SUBDIFFERENTIABLE FUNCTIONS SATISFY LUSIN PROPERTIES OF CLASS C^1 OR C^2

D. AZAGRA, J. FERRERA, M. GARCÍA-BRAVO, AND J. GÓMEZ-GIL

ABSTRACT. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function. Assume that for a measurable set Ω and almost every $x \in \Omega$ there exists a vector $\xi_x \in \mathbb{R}^n$ such that

$$\liminf_{h\to 0} \frac{f(x+h) - f(x) - \langle \xi_x, h\rangle}{|h|^2} > -\infty.$$

Then we show that f satisfies a Lusin-type property of order 2 in Ω, that is to say, for every $\varepsilon > 0$ there exists a function $g \in C^2(\mathbb{R}^n)$ such that $\mathcal{L}^n(\{x \in \Omega : f(x) \neq g(x)\} \leq \varepsilon$. In particular every function which has a nonempty proximal subdifferential almost everywhere also has the Lusin property of class C^2. We also obtain a similar result (replacing C^2 with C^1) for the Fréchet subdifferential. Finally we provide some examples showing that this kind of results are no longer true for *Taylor subexpansions* of higher order.

A classical theorem of Lusin [27] states that for every Lebesgue measurable function $f : \mathbb{R}^n \to \mathbb{R}$ and every $\varepsilon > 0$ there exists a continuous function $g : \mathbb{R}^n \to \mathbb{R}$ such that

$$\mathcal{L}^n \left(\{x \in \mathbb{R}^n : f(x) \neq g(x)\}\right) \leq \varepsilon. \tag{1}$$

Here, as in the rest of this note, $\mathcal{L}^n$ denotes the Lebesgue measure in $\mathbb{R}^n$.

Several authors have shown that one can take g of class C^k, provided that f has some regularity properties of order k (for instance, locally bounded distributional derivatives up to the order k, or Taylor expansions of order k almost everywhere). If, given a differentiability class $\mathcal{C}$ and a function $f : \mathbb{R}^n \to \mathbb{R}$ we can find, for each $\varepsilon > 0$, a function $g \in \mathcal{C}$ satisfying (1), we will say that *f has the Lusin property of class $\mathcal{C}$*.

The first of such results was discovered by Federer [15, p. 442], who showed that a.e differentiable functions (and in particular locally Lipschitz functions) have the Lusin property of class C^1. H. Whitney [31] improved this result by showing that a function $f : \mathbb{R}^n \to \mathbb{R}$ has approximate partial derivatives of first order a.e. if and only if f has the Lusin property of class C^1.

In [11, Theorem 13] Calderon and Zygmund established analogous results of order k for the classes of Sobolev functions $W^{k,p}(\mathbb{R}^n)$. Other authors, including Liu [25], Bagby, Michael and Ziemer [5, 28, 32], Bojarski, Hajłasz and Strzelecki [6, 7], and Bourgain, Korobkov and Kristensen [8] have improved Calderon and Zygmund's result in different ways, 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 [8] 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. The Whitney extension technique [30], and some related techniques as the *Whitney smoothing* introduced in [7], play a key role in the proofs of all of these results.

Key words and phrases. Lusin property of order 2, Proximal subdifferential, Fréchet subdifferential.

