remainder). Thus it remains to show that $g(x) - f(x) = o(|x|^2)$.

Since 0 is a density point of the set $\{f = g\}$, for any x we can find $y \in \{f = g\}$ such that $|x - y| = o(|x|)$. For if not, there is $\varepsilon > 0$ and $x_k \rightarrow 0$ such that $B(x_k, \varepsilon|x_k|) \cap \{f = g\} = \emptyset$ and that contradicts the fact that 0 is a density point of $\{f = g\}$.

Clearly, $f(y) = g(y)$ and $Df(y) = Dg(y)$ by Lemma 3.5. Therefore,

$$f(x) \geq f(y) + Df(y)(x - y) = g(y) + Dg(y)(x - y),$$

where the inequality is a consequence of convexity of f . Since $f \leq g$, the above inequality and (4) yield

$$0 \leq g(x) - f(x) \leq g(x) - g(y) - Dg(y)(x - y) \leq M|x - y|^2 = o(|x|^2).$$

The proof is complete. $\square$

Proof of Theorem 1.1. Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be convex. Let $R > 0$ and $\varepsilon > 0$ and let g be as in Corollary 3.10. It follows from Lemma 4.1 that for almost all $x \in \{f = g\}$, (1) is satisfied with $D^2f(x) := D^2g(x)$. Hence (1) holds true in $B(0, R)$ outside a set of measure less than ε . Since it is true for any $R > 0$ and $\varepsilon > 0$, it follows that (1) is satisfied almost everywhere. $\square$

5. PROOF OF THEOREM 1.2

If f is twice differentiable at 0 as in (1), then we have

$$f(x) = f(0) + Df(0)x + \frac{1}{2}x^T D^2f(0)x + R(x) = f(0) + Df(0)x + \langle Ax, x \rangle + R(x),$$

where $A = \frac{1}{2}D^2f(0)$ and $R(x) = o(|x|^2)$. Note that

$$a(r) := \sup_{0 < |x| \leq 2r} \frac{|R(x)|}{|x|^2} \rightarrow 0 \quad \text{as } r \rightarrow 0^+.$$

Moreover,

$$|R(x)| \leq a\left(\frac{|x|}{2}\right) |x|^2 \leq a(|x|) |x|^2.$$

Proof of Theorem 1.2. Let f be twice differentiable at x as in (1). We need to prove (2). Without loss of generality we may assume that $x = 0$, and hence we need to prove that

$$\lim_{x \rightarrow 0} \frac{\sigma_x - Df(0) - D^2f(0)x}{|x|} = 0 \quad \text{for any } \sigma_x \in \partial f(x).$$

For $x, y \neq 0$, we have

$$f(x) = f(0) + Df(0)x + \langle Ax, x \rangle + R(x), \quad f(y) = f(0) + Df(0)y + \langle Ay, y \rangle + R(y).$$

Since $f(x) + \langle \sigma_x, y - x \rangle \leq f(y)$, we have

$$\langle \sigma_x, y - x \rangle \leq f(y) - f(x) = Df(0)(y - x) + \langle A(x + y), y - x \rangle + R(y) - R(x).$$

We used here the fact that A is symmetric and hence $\langle Ax, y \rangle = \langle Ay, x \rangle$. Let

$$y = x + w, \quad \text{where } w = \sqrt{a(|x|)} |x|z, \quad |z| = 1.$$

Then

$$\begin{aligned} \langle \sigma_x, w \rangle &\leq Df(0)w + \langle A(2x + w), w \rangle + R(y) - R(x), \\ \langle \sigma_x - Df(0) - 2Ax, w \rangle &\leq \langle Aw, w \rangle + R(y) - R(x). \end{aligned}$$

If $|x|$ is sufficiently small, then $a(|x|) \leq 1$ and hence $|w| \leq |x|$, so $|y| \leq 2|x|$. Therefore,

$$|R(y)| \leq a\left(\frac{|y|}{2}\right) |y|^2 \leq 4a(|x|) |x|^2, \quad |R(y) - R(x)| \leq 5a(|x|) |x|^2.$$

