

# LUSIN-TYPE PROPERTIES OF CONVEX FUNCTIONS AND CONVEX BODIES

DANIEL AZAGRA AND PIOTR HAJLASZ

ABSTRACT. We prove that if  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is convex and  $A \subset \mathbb{R}^n$  has finite measure, then for any  $\varepsilon > 0$  there is a convex function  $g : \mathbb{R}^n \rightarrow \mathbb{R}$  of class  $C^{1,1}$  such that  $\mathcal{L}^n(\{x \in A : f(x) \neq g(x)\}) < \varepsilon$ . As an application we deduce that if  $W \subset \mathbb{R}^n$  is a compact convex body then, for every  $\varepsilon > 0$ , there exists a convex body  $W_\varepsilon$  of class  $C^{1,1}$  such that  $\mathcal{H}^{n-1}(\partial W \setminus \partial W_\varepsilon) < \varepsilon$ . We also show that if  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is a convex function and  $f$  is not of class  $C_{\text{loc}}^{1,1}$ , then for any  $\varepsilon > 0$  there is a convex function  $g : \mathbb{R}^n \rightarrow \mathbb{R}$  of class  $C_{\text{loc}}^{1,1}$  such that  $\mathcal{L}^n(\{x \in \mathbb{R}^n : f(x) \neq g(x)\}) < \varepsilon$  if and only if  $f$  is essentially coercive, meaning that  $\lim_{|x| \rightarrow \infty} f(x) - \ell(x) = \infty$  for some linear function  $\ell$ . A consequence of this result is that, if  $S$  is the boundary of some convex set with nonempty interior (not necessarily bounded) in  $\mathbb{R}^n$  and  $S$  does not contain any line, then for every  $\varepsilon > 0$  there exists a convex hypersurface  $S_\varepsilon$  of class  $C_{\text{loc}}^{1,1}$  such that  $\mathcal{H}^{n-1}(S \setminus S_\varepsilon) < \varepsilon$ .

## 1. INTRODUCTION AND MAIN RESULTS

Let  $\mathcal{A}$  and  $\mathcal{C}$  be two classes of real valued functions defined on  $\mathbb{R}^n$  (or on an open subset of  $\mathbb{R}^n$ ). If for a given  $f \in \mathcal{A}$  and every  $\varepsilon > 0$  we can find a function  $g \in \mathcal{C}$  such that

$$(1.1) \quad \mathcal{L}^n(\{x : f(x) \neq g(x)\}) < \varepsilon,$$

we say that  $f$  has the *Lusin property of class*  $\mathcal{C}$ . Here, and in what follows  $\mathcal{L}^n$  denotes the Lebesgue measure in  $\mathbb{R}^n$ . If every function  $f \in \mathcal{A}$  satisfies this property, we also say that  $\mathcal{A}$  has the Lusin property of class  $\mathcal{C}$ .

This terminology comes from the well known theorem of Lusin which asserts that for every Lebesgue measurable function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  and every  $\varepsilon > 0$  there exists a continuous function  $g : \mathbb{R}^n \rightarrow \mathbb{R}$  such that  $\mathcal{L}^n(\{x : f(x) \neq g(x)\}) < \varepsilon$ . That is, measurable functions have the Lusin property of class  $C(\mathbb{R}^n)$ .

Several authors have shown that one can take  $g$  of class  $C^k$  if  $f$  has some weaker regularity properties of order  $k$ . For instance, Federer [19, p. 442] showed that almost everywhere differentiable functions (and in particular locally Lipschitz functions) have the Lusin property of class  $C^1$ . Whitney [30] improved this result by showing that a function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  has approximate partial derivatives of first order a.e. if and only if  $f$  has the Lusin property of class  $C^1$ ; see also [1, 26] for related results.

---

*Date:* November 11, 2020.

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

*Key words and phrases.* convex function, convex body, convex hypersurface, Lusin property, Whitney extension theorems.

P.H. was supported by NSF grant DMS-1800457.

In [15, Theorem 13] Calderón and Zygmund proved analogous results for  $\mathcal{A} = W^{k,p}(\mathbb{R}^n)$  (the class of Sobolev functions) and  $\mathcal{C} = C^k(\mathbb{R}^n)$ . Other authors, including Liu [24], Bagby, Michael and Ziemer [10, 27, 31], Bojarski, Hajłasz and Strzelecki [12, 13], and Bourgain, Korobkov and Kristensen [14] improved Calderón and Zygmund's result in several directions, by obtaining additional estimates for  $f - g$  in the Sobolev norms, as well as the Bessel capacities or the Hausdorff contents of the exceptional sets where  $f \neq g$ . In [14] some Lusin properties of the class  $BV_k(\mathbb{R}^n)$  (of integrable functions whose distributional derivatives of order up to  $k$  are Radon measures) are also established. On the other hand, generalizing Whitney's result [30] to higher orders of differentiability, Isakov [22] and Liu and Tai [25] independently established that a function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  has the Lusin property of class  $C^k$  if and only if  $f$  is approximately differentiable of order  $k$  almost everywhere (and if and only if  $f$  has an approximate  $(k - 1)$ -Taylor polynomial at almost every point). See also [20, 30] for related results.

The Whitney extension technique [29], or other related techniques, such as the Whitney smoothing [13], play a key role in the proofs of all of these results.

For the special class of *convex* functions  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , Alberti, Imomkulov [2, 21] and Evans and Gangbo [17, Proposition A.1] showed that every convex function has the Lusin property of class  $C^2$ ; however, given a convex function  $f$ , the function  $g \in C^2$  satisfying (1.1) that they obtained is not necessarily convex. Indeed, their arguments were based on the Whitney extension theorem and Whitney's construction does not preserve convexity. Thus it is natural to consider the following problems.

Denote by  $C^{1,1}(\mathbb{R}^n)$ ,  $C_{\text{loc}}^{1,1}(\mathbb{R}^n)$  and  $C_{\text{conv}}(\mathbb{R}^n)$  respectively the class of real valued  $C^1$  functions with Lipschitz continuous gradient, the class of  $C^1$  functions with locally Lipschitz continuous gradient and the class of convex functions, and define

$$\begin{aligned} C_{\text{conv}}^{1,1}(\mathbb{R}^n) &= C^{1,1}(\mathbb{R}^n) \cap C_{\text{conv}}(\mathbb{R}^n); \\ C_{\text{conv}}^{1,1\text{loc}}(\mathbb{R}^n) &= C_{\text{loc}}^{1,1}(\mathbb{R}^n) \cap C_{\text{conv}}(\mathbb{R}^n); \\ C_{\text{conv}}^2(\mathbb{R}^n) &= C^2(\mathbb{R}^n) \cap C_{\text{conv}}(\mathbb{R}^n). \end{aligned}$$

**Problem 1.1.** Given  $\mathcal{C} \in \{C_{\text{conv}}^{1,1}(\mathbb{R}^n), C_{\text{conv}}^{1,1\text{loc}}(\mathbb{R}^n), C_{\text{conv}}^2(\mathbb{R}^n)\}$ , does  $C_{\text{conv}}(\mathbb{R}^n)$  have the Lusin property of class  $\mathcal{C}$ ?

Example 1.8 and Proposition 1.10 show that the answer to this question is in the negative. Thus we ask:

**Problem 1.2.** For every such  $\mathcal{C}$ , can we characterize the subclass of  $C_{\text{conv}}(\mathbb{R}^n)$  which has the Lusin property of class  $\mathcal{C}$ ?

In Theorem 1.12 we provide a complete answer to this question when  $\mathcal{C} = C_{\text{conv}}^{1,1\text{loc}}(\mathbb{R}^n)$ .