For the special class of convex functions $f:\mathbb{R}^n\to\mathbb{R}$, Alberti and Imonkulov [2, 21] showed that every convex function has the Lusin property of class C^2 (with g not necessarily convex in (1)); see also [1] for a related problem. More recently Azagra and Hajłasz [4] have proved that g can be taken to be C^1 and convex in (1) if and only if either f is essentially coercive (meaning that f is coercive up to a linear perturbation) or else f is already C^1 (in which case taking $g=f$ is the only possible option); they have also shown that if $f:\mathbb{R}^n\to\mathbb{R}$ is strongly convex then for every $A\subset\mathbb{R}^n$ of finite measure and every $\varepsilon>0$ there exists $g:\mathbb{R}^n\to\mathbb{R}$ convex and $C^{1,1}$ such that $\mathcal{L}^n\left(\{x\in A: f(x)\neq g(x)\}\right)\leq\varepsilon$.
On the other hand, generalizing Whitney's result [31] to higher orders of differentiability, Isakov [22] and Liu and Tai [26] independently established that a function $f:\mathbb{R}^n\to\mathbb{R}$ 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).
In this note we will answer the following question (which we think may be quite natural for people working on nonsmooth analysis or viscosity solutions to PDE such as Hamilton-Jacobi equations): do functions with nonempty subdifferentials a.e. have Lusin properties of order C^1 or C^2? By subdifferentials we mean the Fréchet subdifferential, or the proximal subdifferential, or the second order viscosity subdifferential; see [12, 13, 16] and the references therein for information about subdifferentials and their applications. As we will see the answer is positive: Fréchet subdifferentiable functions have the Lusin property of class C^1, and functions with nonempty proximal subdifferentials a.e. (in particular functions with a.e. nonempty viscosity subdifferentials of order 2) have the Lusin property of class C^2.
This question can be formulated in a more general form (perhaps appealing to a wider audience) as a problem about *Taylor subexpansions*: given $k\in\mathbb{N}$ and a function $f:\mathbb{R}^n\to\mathbb{R}$, assume that for almost every $x\in\mathbb{R}^n$ there exists a polynomial P_x of degree less than or equal to $k-1$ such that

$$\liminf_{y\to x}\frac{f(y)-P_x(y)}{|y-x|^k}>-\infty.$$

Is it then true that f has the Lusin property of order k?
The results of this note will show that the answer to this question is positive for $k=1,2$, but negative for $k\geq 3$.
In the case $k=1$ the proof is very simple and natural.

Theorem 1. *Let $\Omega\subset\mathbb{R}^n$ be a Lebesgue measurable set, and $f:\Omega\to\mathbb{R}$ a function. Assume that for almost every $x\in\Omega$ we have*

$$\liminf_{y\to x}\frac{f(y)-f(x)}{|y-x|}>-\infty. \tag{2}$$

Then, for every $\varepsilon>0$ there exists a function $g\in C^1(\mathbb{R}^n)$ such that

$$\mathcal{L}^n\left(\{x\in\Omega: f(x)\neq g(x)\}\right)\leq\varepsilon.$$

In order to facilitate the proof of Theorem 1, as well as that of Theorem 6 below, let us state the following technical lemma, which is standard. We include its proof for the readers' convenience.

Lemma 2. *Let Ω be a Lebesgue measurable subset of $\mathbb{R}^n$, $k\in\mathbb{N}$, and $f:\Omega\to\mathbb{R}$ be measurable. Then f has the Lusin property of class C^k (meaning that for every $\varepsilon>0$ there exists $g\in C^k(\mathbb{R}^n)$ such that $\mathcal{L}^n\{x\in\Omega: f(x)\neq g(x)\})\leq\varepsilon$) if and only if the restriction of f to each compact subset of Ω has the Lusin property of class C^k.*

Proof. It is obvious that if $f:\Omega\to\mathbb{R}$ has the Lusin property of class C^k then, for every compact subset K of Ω, the function $f_{|_K}:K\to\mathbb{R}$ has the Lusin property of class C^k. Let us prove the converse. Assume first that Ω is bounded. By the regularity of the measure $\mathcal{L}^n$, for every $\varepsilon>0$ we may find K_ε, a compact subset of Ω, such that $\mathcal{L}^n\left(\Omega\setminus K_\varepsilon\right)\leq\varepsilon/2$. By assumption, there exists a function $g=g_{K_\varepsilon}\in C^k(\mathbb{R}^n)$ such that $\mathcal{L}^n\left(\{x\in K_\varepsilon: f(x)\neq g(x)\}\right)\leq\varepsilon/2$. Then we have

$$\mathcal{L}^n\left\{x\in\Omega: f(x)\neq g(x)\}\right)\leq\mathcal{L}^n\left(\Omega\setminus K_\varepsilon\right)+\mathcal{L}^n\left(\{x\in K_\varepsilon: f(x)\neq g(x)\}\right)\leq\varepsilon,$$

and therefore $f:\Omega\to\mathbb{R}$ has the Lusin property of class C^k.