Taking the supremum over all z with $|z| = 1$ we get

$$|\sigma_x - Df(0) - 2Ax| \sqrt{a(|x|)} |x| \leq |A|a(|x|) |x|^2 + 5a(|x|) |x|^2,$$

and hence

$$\frac{|\sigma_x - Df(0) - 2Ax|}{|x|} \leq (|A| + 5) \sqrt{a(|x|)} \rightarrow 0 \quad \text{as } x \rightarrow 0.$$

Since $2A = D^2f(0)$, the result follows. $\square$

6. PROOF OF THEOREM 1.3

One of the differences between Corollary 3.10 and Theorem 1.3 is that the function g in Corollary 3.10 is defined on the ball $B^n(0, R)$ only and the main step in the proof of Theorem 1.3 will be to show that the function g can be extended from a ball $B^n(0, R - \delta)$ to a convex function of class $C^{1,1}(\mathbb{R}^n)$. We will do it by gluing the function g with a quadratic function of the form $a|x|^2 - b$ and we need to know how to glue convex functions while maintaining their smoothness.

The maximum of two convex functions

$$\max\{u, v\} = \frac{u + v + |u - v|}{2}$$

is convex, but even if $u, v \in C^\infty$, the maximum $\max\{u, v\}$ need not be C^1 . To overcome this difficulty, we will use the so called smooth maximum that was introduced in [2].

Let $\theta \in C^\infty(\mathbb{R})$ be such that $\theta(t) = |t|$ if and only if $|t| \geq 1$, θ is convex, $\theta(t) = \theta(-t)$ for all t , and 1-Lipschitz.

It easily follows that $\theta(t) > 0$ for all t and $|\theta'(t)| < 1$ if and only if $|t| < 1$. Then, we define the *smooth maximum* function $\mathcal{M} : \mathbb{R}^2 \rightarrow \mathbb{R}$ as,

$$\mathcal{M}(x, y) := \frac{x + y + \theta(x - y)}{2}.$$

It is easy to see that $\mathcal{M}$ is smooth, convex and

$$\mathcal{M}(x, y) = \max\{x, y\} \quad \text{whenever} \quad |x - y| \geq 1. \quad (14)$$

It is also not difficult to prove that $\mathcal{M}(x, y)$ is non-decreasing in x and y , because partial derivatives of $\mathcal{M}$ are non-negative, see [2, Lemma 2.1(viii)]. This observation and convexity of $\mathcal{M}$ yield (see [2, Proposition 2.2(i)])

Lemma 6.1. *If $u, v : U \rightarrow \mathbb{R}$ are convex functions defined in an open convex set $U \subset \mathbb{R}^n$, then $\mathcal{M}(u, v) : U \rightarrow \mathbb{R}$ is convex.*

It is also obvious that if $u, v \in C_{\text{loc}}^{1,1}(U)$, then $\mathcal{M}(u, v) \in C_{\text{loc}}^{1,1}(U)$.

We will use the smooth maximum to prove the following extension result.

Proposition 6.2. *Let $h \in C_{\text{loc}}^{1,1}(B^n(0, R))$ be a convex function. Then, for every $r \in (0, R)$, there is a convex function $H \in C^{1,1}(\mathbb{R}^n)$, such that*

$$H(x) = h(x) \quad \text{whenever} \quad |x| \leq r. \quad (15)$$

Remark 6.3. If $h \in C^k$, $k \in \mathbb{N} \cup \{\infty\}$, then $H \in C^k(\mathbb{R}^n)$. The proof remains the same.