**Problem 1.3.** What happens if we replace  $\mathbb{R}^n$  with an open bounded convex subset  $\Omega$  of  $\mathbb{R}^n$ ?

In Corollary 1.6 we show that the answer is in the positive when  $\mathcal{C} = C_{\text{conv}}^{1,1}(\Omega)$  (and hence when  $\mathcal{C} = C_{\text{conv}}^{1,1\text{loc}}(\Omega)$ ). See also Theorem 1.4

Our proofs are based on some results and techniques concerning  $C_{\text{conv}}^{1,1}(\mathbb{R}^n)$  and  $C_{\text{conv}}^{1,1\text{loc}}(\mathbb{R}^n)$  extensions of 1-jets recently obtained in [6, 7, 8, 9]. Problems 1.2 and 1.3 remain open for  $\mathcal{C} = C_{\text{conv}}^2(\mathbb{R}^n)$ , and they look rather hard in the absence of a characterization of the 2-jets (defined on an arbitrary subset of  $\mathbb{R}^n$ ) which admit  $C_{\text{conv}}^2(\mathbb{R}^n)$  extensions.

As an application of our results, we will show that all boundaries of compact convex bodies in  $\mathbb{R}^n$  have the Lusin property of class  $C_{\text{conv}}^{1,1}$ , meaning that they are equal, up to subsets of arbitrarily small measures, to boundaries of convex bodies of class  $C^{1,1}$ ; see Corollary 1.7 below.

Our first main result is as follows.

**Theorem 1.4.** *Let  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  be a convex function. For every measurable subset  $A$  of finite Lebesgue measure in  $\mathbb{R}^n$ , and for every  $\varepsilon > 0$ , there exists a convex function  $g : \mathbb{R}^n \rightarrow \mathbb{R}$  of class  $C^{1,1}$  such that*

$$(1.2) \quad \mathcal{L}^n(\{x \in A : f(x) \neq g(x)\}) < \varepsilon.$$

Note that if  $f = g$  in a measurable set  $E$ , then it is easy to show that  $\nabla f = \nabla g$  a.e. in  $E$ , so condition (1.2) is equivalent to a seemingly stronger one:

$$\mathcal{L}^n(\{x \in A : f(x) \neq g(x) \text{ or } \nabla f(x) \neq \nabla g(x)\}) < \varepsilon$$

which says that outside a set of measure less than  $\varepsilon$ , we have  $f = g$  and  $\nabla f = \nabla g$ .

As a corollary we obtain a new proof of a result mentioned above.

**Corollary 1.5.** *If  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is convex, then for any  $\varepsilon > 0$  there is  $g \in C^2(\mathbb{R}^n)$  such that  $\mathcal{L}^n(\{x : f(x) \neq g(x)\}) < \varepsilon$ .*

We do not claim however, that  $g$  is convex.

**Corollary 1.6.** *Let  $\Omega$  be a bounded, open and convex subset of  $\mathbb{R}^n$ . Then for every convex function  $f : \Omega \rightarrow \mathbb{R}$  and every  $\varepsilon > 0$  there exists a convex function  $g : \mathbb{R}^n \rightarrow \mathbb{R}$  of class  $C^{1,1}$  such that  $\mathcal{L}^n(\{x \in \Omega : f(x) \neq g(x)\}) < \varepsilon$ .*

Recall that a subset  $W$  of  $\mathbb{R}^n$  is a *compact convex body* if  $W$  is compact and convex, with nonempty interior. We will say that a compact convex body  $W$  is of class  $C^{1,1}$  provided that its boundary  $\partial W$  is a  $C^1$  hypersurface of  $\mathbb{R}^n$  such that the outer unit normal  $N : \partial W \rightarrow \mathbb{S}^{n-1}$  is a Lipschitz mapping. If  $W$  is a compact convex body, this is equivalent to saying that the Minkowski functional of  $W$  is of class  $C^{1,1}$  on  $\mathbb{R}^n \setminus B(0, \varepsilon)$  for some  $\varepsilon > 0$ , or that  $W$  can be locally parameterized as a graph  $(x_1, \dots, x_{n-1}, g(x_1, \dots, x_{n-1}))$  (coordinates taken with respect to an appropriate permutation of the canonical basis of  $\mathbb{R}^n$ ), where  $g$  is of class  $C^{1,1}$ .

A consequence of Theorem 1.4 is that the boundary of every compact convex body in  $\mathbb{R}^n$  is of class  $C^{1,1}$  up to a subset of arbitrarily small  $(n-1)$ -dimensional Hausdorff measure.

**Corollary 1.7.** *Let  $W$  be a compact convex body in  $\mathbb{R}^n$ . Then for every  $\varepsilon > 0$  there exists a compact convex body  $W_\varepsilon$  of class  $C^{1,1}$  such that  $\mathcal{H}^{n-1}(\partial W \setminus \partial W_\varepsilon) < \varepsilon$ .*

The following examples show that the assumption  $\mathcal{L}^n(A) < \infty$  in Theorem [1.4](#) cannot be dispensed with (unless other hypotheses on the global behaviour and the growth of  $f$  are put in their place).

**Example 1.8.** Let  $f : \mathbb{R} \rightarrow \mathbb{R}$  be defined by  $f(x) = x^4$ . Then there is no function  $g \in C^{1,1}(\mathbb{R})$  such that  $\mathcal{L}^1(\{x \in \mathbb{R} : f(x) \neq g(x)\}) < \infty$ .

Indeed, the second derivative of  $f$  is bounded by a constant only on a set of finite measure.

**Example 1.9.** Let  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  be defined by  $f(x, y) = |x|$ . Then  $f$  is the only convex function  $g : \mathbb{R}^2 \rightarrow \mathbb{R}$  such that  $\mathcal{L}^n(\{x \in \mathbb{R}^n : f(x) \neq g(x)\}) < \infty$ .

More generally we have the following.

**Proposition 1.10.** Let  $P : \mathbb{R}^n \rightarrow X$  be the orthogonal projection onto a linear subspace  $X$  of  $\mathbb{R}^n$  of dimension  $k$ , with  $1 \leq k \leq n - 1$ , let  $c : X \rightarrow \mathbb{R}$  be a convex function, and define  $f(x) = c(P(x))$ . Then  $f$  is the only convex function  $g : \mathbb{R}^n \rightarrow \mathbb{R}$  such that  $\mathcal{L}^n(\{x \in \mathbb{R}^n : f(x) \neq g(x)\}) < \infty$ .

*Proof.* Let  $g : \mathbb{R}^n \rightarrow \mathbb{R}$  be a convex function such that  $\mathcal{L}^n(A) < \infty$ , where

$$A := \{x \in \mathbb{R}^n : f(x) \neq g(x)\}.$$

Let  $X^\perp$  stand for the orthogonal complement of  $X$  in  $\mathbb{R}^n$ . By Fubini's theorem, for  $\mathcal{H}^k$ -almost every point  $x \in X$ , we have that for  $\mathcal{H}^{n-k-1}$ -almost every direction  $v \in X^\perp$ ,  $|v| = 1$ , the line

$$L(x, v) := \{x + tv : t \in \mathbb{R}\}$$

intersects  $A$  in a set of finite 1-dimensional measure. This implies that for all such  $x \in X$ ,  $v \in X^\perp$ , the set  $L(x, v) \cap (\mathbb{R}^n \setminus A)$  contains sequences

$$x_j^\pm := x + t_{x,j}^\pm v \in \mathbb{R}^n \setminus A, j \in \mathbb{N}$$

with  $\lim_{j \rightarrow \pm\infty} t_{x,j}^\pm = \pm\infty$ . Since  $f = f \circ P$ , this means that

$$f(x) = f(x + t_{x,j}^\pm v) = g(x + t_{x,j}^\pm v),$$

and because  $t \mapsto g(x + tv)$  is convex we see that

$$f(x + tv) = f(x) = g(x + tv)$$

for all  $t \in \mathbb{R}$  and every such  $x, v$ . By continuity of  $f$  and  $g$  this implies that

$$f(x + tv) = g(x + tv)$$

for all  $x \in X$ ,  $v \in X^\perp$ , and this shows that  $f = g$  on  $\mathbb{R}^n$ . □

In light of Example [1.8](#), Theorem [1.4](#) and Corollary [1.6](#) seem the best possible results for the Lusin property of class  $C_{\text{conv}}^{1,1}$ .