Now let us consider the general case that Ω is not necessarily bounded. We can write

$$\Omega=\bigcup_{j=1}^{\infty}\Omega_j, \text{ where } \Omega_1=\Omega\cap\text{int}B(0,1), \text{ and } \Omega_{j+1}:=\Omega\cap\text{int}B(0,j+1)\setminus B(0,j),$$

where $B(x,r)$ denotes the closed ball of center x and radius r. According to the previous argument, for each $j\in\mathbb{N}$ there exists a function $g_j\in C^k(\mathbb{R}^n)$ such that

$$\mathcal{L}^n\left(\{x\in\Omega_j: g_j(x)\neq f(x)\}\right)\leq\frac{\varepsilon}{6^j}.$$

Let $(\psi_j)_{j=1}^{\infty}$ be a C^∞ smooth partition of unity subordinated to the covering $\{\text{int}B(0,j+1)\setminus B(0,j-1)\}_{j=1}^{\infty}\cup\{\text{int}B(0,1)\}$ of $\mathbb{R}^n$ (see for instance [20, Ch. 2, Theorem 2.1]), and let us define

$$g(x)=\sum_{j=1}^{\infty}\psi_j(x)g_j(x).$$

Notice that

$$\{x\in\Omega_j: f(x)\neq g(x)\}\subseteq\bigcup_{i=j-1}^{j}\{x\in\Omega_j: f(x)\neq g_i(x)\}.$$

This implies that

$$\mathcal{L}^n\left(\{x\in\Omega: f(x)\neq g(x)\}\right)\leq 2\sum_{j=1}^{\infty}\mathcal{L}^n\left(\{x\in\Omega_j: f(x)\neq g_j(x)\}\right)\leq 2\sum_{j=1}^{\infty}\frac{\varepsilon}{6^j}\leq\varepsilon,$$

and concludes the proof of the Lemma. □

Now let us present the proof of Theorem 1. Let us call $N\subset\Omega$ the set of points for which (2) does not hold. Since N has measure zero, proving Lusin property of class C^1 for the restriction of f to $\Omega\setminus N$ would immediately lead to Lusin property of class C^1 for f. So we may and do assume in what follows that $N=\emptyset$, and in particular that

$$\liminf_{y\to x}\frac{f(y)-f(x)}{|y-x|}>-\infty$$

for every $x\in\Omega$. Note that this inequality implies that f is lower semicontinuous on Ω, and in particular f is measurable. Now, according to Lemma 2, it is enough to check that the restriction of f to every compact subset of Ω has the Lusin property of class C^1, and therefore we may also assume without loss of generality that Ω is compact. Define for each $j\in\mathbb{N}$,

$$E_j:=\left\{x\in\Omega:\ f(y)-f(x)\geq -j|y-x| \text{ for all } y\in B\left(x,\frac{1}{j}\right)\cap\Omega\right\}\cap\left\{x\in\Omega:\ |f(x)|\leq j\right\}.$$

Because f is lower semicontinuous the sets

$$\left\{x \in \Omega : \; f(y) - f(x) \geq -j|y-x| \text{ for all } y \in B\left(x, \frac{1}{j}\right) \cap \Omega\right\}$$

are closed, and by using the measurability of f this implies that each set E_j is measurable. These sets form an increasing sequence such that

$$\Omega = \bigcup_{j=1}^{\infty} E_j,$$

so we have

$$\lim_{j\to\infty} \mathcal{L}^n\left(\Omega \setminus E_j\right) = 0,$$

and therefore, for a given $\varepsilon > 0$ we may find $j_0 \in \mathbb{N}$ large enough such that $\mathcal{L}^n(\Omega \setminus E_{j_0}) < \frac{\varepsilon}{2}$. Take now $x, y \in E_{j_0}$. If $|y-x| \leq \frac{1}{j_0}$ then we have

$$|f(y) - f(x)| \leq j_0|y-x| \text{ and } |f(x)| \leq j_0.$$

On the other hand, if $x, y \in E_{j_0}$ and $|y-x| > 1/j_0$ then we trivially get

$$|f(y) - f(x)| \leq 2 \sup_{z\in E_{j_0}} |f(z)| \leq M_0|y-x|,$$

where $M_0 := 2j_0\left(1 + \sup_{z\in\Omega}|f(z)|\right)$.
Observe that $M_0 \geq j_0$. Thus in either case we see that

$$|f(y) - f(x)| \leq M_0|y-x| \text{ and } |f(x)| \leq M_0, \text{ for all } x, y \in E_{j_0}.$$