Proof. Choose $\rho \in (r, R)$ and let

$$m := \inf_{|x| \leq r} h, \quad M := \sup_{|x| = \rho} h.$$

Then, we can find $a, b > 0$ such that the function $q(x) := a|x|^2 - b$ satisfies

$$q(x) < m - 1 \quad \text{if} \quad |x| \leq r \quad (16)$$

$$q(x) > M + 1 \quad \text{if} \quad |x| = \rho, \quad (17)$$

and we define

$$H(x) := \begin{cases} \mathcal{M}(h(x), q(x)) & \text{if } |x| \leq \rho, \\ q(x) & \text{if } |x| > \rho. \end{cases}$$

It follows from (16) that $h(x) > q(x) + 1$ if $|x| \leq r$, so by (14), we have $H(x) = \mathcal{M}(h(x), q(x)) = h(x)$ if $|x| \leq r$ and the condition (15) is satisfied. It follows from (17) that there is $\varepsilon > 0$ such that $q(x) > h(x) + 1$ if $\rho \leq |x| \leq \rho + \varepsilon$ and hence by (14), $\mathcal{M}(h(x), q(x)) = q(x)$ when $\rho \leq |x| \leq \rho + \varepsilon$. Therefore, the convex functions $q(x) \in C^{1,1}(\mathbb{R}^n)$ and $\mathcal{M}(h(x), q(x)) \in C_{\text{loc}}^{1,1}(B^n(0, R))$ coincide in the annulus $\rho \leq |x| \leq \rho + \varepsilon$ and hence H is convex in $\mathbb{R}^n$ with $H \in C_{\text{loc}}^{1,1}(\mathbb{R}^n)$. Since $H = q \in C^{1,1}$ outside the compact ball $\bar{B}^n(0, \rho)$, it follows that $H \in C^{1,1}(\mathbb{R}^n)$. $\square$

Proof of Theorem 1.3. Let $R > 0$ be such that $|A \setminus \bar{B}^n(0, R)| < \varepsilon/2$. According to Corollary 3.10 there is a convex function $\tilde{g} \in C^{1,1}(B^n(0, 2R))$ such that

$$|\{x \in B^n(0, 2R) : f(x) \neq \tilde{g}(x)\}| < \frac{\varepsilon}{2}.$$

Now, Proposition 6.2 yields a convex function $g \in C^{1,1}(\mathbb{R}^n)$ such that $g(x) = \tilde{g}(x)$ for $x \in \bar{B}^n(0, R)$ and we have

$$|\{x \in A : f(x) \neq g(x)\}| \leq |A \setminus \bar{B}^n(0, R)| + |\{x \in B^n(0, R) : f(x) \neq \tilde{g}(x)\}| < \varepsilon.$$

□

7. PROOFS OF THEOREMS 1.5, 1.6 AND 1.8

This section is less self-contained than the others. The proof of the implication (1) $\Rightarrow$ (2) in Theorem 1.5 is easy and we will not show it here; see [5, Proposition 1.10 and Theorem 2.5]. The implication (2) $\Rightarrow$ (1) in Theorem 1.5 is equivalent to the following result.

Theorem 7.1. *Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a convex function such that $\lim_{|x| \rightarrow \infty} f(x) = +\infty$. Then for every $\varepsilon > 0$ there exists a convex function $g : \mathbb{R}^n \rightarrow \mathbb{R}$ of class $C_{\text{loc}}^{1,1}(\mathbb{R}^n)$ such that $g \geq f$ and $|\{x \in \mathbb{R}^n : f(x) \neq g(x)\}| < \varepsilon$.*

Next, we give a proof of Theorem 7.1 that greatly simplifies the one provided by [5]. Its main ingredients are Corollary 3.10 above and the following lemma (whose elementary proof can be found in [3, Lemma 5.3], which in turn is a refinement of the result of [12]).

Lemma 7.2. *Let $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ be a continuous function such that $\lim_{|x| \rightarrow \infty} \varphi(x) = +\infty$ and such that for every $R > 0$ there exists $C_R > 0$ so that for every $x, h \in B^n(0, R)$ we have*

$$\varphi(x+h) + \varphi(x-h) - 2\varphi(x) \leq C_R |h|^2.$$

Then the function $F = \text{conv}(\varphi)$ has a similar property: for every $R > 0$ there exists $C'_R > 0$ such that for every $x, h \in B^n(0, R)$ we have

$$F(x+h) + F(x-h) - 2F(x) \leq C'_R |h|^2.$$

Therefore $F \in C_{\text{loc}}^{1,1}(\mathbb{R}^n)$.