Thus it is natural to consider the Lusin property of class  $C_{\text{conv}}^{1,1 \text{ loc}}$  or  $C_{\text{conv}}^2$ , where we do not have any restrictions on the growth of the second derivative. Clearly, if the function  $f$  from Proposition [1.10](#) is not already  $C_{\text{loc}}^{1,1}$ , it does not have the Lusin property of class  $C_{\text{conv}}^{1,1 \text{ loc}}$  or

$C_{\text{conv}}^2$ . Under the assumptions of Proposition [1.10](#), if  $0 \neq v \in X^\perp$ , then the function  $f$  is constant along the line  $t \rightarrow tv$  and hence there is no linear function  $\ell : \mathbb{R}^n \rightarrow \mathbb{R}$  such that  $f(x) - \ell(x) \rightarrow \infty$  as  $|x| \rightarrow \infty$ . That is,  $f$  is not essentially coercive, where the essential coercivity is defined as follows:

**Definition 1.11.** *We say that a convex function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is essentially coercive provided there exists a linear function  $\ell : \mathbb{R}^n \rightarrow \mathbb{R}$  such that*

$$\lim_{|x| \rightarrow \infty} (f(x) - \ell(x)) = \infty.$$

It is natural to wonder whether the lack of essential coercivity of a convex function defined on  $\mathbb{R}^n$  is the only obstruction for it to satisfy a Lusin property of type  $C_{\text{conv}}^{1,1,\text{loc}}(\mathbb{R}^n)$  or  $C_{\text{conv}}^2(\mathbb{R}^n)$ . Our second main result shows that this is indeed so in the  $C_{\text{conv}}^{1,1,\text{loc}}(\mathbb{R}^n)$  case.

**Theorem 1.12.** *Let  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  be a convex function, and assume that  $f$  is not of class  $C_{\text{loc}}^{1,1}(\mathbb{R}^n)$ . Then  $f$  is essentially coercive if and only if 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  $\mathcal{L}^n(\{x \in \mathbb{R}^n : f(x) \neq g(x)\}) < \varepsilon$ .*

As an easy application we obtain the following generalization of Corollary [1.7](#). We call  $S$  a convex hypersurface of  $\mathbb{R}^n$  provided that  $S$  is the boundary of a closed convex set  $W$  with nonempty interior (not necessarily bounded) in  $\mathbb{R}^n$ , and we say that  $S$  is of class  $C_{\text{loc}}^{1,1}$  whenever  $S$  is a 1-codimensional  $C^1$  submanifold of  $\mathbb{R}^n$  such that the outer unit normal  $N : S \rightarrow \mathbb{S}^{n-1}$  is a locally Lipschitz mapping. Equivalently, the Minkowski functional  $\mu_W$  of  $W$  is of class  $C_{\text{loc}}^{1,1}$  on  $\mathbb{R}^n \setminus \mu_W^{-1}(0)$ .

**Corollary 1.13.** *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)  $S$  does not contain any line.
- (2) For every  $\varepsilon > 0$  there exists a convex hypersurface  $S_\varepsilon$  of class  $C_{\text{loc}}^{1,1}$  of  $\mathbb{R}^n$  such that  $\mathcal{H}^{n-1}(S \setminus S_\varepsilon) < \varepsilon$ .

## 2. PRELIMINARIES AND TOOLS

In this section we explain some known results and techniques which we will use in the proofs of our main results.

A convex function  $f : \Omega \rightarrow \mathbb{R}$  defined on an open and convex set  $\Omega \subset \mathbb{R}^n$  is locally Lipschitz and hence differentiable almost everywhere by Rademacher's theorem. In fact the following result is true:

**Theorem 2.1.** *If a convex function  $f : \Omega \rightarrow \mathbb{R}$  is defined on an open and convex set  $\Omega \subset \mathbb{R}^n$ , and  $D \subset \Omega$  is the set of points where  $f$  is differentiable, then  $\nabla f|_D$  is continuous.*

For a proof see e.g. [\[28\]](#), Theorem IV.E]. An elementary and a straightforward argument can also be found in [\[23\]](#), p.727].

According to Aleksandrov's theorem [4], at almost every point  $x$ , where  $f$  is differentiable, there is a symmetric  $n \times n$  matrix  $\nabla^2 f(x)$  such that

$$(2.1) \quad \lim_{y \rightarrow x} \frac{f(y) - f(x) - \langle \nabla f(x), y - x \rangle - \frac{1}{2} \langle \nabla^2 f(x)(y - x), y - x \rangle}{|y - x|^2} = 0.$$

For modern proofs, see for example [3, Theorem 7.10], [11], [16, Theorem A.2] [18, Theorem 6.9].

A common technique of showing that a class of functions has a Lusin property of class  $C^k$  is based on the Whitney extension theorem. For example, it follows from the Aleksandrov theorem that a convex function satisfies the assumptions of a  $C^2$ -version of the Whitney extension theorem outside a set of an arbitrarily small measure and hence the class  $C_{\text{conv}}(\mathbb{R}^n)$  has the Lusin property of class  $C^2(\mathbb{R}^n)$ , see [2, 17, 21]. Unfortunately, Whitney's construction does not preserve convexity. Instead we will use results and techniques from [6, 7, 8, 9] which we next review.

**Theorem 2.2.** ([8, Corollary 1.3] and [7, Theorem 2.4]). *Let  $E$  be an arbitrary subset of  $\mathbb{R}^n$ . Let  $f : E \rightarrow \mathbb{R}$ ,  $G : E \rightarrow \mathbb{R}^n$  be given functions. Then there exists a convex function  $F \in C^{1,1}(\mathbb{R}^n)$  with  $F = f$  and  $\nabla F = G$  on  $E$  if and only if there exists a number  $M > 0$  such that*

$$(2.2) \quad f(x) - f(y) - \langle G(y), x - y \rangle \geq \frac{1}{2M} |G(x) - G(y)|^2$$

for all  $x, y \in E$ . In fact, the formula

$$F = \text{conv} \left( x \mapsto \inf_{y \in E} \left\{ f(y) + \langle G(y), x - y \rangle + \frac{M}{2} |x - y|^2 \right\} \right)$$

defines such an extension, with the additional property that  $\text{Lip}(\nabla F) \leq M$ .

Here  $\text{conv}(x \mapsto g(x))$  denotes the convex envelope of the function  $g$ , that is, the largest convex function  $\varphi$  such that  $\varphi \leq g$ .

This result was first proved in [8, Corollary 1.3], but the proof given in [7, Theorem 2.4] is elementary and much simpler. The next elementary lemma shows another condition that is equivalent to (2.2). In fact, in our proofs we will apply Theorem 2.2 by verifying condition (b) from below.