That is, f is bounded and M_0-Lipschitz on E_{j_0}. Then we can extend f to a Lipschitz function F on $\mathbb{R}^n$, for instance by using the McShane-Whitney formula

$$F(x) = \inf_{y\in E_{j_0}} \{f(y) + M_0|x-y|\},$$

which defines an M_0-Lipschitz function on $\mathbb{R}^n$ that coincides with f on E_{j_0}. Obviously we have

$$\mathcal{L}^n(\{x \in \Omega : \; f(x) \neq F(x)\}) \leq \mathcal{L}^n(\Omega \setminus E_{j_0}) < \frac{\varepsilon}{2}.$$

But according to the result of Federer's that we mentioned above (see also [14, Theorem 6.11]), Lipschitz functions have the C^1 Lusin property, so we may find another function $g \in C^1(\mathbb{R}^n)$ such that $\mathcal{L}^n(\{x \in \Omega : \; F(x) \neq g(x)\}) < \frac{\varepsilon}{2}$. Thus we conclude that

$$\begin{aligned}
&\mathcal{L}^n(\{x \in \Omega : \; f(x) \neq g(x)\}) = \\
&\qquad = \mathcal{L}^n(\{x \in E_{j_0} : \; F(x) \neq g(x)\} \cup \{x \in \Omega \setminus E_{j_0} : \; f(x) \neq g(x)\}) \leq \\
&\qquad\qquad \leq \mathcal{L}^n(\{x \in E_{j_0} : \; F(x) \neq g(x)\}) + \mathcal{L}^n(\Omega \setminus E_{j_0}) \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.
\end{aligned}$$

□

Corollary 3. *Let U be a measurable subset of $\mathbb{R}^n$, $f : U \to \mathbb{R}$ be a measurable function, and define $\Omega = \{x \in U : D^- f(x) \neq \emptyset\}$. Then for every $\varepsilon > 0$ there exists a function $g \in C^1(\mathbb{R}^n)$ such that*

$$\mathcal{L}^n\left(\{x \in \Omega : f(x) \neq g(x)\}\right) \leq \varepsilon.$$

Here $D^- f(x)$ denotes the Fréchet subdifferential of f at x, that is the set of vectors $\zeta \in \mathbb{R}^n$ such that

$$\liminf_{h\to 0} \frac{f(x+h)-f(x)-\langle \zeta, h\rangle}{|h|} \geq 0.$$

Remark 4. *In the above corollary we also have $D^- f(x) = \{\nabla g(x)\}$ for almost every $x \in \Omega$ with $f(x) = g(x)$.*

Proof. Almost every point of the set $A = \{x \in \Omega : f(x) = g(x)\}$ is a point of density 1 of A, and for every such point x and every $\xi_x \in D^- f(x)$ we have

$$0 \leq \liminf_{y\to x, y\in A} \frac{f(y)-f(x)-\langle \xi_x, y-x\rangle}{|y-x|} = \liminf_{y\to x, y\in A} \frac{g(y)-g(x)-\langle \xi_x, y-x\rangle}{|y-x|},$$

and

$$\lim_{y\to x, y\in A} \frac{g(y)-g(x)-\langle \nabla g(x), y-x\rangle}{|y-x|} = 0,$$

hence also

$$\liminf_{y\to x, y\in A} \frac{\langle \nabla g(x)-\xi_x, y-x\rangle}{|y-x|} \geq 0, \tag{3}$$

which, because x is a point of density 1 of A and $h \mapsto \langle \nabla g(x) - \xi_x, h\rangle$ is linear, implies that $\nabla g(x) = \xi_x$. Indeed, we have

$$\lim_{r\to 0^+} \frac{\mathcal{L}^n\left(A\cap B(x,r)\right)}{\mathcal{L}^n\left(B(x,r)\right)} = 1. \tag{4}$$