Here $\text{conv}(\varphi)$ denotes the convex envelope of φ , defined as the supremum of all convex functions less than or equal to φ .

Proof of Theorem 7.1. By Corollary 3.10, for every $k \in \mathbb{N}$ we can find a convex function $g_k \in C^{1,1}(B^n(0, 2k))$ such that $f \leq g_k$ and

$$|\{x \in B^n(0, 2k) : f(x) \neq g_k(x)\}| < \varepsilon/2^k.$$

For every $k \in \mathbb{N}$, let $\theta_k : (k-2, k+1) \rightarrow [0, \infty)$ be a C^∞ convex function such that:

- (1) $\theta_k(t) = 0$ iff $k-1 \leq t \leq k$;
- (2) $\lim_{t \rightarrow (k-2)^+} \theta_k(t) = +\infty$, and
- (3) $\lim_{t \rightarrow (k+1)^-} \theta_k(t) = +\infty$.

Define $\varphi_k : \mathbb{R}^n \rightarrow (-\infty, +\infty]$ by

$$\varphi_k(x) = g_k(x) + \theta_k(|x|) \text{ if } k-2 < |x| < k+1, \text{ and } \varphi_k(x) = +\infty \text{ otherwise.}$$

When $n \geq 2$ and $k \geq 2$, the function φ_k is not convex, but we do not need it to be.

Note that $\varphi_k(x) = g_k(x)$ on the annulus $A_k := \{x : k-1 \leq |x| \leq k\}$ (or ball in the special case $k = 1$), and consider $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ defined by $\varphi(x) = \inf_{k \in \mathbb{N}} \varphi_k(x)$. It is clear that

$$f \leq \varphi \text{ on } \mathbb{R}^n, \text{ and } \varphi \leq g_k \text{ on } A_k, \text{ for each } k \in \mathbb{N},$$

and in particular $\lim_{|x| \rightarrow \infty} \varphi(x) = +\infty$, though φ is finite everywhere. As a matter of fact, it is easily seen that φ is locally the minimum of at most three continuous functions, and therefore it is continuous on $\mathbb{R}^n$. More precisely, for each $k \in \mathbb{N}$ we have that

$$\varphi(x) = \min\{\varphi_{k-1}(x), \varphi_k(x), \varphi_{k+1}(x)\} \text{ for every } x \in A_k.$$

Moreover, since $\lim_{|x| \rightarrow k-} \varphi_{k-1}(x) = +\infty = \lim_{|x| \rightarrow k+} \varphi_{k+2}(x)$ and φ_k and φ_{k+1} are bounded and $C^{1,1}$ on a neighborhood of the sphere $\{x : |x| = k\}$, there exist some $M_k, \delta_k > 0$ such that $\varphi(x) = \min\{\varphi_k(x), \varphi_{k+1}(x)\}$ and $\varphi_j(x+h) + \varphi_j(x-h) - 2\varphi_j(x) \leq M_k|h|^2$ for all $k - \delta_k \leq |x| \leq k + \delta_k$, $|h| \leq \delta_k$, and $j = k, k+1$. These inequalities easily follow from (4). This implies that

$$\varphi(x+h) + \varphi(x-h) - 2\varphi(x) \leq M_k|h|^2$$

for all $k - \delta_k \leq |x| \leq k + \delta_k$, and $|h| \leq \delta_k$. Similarly, there exist $M'_k, \delta'_k > 0$ such that $\varphi(x) = \min\{\varphi_{k-1}(x), \varphi_k(x), \varphi_{k+1}(x)\}$ and $\varphi_j(x+h) + \varphi_j(x-h) - 2\varphi_j(x) \leq M'_k|h|^2$ for all $k-1+\delta_{k-1}-\delta'_k \leq |x| \leq k - \delta_k + \delta'_k$, $|h| \leq \delta'_k$, and $j = k-1, k, k+1$, implying that

$$\varphi(x+h) + \varphi(x-h) - 2\varphi(x) \leq M'_k|h|^2$$

for all $k-1+\delta_{k-1}-\delta'_k \leq |x| \leq k - \delta_k + \delta'_k$, and $|h| \leq \delta'_k$. Since every ball is contained in a finite union of sets A_k , these estimates imply that for every $R > 0$ there exist $C_R > 0$ and $\delta_R > 0$ so that for every $x \in B^n(0, R)$ and $|h| < \delta_R$ we have

$$E_h(x) := \varphi(x+h) + \varphi(x-h) - 2\varphi(x) \leq C_R|h|^2. \quad (18)$$