**Lemma 2.3.** *Let  $E$  be an arbitrary subset of  $\mathbb{R}^n$ . Let  $f : E \rightarrow \mathbb{R}$ ,  $G : E \rightarrow \mathbb{R}^n$  be given functions and let  $M > 0$  be a given constant. Then the following conditions are equivalent:*

- (a)  $f(x) - f(y) - \langle G(y), x - y \rangle \geq \frac{1}{2M} |G(x) - G(y)|^2$  for all  $x, y \in E$ ;
- (b)  $f(z) + \langle G(z), x - z \rangle \leq f(y) + \langle G(y), x - y \rangle + \frac{M}{2} |x - y|^2$  for all  $y, z \in E$  and all  $x \in \mathbb{R}^n$ .

*Proof.* Implication from (a) to (b) is just [7, Lemma 2.6]. Thus it remains to show that (b) implies (a). Renaming variables allows us to rewrite (b) as

$$f(y) + \langle G(y), \xi - y \rangle \leq f(x) + \langle G(x), \xi - x \rangle + \frac{M}{2} |\xi - x|^2 \quad \text{for } x, y \in E \text{ and } \xi \in \mathbb{R}^n,$$

which is equivalent to

$$f(x) - f(y) - \langle G(y), x - y \rangle \geq \langle G(y) - G(x), \xi - x \rangle - \frac{M}{2} |\xi - x|^2, \quad \text{for } x, y \in E \text{ and } \xi \in \mathbb{R}^n.$$

Since the inequality is true for all  $\xi \in \mathbb{R}^n$ , we can take  $\xi$  such that  $\xi - x = \frac{1}{M}(G(y) - G(x))$  and (a) follows.  $\square$

In the same spirit, a more complicated version of Theorem 2.2 for  $C_{\text{loc}}^{1,1}$  convex extensions of 1-jets has been established in [6, Theorem 1.10]. However, we will need the following special case which is easier to state.

**Theorem 2.4.** ([6, Theorem 1.3]). *Let  $E$  be an arbitrary nonempty subset of  $\mathbb{R}^n$ . Let  $f : E \rightarrow \mathbb{R}$ ,  $G : E \rightarrow \mathbb{R}^n$  be functions such that*

$$(2.3) \quad \text{span}\{G(x) - G(y) : x, y \in E\} = \mathbb{R}^n.$$

*Then there exists a convex function  $F \in C_{\text{loc}}^{1,1}(\mathbb{R}^n)$  such that  $F|_E = f$  and  $(\nabla F)|_E = G$  if and only if for each  $k \in \mathbb{N}$  there exists a number  $A_k \geq 2$  such that*

$$(2.4) \quad f(z) + \langle G(z), x - z \rangle \leq f(y) + \langle G(y), x - y \rangle + \frac{A_k}{2} |x - y|^2$$

*for every  $z \in E$ ,  $y \in E \cap B(0, k)$ ,  $x \in B(0, 4k)$ .*

We will also need:

**Theorem 2.5.** ([9, Theorem 1.11] and [5, Lemma 4.2]). *For every convex function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , there exist a unique linear subspace  $X$  of  $\mathbb{R}^n$ , a unique vector  $v \in X^\perp$ , and a unique essentially coercive function  $c : X \rightarrow \mathbb{R}$  such that  $f$  can be written in the form*

$$f(x) = c(P(x)) + \langle v, x \rangle \quad \text{for all } x \in \mathbb{R}^n,$$

*where  $P : \mathbb{R}^n \rightarrow X$  is the orthogonal projection.*

The next extension lemma is well known, but for the sake of completeness we will provide a proof.

**Lemma 2.6.** *Suppose that  $f : W \rightarrow \mathbb{R}$  is convex and  $K$ -Lipschitz, where  $W \subset \mathbb{R}^n$  is convex. Then,*

$$(2.5) \quad \tilde{f}(x) = \inf_{z \in W} \{f(z) + K|x - z|\}, \quad x \in \mathbb{R}^n$$

*is convex and  $K$ -Lipschitz on  $\mathbb{R}^n$ , and  $\tilde{f} = f$  on  $W$ .*

*Proof.*  $K$ -Lipschitz continuity of  $f$  on  $W$  implies that for  $x, z \in W$  we have

$$f(x) \leq f(z) + K|x - z| \quad \text{so} \quad f(x) \leq \tilde{f}(x).$$

On the other hand,  $\tilde{f}(x) \leq f(x) + K|x - x| = f(x)$  so  $f = \tilde{f}$  on  $W$ . To prove that  $\tilde{f}$  is  $K$ -Lipschitz, let  $x, y \in \mathbb{R}^n$  and assume that  $\tilde{f}(x) \geq \tilde{f}(y)$ . For  $z \in W$  we have

$$\tilde{f}(x) \leq f(z) + K|x - z| \leq (f(z) + K|y - z|) + K|x - y|,$$

so taking infimum on the right hand side over  $z \in W$  yields  $K$ -Lipschitz continuity of  $\tilde{f}$ . It remains to show that  $\tilde{f}$  is convex. If  $x, y \in \mathbb{R}^n$  and  $\lambda \in [0, 1]$ , then for any  $z, w \in W$  we have

$$\begin{aligned} \tilde{f}(\lambda x + (1 - \lambda)y) &\leq f(\lambda z + (1 - \lambda)w) + K|(\lambda x + (1 - \lambda)y) - (\lambda z + (1 - \lambda)w)| \\ &\leq \lambda(f(z) + K|x - z|) + (1 - \lambda)(f(w) + K|y - w|) \end{aligned}$$

and taking infimum over  $z, w \in W$  yields convexity of  $\tilde{f}$ .  $\square$

### 3. PROOFS OF THE MAIN RESULTS

**3.1. Proof of Theorem 1.4.** Let  $E \subset A$  be a compact set such that  $\mathcal{L}^n(A \setminus E) < \varepsilon/2$  and that all points  $x \in E$  satisfy (2.1). It is easy to see that

$$E = \bigcup_{j=1}^{\infty} E_j, \quad E_1 \subset E_2 \subset \dots,$$

where

$$E_j = \left\{ y \in E : f(x) - f(y) - \langle \nabla f(y), x - y \rangle \leq j|x - y|^2 \text{ for all } x \in \mathbb{R}^n \text{ s.t. } |x - y| \leq \frac{1}{j} \right\}.$$

Since by Theorem 2.1,  $\nabla f$  is continuous on  $E$ , it is easy to check that the sets  $E_j$  are closed and hence compact. Since  $\mathcal{L}^n(E) < \infty$ ,  $\mathcal{L}^n(E \setminus E_N) < \varepsilon/2$  for some  $N$ , so  $\mathcal{L}^n(A \setminus E_N) < \varepsilon$ . Thus

$$(3.1) \quad \forall y \in E_N \quad \forall x \in \mathbb{R}^n \quad (|x - y| \leq N^{-1} \Rightarrow f(x) - f(y) - \langle \nabla f(y), x - y \rangle \leq N|x - y|^2).$$

Since  $E_N$  is compact,  $E_N \subset \overline{B}(0, r)$  for some  $r > 0$ . Let  $R > r + N^{-1}$  and let  $W = \overline{B}(0, R)$ . Since convex functions are locally Lipschitz,  $f|_W$  is  $K$ -Lipschitz for some  $K > 0$  and we may assume that  $K \geq 1$ .

Let  $\tilde{f}$  be defined by (2.5). Then  $\tilde{f}$  is convex and  $K$ -Lipschitz on  $\mathbb{R}^n$ . Since  $\tilde{f} = f$  in  $W$ , the function  $\tilde{f}$  satisfies (3.1). Observe that

$$(3.2) \quad \tilde{f}(x) - \tilde{f}(y) - \langle \nabla \tilde{f}(y), x - y \rangle \leq 2KN|x - y|^2 \quad \text{for all } y \in E_N \text{ and all } x \in \mathbb{R}^n.$$

Indeed, if  $|x - y| \leq N^{-1}$ , (3.2) follows from (3.1). If  $|x - y| > N^{-1}$ , then

$$\tilde{f}(x) - \tilde{f}(y) - \langle \nabla \tilde{f}(y), x - y \rangle \leq K|x - y| + |\nabla \tilde{f}(y)| |x - y| \leq 2K|x - y| < 2KN|x - y|^2.$$

Also convexity of  $\tilde{f}$  yields

$$(3.3) \quad \tilde{f}(z) + \langle \nabla \tilde{f}(z), x - z \rangle \leq \tilde{f}(x) \quad \text{for all } z \in E_N \text{ and all } x \in \mathbb{R}^n.$$



By combining (3.3) and (3.2) we get

$$\tilde{f}(z) + \langle \nabla \tilde{f}(z), x - z \rangle \leq \tilde{f}(y) + \langle \nabla \tilde{f}(y), x - y \rangle + 2KN|x - y|^2 \quad \text{for } y, z \in E_N \text{ and } x \in \mathbb{R}^n.$$

This is condition (b) from Lemma 2.3. Since the condition is equivalent to (2.2), Theorem 2.2 gives that there is a  $C^{1,1}$  function  $F$  defined in  $\mathbb{R}^n$  such that  $F = f = \tilde{f}$  and  $\nabla F = \nabla \tilde{f} = \nabla f$  in  $E_N$ . This and the fact that  $\mathcal{L}^n(A \setminus E_N) < \varepsilon$  complete the proof.  $\square$

**3.2. Proof of Corollary 1.5.** According to Theorem 1.4 for each positive integer  $i$ , there is  $g_i \in C^{1,1}(\mathbb{R}^n)$  such that  $\mathcal{L}^n(\{x \in B(0, i) : f(x) \neq g_i(x)\}) < \varepsilon/2^{i+1}$ . Let  $\{\varphi_i\}_{i=1}^\infty$  be a smooth partition of unity subordinate to the covering  $\{B(0, i)\}_{i=1}^\infty$  of  $\mathbb{R}^n$ . Then  $g = \sum_{i=1}^\infty g_i \varphi_i \in C_{\text{loc}}^{1,1}$  satisfies  $\mathcal{L}^n(\{x \in \mathbb{R}^n : g(x) \neq f(x)\}) < \varepsilon/2$  and the result follows from Whitney's theorem [30, Theorem 4] according to which a  $C_{\text{loc}}^{1,1}$  function on  $\mathbb{R}^n$  coincides with a  $C^2$  function outside a set of measure less than  $\varepsilon/2$ .  $\square$

**3.3. Proof of Corollary 1.6.** Take a compact convex body  $W_\varepsilon$  such that

$$(3.4) \quad W_\varepsilon \subset \Omega, \quad \text{and } \mathcal{L}^n(\Omega \setminus W_\varepsilon) < \frac{\varepsilon}{2}.$$

Denote the Lipschitz constant of  $f|_{W_\varepsilon}$  by  $K$  (notice that  $W_\varepsilon$  is at positive distance from the boundary of  $\Omega$ , so this Lipschitz constant exists). According to Lemma 2.6,

$$\tilde{f}(x) = \inf_{z \in W_\varepsilon} \{f(z) + K|x - z|\}$$

is convex on  $\mathbb{R}^n$  and  $\tilde{f} = f$  on  $W_\varepsilon$ . Therefore, we can apply Theorem 1.4 to  $\tilde{f}$  and find a function  $g \in C_{\text{conv}}^{1,1}(\mathbb{R}^n)$  such that

$$\mathcal{L}^n(\{x \in W_\varepsilon : g(x) \neq f(x)\}) = \mathcal{L}^n(\{x \in W_\varepsilon : g(x) \neq \tilde{f}(x)\}) < \frac{\varepsilon}{2}.$$

This and (3.4) imply that  $\mathcal{L}^n(\{x \in \Omega : g(x) \neq f(x)\}) < \varepsilon$ .  $\square$

**3.4. Proof of Corollary 1.7.** We can assume that  $0 \in \text{int}(W)$  and consider the Minkowski functional of  $W$ , defined by

$$\mu(x) = \inf\{\lambda \geq 0 : x \in \lambda W\},$$

which is a Lipschitz convex function on  $\mathbb{R}^n$ . Let  $L$  be the Lipschitz constant of  $\mu$ . By using Corollary 1.6 we may find a function  $g \in C_{\text{conv}}^{1,1}(\mathbb{R}^n)$  such that

$$\mathcal{L}^n(\{x \in 2W : \mu(x) \neq g(x)\}) < \frac{\varepsilon}{L}.$$

Now consider the annulus

$$C_{1,2} := 2W \setminus W = \{x \in \mathbb{R}^n : 1 < \mu(x) \leq 2\},$$

and define

$$A = \{x \in C_{1,2} : \mu(x) \neq g(x)\}.$$

By the coarea formula for Lipschitz functions (see [18, Theorem 3.10] for instance) we have

$$\varepsilon > L \mathcal{L}^n(A) \geq \int_A |\nabla \mu(x)| dx = \int_1^2 \mathcal{H}^{n-1}(A \cap \mu^{-1}(t)) dt.$$

This inequality and Sard's theorem imply that there exists a regular value  $t_0 \in (1, 2)$  of  $g \in C^{1,1}$  such that

$$\mathcal{H}^{n-1}(A \cap \mu^{-1}(t_0)) < \varepsilon.$$

Then we can define

$$W_\varepsilon = \frac{1}{t_0} g^{-1}(-\infty, t_0],$$

so that  $W_\varepsilon$  is a convex body of class  $C^{1,1}$ , with boundary

$$\partial W_\varepsilon = \frac{1}{t_0} g^{-1}(t_0) \quad \text{and hence} \quad t_0(\partial W \setminus \partial W_\varepsilon) = A \cap \mu^{-1}(t_0).$$

This yields

$$\mathcal{H}^{n-1}(\partial W \setminus \partial W_\varepsilon) \leq t_0^{n-1} \mathcal{H}^{n-1}(\partial W \setminus \partial W_\varepsilon) = \mathcal{H}^{n-1}(A \cap \mu^{-1}(t_0)) < \varepsilon.$$

□

**3.5. Proof of Theorem 1.12.** The necessity of the essential coercitivity assumption is clear from Proposition 1.10 and Theorem 2.5. The fact that this assumption is sufficient will follow from Theorem 2.4. The rest of the proof is similar to that of Theorem 1.4 but more complicated.

Thus assume that  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is convex,  $f \notin C_{\text{loc}}^{1,1}$  and  $f$  is essentially coercive. Fix  $0 < \varepsilon < 1$ . It remains to show that there is a convex  $g \in C_{\text{loc}}^{1,1}$  such that  $\mathcal{L}^n(\{f \neq g\}) < \varepsilon$ .