On the other hand, for $\delta_R \leq |h| \leq R$, we obviously have $E_h(x) \leq 4M \leq \widetilde{C}_R|h|^2$, where $M := \sup_{z \in B(0, 2R)} \varphi(z)$ and $\widetilde{C}_R = 4M/\delta_R^2$. So by replacing C_R with $\max\{C_R, \widetilde{C}_R\}$ we certainly have

$$\varphi(x+h) + \varphi(x-h) - 2\varphi(x) \leq C_R|h|^2 \quad \text{for all } x, h \in B^n(0, R). \quad (19)$$

Therefore, (19) and Lemma 7.2, imply that the function $g := \text{conv}(\varphi)$ is of class $C_{\text{loc}}^{1,1}$ (and it obviously satisfies $f \leq g \leq \varphi$). Since $|\{x \in B^n(0, 2k) : f(x) \neq g_k(x)\}| < \varepsilon/2^k$ and $f \leq g \leq \varphi \leq g_k$ on A_k , it follows that $|\{x \in A_k : f(x) \neq g(x)\}| < \varepsilon/2^k$ for every $k \in \mathbb{N}$, which implies that $|\{x \in \mathbb{R}^n : f(x) \neq g(x)\}| \leq \varepsilon$. $\square$

Proof of Theorem 1.6. For the proof of the implication (1) $\Rightarrow$ (2), see [5, Corollary 1.13]. Regarding the implication (2) $\Rightarrow$ (1), the same proof as in [5, Corollary 1.13] gives us a (possibly unbounded) convex body W_ε with boundary S_ε such that $W_\varepsilon = \frac{1}{t_0}g^{-1}(-\infty, t_0]$ for some $t_0 \in (1, 2)$ and some convex function $g \in C_{\text{loc}}^{1,1}(\mathbb{R}^n)$ such that $\mathcal{H}^{n-1}(S \setminus S_\varepsilon) < \varepsilon/2$, $S = \partial W$, and $\mu \leq g$, where μ is the Minkowski functional of W , hence $W_\varepsilon \subset W$.

Now it suffices to show that $\pi_{W_\varepsilon}(S) = S_\varepsilon$ ¹, where π_{W_ε} is the nearest point projection defined in (6), because this fact and Lemma 2.1 will imply $\mathcal{H}^{n-1}(S_\varepsilon \setminus S) \leq \mathcal{H}^{n-1}(S \setminus S_\varepsilon)$, and hence $\mathcal{H}^{n-1}(S_\varepsilon \triangle S) < \varepsilon$.

Therefore, it remains to show that if $x \in S_\varepsilon$, then there is $z \in S$ such that $\pi_{W_\varepsilon}(z) = x$. Let $\nu(x)$ be the unit outward normal to S_ε at x . It suffices to show that the ray $R_x := \{x + t\nu(x) : t \geq 0\}$ intersects S at some point z , because clearly, $\pi_{W_\varepsilon}(z) = x$. Suppose to the contrary that R_x does not intersect with S i.e. $R_x \subset \text{int } W$. The tangent hyperplane to S_ε at x is defined by $T_x := \{x + v : \langle v, \nu(x) \rangle = 0\}$ and clearly $S_\varepsilon \cap F_x = \emptyset$, where $F_x := \{x + v : \langle v, \nu(x) \rangle > 0\}$ is an open half-space bounded by T_x . Since $x \in \text{int } W$, there is $\delta > 0$ such that $D_{2\delta} \subset \text{int } W$, where $D_{2\delta} := \{x + v \in T_x : |v| < 2\delta\}$ is the ball in T_x centered at x and of radius 2δ . Since $D_{2\delta} \cup R_x \subset \text{int } W$, it follows from the convexity of W that $C_x \subset \text{int } W$, where $C_x := \{p + t\nu(x) : p \in \partial D_\delta, t > 0\}$ is the side surface of a half-cylinder.