By Aleksandrov's theorem and by Theorem 2.1<sup>1</sup>, there is a closed set  $A \subset \mathbb{R}^n$  such that

$$\mathcal{L}^n(\mathbb{R}^n \setminus A) < \frac{\varepsilon}{2}, \quad \nabla f|_A \text{ is continuous, and } (2.1) \text{ is true for all } x \in A.$$

As in the proof of Theorem 1.4,

$$A = \bigcup_{j=1}^{\infty} E_j, \quad E_1 \subset E_2 \subset \dots,$$

where

$$E_j = \left\{ y \in A : f(x) - f(y) - \langle \nabla f(y), x - y \rangle \leq j|x - y|^2 \text{ for all } x \in \mathbb{R}^n \text{ s.t. } |x - y| \leq \frac{1}{j} \right\}.$$

Since  $\nabla f$  is continuous on  $A$ , it easily follows that the sets  $E_j$  are closed and hence measurable.

We set  $B_0 = \emptyset$ , and for each  $k \in \mathbb{N}$ , we define

$$B_k := B(0, k), \quad \text{and} \quad A_k := A \cap (B_k \setminus B_{k-1}).$$

Now, for each  $k \in \mathbb{N}$ , since the sequence  $\{E_j\}_{j \in \mathbb{N}}$  is increasing and  $A_k = \bigcup_{j=1}^{\infty} (E_j \cap A_k)$ , we can find  $j_k \in \mathbb{N}$  such that

$$\mathcal{L}^n(A_k \setminus E_{j_k}) < \frac{\varepsilon}{2^{k+1}},$$

and define, for each  $k \in \mathbb{N}$ ,

$$C_k := E_{j_k} \cap A_k,$$

<sup>1</sup>We could use Lusin's theorem instead of Theorem 2.1

and

$$C := \bigcup_{k=1}^{\infty} C_k.$$

We may obviously assume that

$$(3.5) \quad j_k \leq j_{k+1} \text{ for all } k \in \mathbb{N}.$$

We then have that

$$(3.6) \quad \mathcal{L}^n(A \setminus C) = \sum_{k=1}^{\infty} \mathcal{L}^n(A_k \setminus C_k) < \frac{\varepsilon}{2} \quad \text{so} \quad \mathcal{L}^n(\mathbb{R}^n \setminus C) < \varepsilon.$$

**Lemma 3.1.** *For each  $k \in \mathbb{N}$  there exists a number  $\beta_k \geq 1$  such that*

$$(3.7) \quad f(x) - f(y) - \langle \nabla f(y), x - y \rangle \leq \beta_k |x - y|^2 \text{ for all } y \in C \cap B_k \text{ and all } x \in B_{4k}.$$

*Proof.* Take  $y \in C \cap B_k$  and note that since  $(j_k)$  is increasing we have  $C \cap B_k \subset E_{j_k} \cap B_k \subset B_{4k}$ . In particular  $y \in E_{j_k}$ .

If  $x \in \mathbb{R}^n$  is such that  $|x - y| \leq 1/j_k$ , the inequality we seek obviously holds with  $\beta_k = j_k$ , because of the definition of  $E_{j_k}$ . On the other hand, if  $|x - y| > 1/j_k$  and  $x \in B_{4k}$ , then, since  $f$  is Lipschitz on the ball  $B_{4k}$ , we have

$$f(x) - f(y) - \langle \nabla f(y), x - y \rangle \leq 2 \operatorname{Lip} \left( f|_{B_{4k}} \right) |x - y| \leq 2 \operatorname{Lip} \left( f|_{B_{4k}} \right) j_k |x - y|^2.$$

In any case the Lemma is satisfied with  $\beta_k = \max \left\{ j_k, 2j_k \operatorname{Lip} \left( f|_{B_{4k}} \right) \right\}$ .  $\square$

Since  $f$  is convex we have, for all  $z \in C$ ,  $x \in \mathbb{R}^n$ , that

$$(3.8) \quad f(z) + \langle \nabla f(z), x - z \rangle \leq f(x),$$

which combined with the preceding lemma gives us

$$f(z) + \langle \nabla f(z), x - z \rangle \leq f(y) + \langle \nabla f(y), x - y \rangle + \beta_k |x - y|^2$$

for all  $z \in C$ ,  $y \in C \cap B_k$ ,  $x \in B_{4k}$ . That is to say, the jet  $(f(y), \nabla f(y))$ ,  $y \in C$ , satisfies condition (2.4) of Theorem 2.4 with  $A_k = 2\beta_k \geq 2$ .

Finally, let us check condition (2.3) which in our case reads as

$$\operatorname{span}\{\nabla f(x) - \nabla f(y) : x, y \in C\} = \mathbb{R}^n.$$

Fix some  $y_0 \in C$ , and consider the function  $g(x) = f(x) - f(y_0) - \langle \nabla f(y_0), x - y_0 \rangle$ . Since  $f$  is essentially coercive and convex,  $g$  is also essentially coercive and convex. We have that

$$\operatorname{span}\{\nabla f(x) - \nabla f(y) : x, y \in C\} = \operatorname{span}\{\nabla f(x) - \nabla f(y_0) : x \in C\} = \operatorname{span}\{\nabla g(x) : x \in C\}.$$

Thus it suffices to show that  $Y := \operatorname{span}\{\nabla g(x) : x \in C\} = \mathbb{R}^n$ . Seeking a contradiction, suppose that  $Y \neq \mathbb{R}^n$ . We can then take a vector  $0 \neq v \in Y^\perp$  such that

$$(3.9) \quad \langle \nabla g(x), v \rangle = 0 \text{ for all } x \in C.$$

Since  $\mathcal{L}^n(\mathbb{R}^n \setminus C) < \varepsilon$ , an easy application of Fubini's theorem shows that there exists  $x_0$  perpendicular to  $v$  such that the intersection of the line  $L := \{x_0 + tv : t \in \mathbb{R}\}$  with the set

$\mathbb{R}^n \setminus C$  has finite one-dimensional measure. This implies that  $L \cap C$  must contain sequences  $x_j^\pm := x_0 + t_j^\pm v, j \in \mathbb{N}$  with  $\lim_{j \rightarrow \pm\infty} t_j^\pm = \pm\infty$ .

Consider the restriction of  $g$  to the line  $L$  i.e., consider the convex function  $h(t) = g(x_0 + tv)$ . Since by (3.9),

$$h'(t_j^\pm) = \langle \nabla g(x_j^\pm), v \rangle = 0,$$

it follows that  $h$  is constant and hence  $g$  is constant on the line  $L$ . But this contradicts the fact that  $g$  is essentially coercive.

We have thus checked that the 1-jet  $(f(y), \nabla f(y)), y \in C$ , satisfies all the conditions of Theorem 2.4, and therefore there exists a locally  $C^{1,1}$  convex function  $F : \mathbb{R}^n \rightarrow \mathbb{R}$  such that  $F = f$  on  $C$ , and also  $\nabla F = \nabla f$  on  $C$ . In particular we have that

$$\mathcal{L}^n(\{x \in \mathbb{R}^n : f(x) \neq F(x)\}) \leq \mathcal{L}^n(\mathbb{R}^n \setminus C) < \varepsilon.$$

The proof of Theorem 1.12 is complete.  $\square$

**3.6. Proof of Corollary 1.13.** We will need to use the following:

**Lemma 3.2.** *Let  $W$  be a closed convex set such that  $0 \in \text{int}(W)$ , and  $\mu = \mu_W$  denote the Minkowski functional of  $W$ . The following assertions are equivalent:*

- (a)  $W$  does not contain any line.
- (b)  $\partial W$  does not contain any line.
- (c)  $\mu^{-1}(0)$  does not contain any line
- (d)  $\mu$  is essentially coercive.

*Proof.* (a)  $\implies$  (b) is obvious.

(b)  $\implies$  (c): if for some  $x, v$  with  $v \neq 0$  we have  $\mu(x + tv) = 0$  for all  $t \in \mathbb{R}$  then, for any  $y \in \mu^{-1}(1) = \partial W$  we have  $\mu(y + tv) = \mu(y - x + x + tv) \leq \mu(y - x) + \mu(x + tv) = \mu(y - x)$  for all  $t \in \mathbb{R}$ . In particular the convex function  $\mathbb{R} \ni t \mapsto \mu(y + tv) \in \mathbb{R}$  is bounded above, hence it is constant. That is to say,  $\mu(y + tv) = \mu(y) = 1$  for all  $t \in \mathbb{R}$ , which means that  $\partial W$  contains the line  $\{y + tv : t \in \mathbb{R}\}$ .

(c)  $\implies$  (d): If  $\mu$  is not essentially coercive then by Theorem 2.5, for some  $w \neq 0$  we have that  $t \mapsto \mu(tw)$  is linear, and this may only happen if  $\mu(tw) = 0$  for all  $t \in \mathbb{R}$ , because  $\mu \geq 0$ .

(d)  $\implies$  (a). If  $W$  contains a line  $\{x + tv : t \in \mathbb{R}\}$ , then  $t \mapsto \mu(x + tv)$  convex and bounded from above by 1 so is constant. This prevents  $\mu$  from being essentially coercive.  $\square$

Now we can prove Corollary 1.13.

(1)  $\implies$  (2): Let  $W$  be the closed convex set with nonempty interior such that  $S = \partial W$ . Without loss of generality we may assume that  $0$  is an interior point of  $W$ . Denote the Minkowski functional of  $W$  by  $\mu$ . By the assumption (1) and the preceding lemma,  $\mu$  is essentially coercive. Also observe that  $\mu$  is  $L$ -Lipschitz, where  $1/L = d(0, S)$ . Then the

same proof as in Corollary 1.7 (replacing  $C^{1,1}$  with  $C_{\text{loc}}^{1,1}$  at appropriate places, and using Theorem 1.12 instead of Corollary 1.6) shows (2).

(2)  $\implies$  (1): Suppose to the contrary that  $S$  contains a line, or equivalently that  $\mu^{-1}(0)$  contains a line  $L_0 = \{tv : t \in \mathbb{R}\}$ ,  $v \neq 0$ .

Given  $\varepsilon > 0$ , let  $S_\varepsilon$  be a convex hypersurface of class  $C_{\text{loc}}^{1,1}$  such that  $\mathcal{H}^{n-1}(S \setminus S_\varepsilon) < \varepsilon$ . Let  $W_\varepsilon$  be the closed convex set such that  $\partial W_\varepsilon = S_\varepsilon$ . If  $\varepsilon > 0$  is small enough, we may assume that  $0 \in \text{int}(W)$  and  $0 \in \text{int}(W_\varepsilon)$ .

Indeed, since  $S \notin C_{\text{loc}}^{1,1}$ , it is not a flat hyperplane and we can find points  $p_1, \dots, p_{n+1} \in S$  such that the simplex  $\text{conv}\{p_1, \dots, p_{n+1}\} \subset W$  has nonempty interior. By translating the coordinate system we may assume that  $0$  belongs to the interior of that simplex and hence  $0 \in \text{int}(W)$ . Then, if  $\varepsilon > 0$  is sufficiently small, we can find points  $p'_1, \dots, p'_{n+1} \in S_\varepsilon$  so close to the points  $p_1, \dots, p_{n+1}$  that  $0$  belongs to the interior of the simplex  $\text{conv}\{p'_1, \dots, p'_{n+1}\} \subset W_\varepsilon$  and hence  $0 \in \text{int}(W_\varepsilon)$ . Denote the Minkowski functional of  $W_\varepsilon$  by  $\mu_\varepsilon$ .

Since  $\mu$  is convex, if  $\nabla\mu(x) = 0$ , then  $\mu$  attains minimum at  $x$ . Since  $\mu \geq 0 = \mu(0)$ , it follows that  $x \in \mu^{-1}(0)$ . Therefore,  $|\nabla\mu| > 0$  almost everywhere in the set  $\mathbb{R}^n \setminus \mu^{-1}(0)$ .

Recall that  $L_0 \subset \mu^{-1}(0)$  so  $\mu(tv) = 0$  for all  $t \in \mathbb{R}$ . This implies that  $\mu$  is constant on every line parallel to  $L_0$ . Indeed,

$$\mu(x + tv) \leq \mu(x) + \mu(tv) = \mu(x)$$

so the convex function  $t \mapsto \mu(x + tv)$  is constant as bounded from above.

Let  $X = L_0^\perp$  be the orthogonal complement of  $L_0$  and let  $P : \mathbb{R}^n \rightarrow X$  be the orthogonal projection. Since  $\mu$  is constant on every line parallel to  $L_0$ ,  $\mu(x) = \mu(P(x))$ , hence  $\nabla\mu(x) = \nabla\mu(P(x))$ , and also

$$(3.10) \quad \mathbb{R}^n \setminus \mu^{-1}(0) = P^{-1}(X \setminus \mu^{-1}(0)).$$

Recall that  $|\nabla\mu(y)| > 0$  exists and is positive for almost all  $y \in X \setminus \mu^{-1}(0)$ .

Suppose that  $E \subset \mathbb{R}^n \setminus \mu^{-1}(0)$  is measurable and

$$\int_E |\nabla\mu| < \infty.$$

Since  $|\nabla\mu|$  is well defined and constant along almost all lines  $P^{-1}(y)$ ,  $y \in X \setminus \mu^{-1}(0)$  parallel to  $L_0$ , we note that  $P(E) \subseteq X \setminus \mu^{-1}(0)$ , and apply Fubini's theorem to obtain

$$\int_{P(E)} |\nabla\mu(y)| \mathcal{L}^1(P^{-1}(y) \cap E) d\mathcal{L}^{n-1}(y) = \int_E |\nabla\mu| d\mathcal{L}^n < \infty.$$

Since  $|\nabla\mu(y)| > 0$  for almost all  $y \in X \setminus \mu^{-1}(0)$ ,  $\mathcal{L}^1(P^{-1}(y) \cap E) < \infty$  for almost all  $y \in X \setminus \mu^{-1}(0)$  i.e., almost every line parallel to  $L_0$  and disjoint from  $\mu^{-1}(0)$  intersects  $E$  along a set of finite length. We shall use this observation below.

Let  $C_a := \{ty : y \in S \cap S_\varepsilon, t \in [0, a]\}$ ,  $a > 0$ . Then for  $0 < t \leq a$ ,

$$\mathcal{H}^{n-1}(\mu^{-1}(t) \setminus C_a) = \mathcal{H}^{n-1}(t(S \setminus S_\varepsilon)) = t^{n-1} \mathcal{H}^{n-1}(S \setminus S_\varepsilon) < \varepsilon t^{n-1},$$

and hence the coarea formula yields

$$\int_{\mu^{-1}((0,a]) \setminus C_a} |\nabla \mu(x)| dx = \int_0^a \mathcal{H}^{n-1}(\mu^{-1}(t) \setminus C_a) dt < \int_0^a \varepsilon t^{n-1} dt = \frac{\varepsilon a^n}{n} < \infty.$$

It follows from an observation that we made earlier, that almost every line parallel to  $L_0$  intersects  $\mu^{-1}((0, a]) \setminus C_a$  along a set of finite length and hence almost every line parallel to  $L_0$  that is contained in  $\mu^{-1}(0, a]$  intersects  $\mathbb{R}^n \setminus C_a$  along a set of finite length. Therefore, for such a line, the set  $L \cap C_a$  contains sequences

$$x_j^\pm := x + t_{x,j}^\pm v \in C_a, \quad \text{with} \quad \lim_{j \rightarrow \pm\infty} t_{x,j}^\pm = \pm\infty.$$