¹This is very easy if W is bounded and we used this fact in the proof of Theorem 1.4.

Since S does not contain any line, it follows that W does not contain any line (cf. the argument at the beginning of the proof of Theorem 1.8). Therefore, for any $p \in R_x$, and any unit vector v parallel to T_x , the lines $L_{p,v} := \{p + tv : t \in \mathbb{R}\}$ must intersect S at least at one point. Denote all such points in S by A . Since $A \subset F_x$, $F_x \cap S_\varepsilon = \emptyset$, we have that $A \subset S \setminus S_\varepsilon$. It is easy to see that $\mathcal{H}^{n-1}(A) = \infty$ and hence $\mathcal{H}^{n-1}(S \setminus S_\varepsilon) = \infty$ which is a contradiction. To show that $\mathcal{H}^{n-1}(A) = \infty$, note that the radial projection π of A onto C_x along lines $L_{p,v}$ is 1-Lipschitz and hence $\mathcal{H}^{n-1}(A) \geq \mathcal{H}^{n-1}(\pi(A))$. Now, for each two antipodal points in each sphere of radius δ in C_x that is parallel to T_x , at least one belongs to $\pi(A)$. The mapping $\Phi : C_x \rightarrow C_x$ that maps points in C_x to antipodal points is an isometry of C_x . Hence $\mathcal{H}^{n-1}(\pi(A)) = \mathcal{H}^{n-1}(\Phi(\pi(A)))$. Therefore,

$$2\mathcal{H}^{n-1}(A) \geq 2\mathcal{H}^{n-1}(\pi(A)) \geq \mathcal{H}^{n-1}(\pi(A) \cup \Phi(\pi(A))) = \mathcal{H}^{n-1}(C_x) = \infty.$$

The proof is complete. $\square$

Proof of Theorem 1.8. If an unbounded convex body V contains a line L , then since V is closed, it is easy to see that V contains all lines parallel to L that intersect with V . In particular ∂V is the union of lines parallel to L .

(2) $\Rightarrow$ (1). Let us define V_f as the closure of the epigraph of f , then V_f is an unbounded convex body in $\mathbb{R}^{n+1}$. Since the graph of f does not contain a line, ∂V_f does not contain any line and we may apply Theorem 1.6 and Remark 1.7 to find a $C_{\text{loc}}^{1,1}$ convex body $W \subseteq V_f$ such that $\mathcal{H}^n(\partial V_f \setminus \partial W) < \varepsilon$. Since the projection $\pi : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^n$ is 1-Lipschitz, it follows that $g(x) := \inf\{y : (x, y) \in W\}$ is a $C_{\text{loc}}^{1,1}$ convex function such that $g \geq f$, $|\{x \in U : g(x) \neq f(x)\}| < \varepsilon$. Note that g is finite on all of U . Indeed, let $A := \{x \in U : g(x) < \infty\}$; we understand that $g(x) = \infty$ if the line $\{(x, t) : t \in \mathbb{R}\}$ does not intersect W . The set A is convex because W is convex. If $A \neq U$ then, for some $x_0 \in U$, some number c and some linear function $\ell : \mathbb{R}^n \rightarrow \mathbb{R}$, we have $\ell(x_0) \geq c \geq \ell(x)$ for all $x \in A$, implying that for all $x \in U \cap \ell^{-1}(c, \infty)$ the vertical line $\{(x, t) : t \in \mathbb{R}\}$ does not intersect W . But it is easy to see that the set $\partial V_f \cap \{(x, t) : x \in \bar{U} \cap \ell^{-1}(c, \infty)\}$ has infinite Hausdorff n -dimensional measure, therefore it must intersect ∂W . Hence $g(x) = f(x) < \infty$ for some $x \in U \cap \ell^{-1}(c, \infty)$, a contradiction.