By convexity of  $\mu_\varepsilon$ , the fact that  $\mu = \mu_\varepsilon$  on  $C_a$ , and that  $\mu$  is constant on  $L$ , it follows that  $\mu(y) = \mu_\varepsilon(y)$  for all  $y \in L$ , and by continuity it follows that  $\mu = \mu_\varepsilon$  on every line parallel to  $L_0$  and contained in  $\mu^{-1}(0, a]$ . Since  $a > 0$  is arbitrary, it follows that  $\mu = \mu_\varepsilon$  on every line parallel to  $L_0$  and disjoint from  $\mu^{-1}(0)$  (because every line  $L$  parallel to  $L_0$  and disjoint from  $\mu^{-1}(0)$  is contained in  $\mu^{-1}(0, a]$  for some  $a > 0$ ; indeed, if  $\mu(x) = a > 0$  for some  $x \in L$ , then  $\mu = a$  on  $L$  and  $L \subset \mu^{-1}(0, a]$ ). Bearing in mind (3.10), we deduce that  $\mu = \mu_\varepsilon$  on  $\mathbb{R}^n \setminus \mu^{-1}(0)$ , and therefore  $S = \mu^{-1}(1) = \mu_\varepsilon^{-1}(1) = S_\varepsilon$ . But this is absurd, because  $S_\varepsilon$  is of class  $C_{\text{loc}}^{1,1}$  and  $S$  is not.

## REFERENCES

- [1] G. Alberti, *A Lusin type theorem for gradients*, J. Funct. Anal. 100 (1991), no. 1, 110–118.
- [2] G. Alberti, *On the structure of singular sets of convex functions*, Calc. Var. Partial Differential Equations 2 (1994), no. 1, 17–27.
- [3] G. Alberti and L. Ambrosio, *A geometrical approach to monotone functions in  $\mathbb{R}^n$* , Math. Z. 230 (1999), no. 2, 259–316.
- [4] A.D. Alexandroff, *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.
- [5] D. Azagra, *Global and fine approximation of convex functions*, Proc. London Math. Soc. 107 (2013), 799–824.
- [6] D. Azagra, *Locally  $C^{1,1}$  convex extensions of 1-jets*, preprint, 2019, arXiv:1905.02127. To appear in Rev. Matemática Iberoamericana.
- [7] D. Azagra, E. Le Gruyer, C. Mudarra, *Explicit formulas for  $C^{1,1}$  and  $C_{\text{conv}}^{1,\omega}$  extensions of 1-jets in Hilbert and superreflexive spaces*, J. Funct. Anal. 274 (2018), 3003–3032.
- [8] D. Azagra and C. Mudarra, *Whitney extension theorems for convex functions of the classes  $C^1$  and  $C^{1,\omega}$* , Proc. London Math. Soc. 114 (2017), 133–158.
- [9] D. Azagra and C. Mudarra, *Global geometry and  $C^1$  convex extensions of 1-jets*, Analysis and PDE 12 (2019) no. 4, 1065–1099.
- [10] T. Bagby and W.P. Ziemer, *Pointwise differentiability and absolute continuity*, Trans. Amer. Math. Soc. 191 (1974), 129–148.
- [11] G. Bianchi, A. Colesanti, C. Pucci, *On the second differentiability of convex surfaces*. Geom. Dedicata 60 (1996), 39–48.
- [12] B. Bojarski and P. Hajlasz, *Pointwise inequalities for Sobolev functions*, Studia Math. 106 (1993), 77–92.
- [13] B. Bojarski, P. Hajlasz, and P. Strzelecki, *Improved  $C^{k,\lambda}$  approximation of higher order Sobolev functions in norm and capacity*. Indiana Univ. Math. J. 51 (2002), 507–540.
- [14] J. Bourgain, M. V. Korobkov and J. Kristensen, *On the Morse-Sard property and level sets of  $W^{n,1}$  Sobolev functions on  $\mathbb{R}^n$* , J. Reine Angew. Math. 700 (2015), 93–112.

- [15] A. P. Calderón and A. Zygmund, *Local properties of solutions of elliptic partial differential equations*, *Studia Math.* 20 (1961), 171–225.
- [16] M.G. Crandall, H. Ishii, P.-L. Lions, *User’s guide to viscosity solutions of second order partial differential equations*. *Bull. Amer. Math. Soc. (N.S.)* 27 (1992), 1–67.
- [17] L.C. Evans, W. Gangbo, *Differential equations methods for the Monge-Kantorovich mass transfer problem*. *Mem. Amer. Math. Soc.* 137 (1999), no. 653.
- [18] L.C. Evans, R.F. Gariepy, *Measure theory and fine properties of functions*. Revised edition. Textbooks in Mathematics. CRC Press, Boca Raton, FL, 2015.
- [19] H. Federer, *Surface area. II*, *Trans. Amer. Math. Soc.* 55, (1944), 438–456.
- [20] G. Francos, *The Luzin theorem for higher-order derivatives*. *Michigan Math. J.* 61 (2012), 507–516.
- [21] S.A. Imomkulov, *Twice differentiability of subharmonic functions*. (Russian) *Izv. Ross. Akad. Nauk Ser. Mat.* 56 (1992), 877–888; translation in *Russian Acad. Sci. Izv. Math.* 41 (1993), 157–167.
- [22] N.M. Isakov, *A global property of approximately differentiable functions*, *Mathematical Notes of the Academy of Sciences of the USSR* (1987) 41 (1987), 280–285.
- [23] B. Kirchheim, J. Kristensen, *Differentiability of convex envelopes*. *C. R. Acad. Sci. Paris Sér. I Math.* 333 (2001), no. 8, 725–728.
- [24] Fon-Che Liu, *A Luzin type property of Sobolev functions*, *Indiana Univ. Math. J.* 26 (1977), 645–651.
- [25] F.-C. Liu and W.-S. Tai, *Approximate Taylor polynomials and differentiation of functions*, *Topol. Methods Nonlinear Anal.* 3 (1994), no. 1, 189–196.
- [26] L. Moonens, W. F. Pfeffer, *The multidimensional Luzin theorem*. *J. Math. Anal. Appl.* 339 (2008), 746–752.
- [27] J. Michael and W.P. Ziemer, *A Luzin type approximation of Sobolev functions by smooth functions*, *Contemp. Math.* 42 (1985), 135–167.
- [28] A.W. Roberts, D.E. Varberg, *Convex functions*. Pure and Applied Mathematics, Vol. 57. Academic Press, New York-London, 1973.
- [29] H. Whitney, *Analytic extensions of differentiable functions defined in closed sets*, *Trans. Amer. Math. Soc.* 36 (1934), 63–89.
- [30] H. Whitney, *On totally differentiable and smooth functions*, *Pacific J. Math.* 1, (1951). 143–159.
- [31] W. P. Ziemer, *Weakly Differentiable Functions*, Springer-Verlag, 1989.

DEPARTAMENTO DE ANÁLISIS MATEMÁTICO Y MATEMÁTICA APLICADA, FACULTAD CIENCIAS MATEMÁTICAS, UNIVERSIDAD COMPLUTENSE, 28040, MADRID, SPAIN.

*Email address:* azagra@mat.ucm.es

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF PITTSBURGH, 301 THACKERAY HALL, PITTSBURGH, PA 15260, USA.

*Email address:* hajlasz@pitt.edu