(1) $\Rightarrow$ (2). Suppose to the contrary that f satisfies (1) and that the graph of f contains a line L . Then, as we observed above, the graph of f is the union of lines parallel to L . Thus, U is the union of lines parallel to $\pi(L)$ and clearly, f is affine on each such a line. Now, an argument similar to the proof of [5, Proposition 1.10] yields that if $g : U \rightarrow \mathbb{R}$ is a convex function such that $|\{x : f(x) \neq g(x)\}| < \infty$, then $g = f$. Hence $g \notin C_{\text{loc}}^{1,1}$, because $f \notin C_{\text{loc}}^{1,1}$ and we arrive to a contradiction with (1). $\square$

REFERENCES

- [1] ALEXANDROFF, A.D.: Almost everywhere existence of the second differential of a convex function and some properties of convex surfaces connected with it. (Russian) Leningrad State Univ. Annals [Uchenye Zapiski] Math. Ser. 6, (1939), 3–35.
- [2] AZAGRA, D.: Global and fine approximation of convex functions. *Proc. Lond. Math. Soc.* (3) 107 (2013), no. 4, 799–824.
- [3] AZAGRA, D.: Locally $C^{1,1}$ convex extensions of 1-jets. *Rev. Mat. Iberoam.* 38 (2022), no. 1, 131–174.
- [4] AZAGRA, D., DRAKE, M., HAJŁASZ, P.: C^2 -Lusin approximation of strongly convex functions. Preprint 2023.
- [5] AZAGRA, D., HAJŁASZ, P.: Lusin-type properties of convex functions and convex bodies. *J. Geom. Anal.* 31 (2021), 11685–11701.
- [6] AZAGRA, D., LE GRUYER, E., MUDARRA, C.: Explicit formulas for $C^{1,1}$ and $C_{\text{cov}}^{1,\omega}$ extensions of 1-jets in Hilbert and superreflexive spaces. *J. Funct. Anal.* 274 (2018), 3003–3032.

- [7] AZAGRA, D., MUDARRA, C.: Global geometry and C^1 convex extensions of 1-jets. *Anal. PDE* 12 (2019), 1065–1099.
- [8] AZAGRA, D., MUDARRA, C.: Whitney extension theorems for convex functions of the classes C^1 and $C^{1,\omega}$. *Proc. Lond. Math. Soc.* 114 (2017), 133–158.
- [9] BIANCHI, G., COLESANTI, A., PUCCI, C.: On the second differentiability of convex surfaces. *Geom. Dedicata* 60 (1996), 39–48.
- [10] HÖRMANDER, L.: *Notions of Convexity*. Progress in Mathematics Series, 127 Birkhäuser Boston, Boston, M.A. 1994
- [11] HIRIART-URRUTY, J.-B., LEMARÉCHAL, C.: *Fundamentals of convex analysis*. Grundlehren Text Editions. Springer-Verlag, Berlin, 2001.
- [12] KIRCHHEIM, B., KRISTENSEN, J.: Differentiability of Convex Envelopes. *C. R. Acad. Sci. Paris Sér. I Math.* 333(8), 725 - 728 (2001), 11685-11701.
- [13] KISELMAN, C. O.: Regularity classes for operations in convexity theory. *Kodai Math. J.* 15 (1992), 354–374.
- [14] LUCAS, K.: Submanifolds of dimension $n - 1$ in E^n with normals satisfying a Lipschitz condition. Thesis (Ph.D.)–University of Kansas. 1957. Studies in Eigenvalue Problems, Technical Report 18, Dept of Mathematics, University of Kansas, 1957, Sect. 2.
- [15] McMULLEN, P.: On the inner parallel body of a convex body. *Israel J. Math.* 19 (1974), 217–219.
- [16] SCHNEIDER, R.: *Convex bodies: the Brunn-Minkowski theory*. Second expanded edition. Encyclopedia of Mathematics and its Applications, 151. Cambridge University Press, Cambridge, 2014.