Assume we had $\zeta := \nabla g(x) - \xi_x \neq 0$, and consider the sets

$$S_\zeta := \{v \in \mathbb{R}^n \,:\, |v| = 1,\, \langle \zeta, v\rangle \leq -\frac{1}{2}|\zeta|\},$$

which determines a region of positive surface measure in the unit sphere, and the associated cone

$$C_{x,\zeta} = \{x + tv \,:\, v \in S_\zeta, t > 0\},$$

of which x is thus a point of positive density. Hence $C_{x,\zeta}$ also satisfies, in view of (4), that

$$\liminf_{r\to 0^+} \frac{\mathcal{L}^n\left(A\cap C_{x,\zeta}\cap B(x,r)\right)}{\mathcal{L}^n\left(B(x,r)\right)} > 0.$$

In particular there exists a sequence $(y_k) = (x + t_k v_k) \subset A \cap C_{x,\zeta}$ (with $t_k > 0$ and $v_k \in S_\zeta$, $k \in \mathbb{N}$) such that $\lim_{k\to\infty} y_k = x$. For this sequence we have, because of the definition of $C_{x,\zeta}$, that

$$\frac{\langle \nabla g(x) - \xi_x, y_k - x\rangle}{|y_k - x|} = \frac{\langle \zeta, t_k v_k\rangle}{t_k} \leq -\frac{1}{2}|\zeta| < 0$$

for all $k \in \mathbb{N}$, which contradicts (3). □

A natural question at this point is the following. Does Corollary 3 hold true if we replace the Frechet subdifferential by the *limiting subdifferential*? Let us recall that the limiting subdifferential $\partial_L f(x)$ of a lower semicontinuous function $f : \mathbb{R}^n \to \mathbb{R}$ at a point x consists of all vectors of the form $\zeta = \lim_n \zeta_n$, where $\zeta_n \in D^- f(x_n)$, for sequences $\{x_n\}$ satisfying $\lim_n x_n = x$, and $\lim_n f(x_n) = f(x)$; see [12, 16], for instance, for elementary properties of this subdifferential.

The question is whether or not the assumption that $\partial_L f(x) \neq \emptyset$ for every $x \in \mathbb{R}^n$ implies that f satisfies the Lusin property of order C^1. Since one trivially has that $D^- f(x) \subset \partial_L f(x)$, such a result would be much stronger than Corollary 3 above. The following example shows that the answer is negative.

Example 5. We consider the classical Takagi function $T : \mathbb{R} \to \mathbb{R}$ defined as follows. If D_n denotes the set of real numbers $\{\frac{k}{2^n} : k \in \mathbb{Z}\}$, and $d(x, D_n)$ is the distance of x to D_n, then

$$T(x) = \sum_{n=1}^{\infty} d(x, D_n)$$

This function was introduced by Takagi, [29], as an easy example of a continuous function which is nowhere differentiable. In [9, Theorem 2] it is proved that T does not agree with any C^1 function on any set of positive measure, and in particular T does not satisfy the Lusin property of order C^1. However, in [17, Corollary 1.4], and also implicitly in [18], it is proved that $\partial_L T(x) = \mathbb{R}$ for every $x \in \mathbb{R}$.

Concerning the Lusin property of class C^2 we have the following result.

Theorem 6. *Let $\Omega \subset \mathbb{R}^n$ be a Lebesgue measurable set, and $f : \Omega \to \mathbb{R}$ be a function such that for almost every $x \in \Omega$ there exists a vector $\xi_x \in \mathbb{R}^n$ such that*

$$\liminf_{y \to x} \frac{f(y) - f(x) - \langle \xi_x, y - x \rangle}{|y - x|^2} > -\infty. \tag{5}$$

Then for every $\varepsilon > 0$ there exists a function $g \in C^2(\mathbb{R}^n)$ such that

$$\mathcal{L}^n \left(\{ x \in \Omega : f(x) \neq g(x) \} \right) \leq \varepsilon.$$

Proof. Let N be the subset of points for which (5) does not hold, and put $\Omega_1 = \Omega \setminus N$. Since N has measure zero, it will be enough to show that the restriction f_1 of f to Ω_1 has the Lusin property of class C^2. Since (5) holds for every $x \in \Omega_1$, it follows that f is lower semicontinuous on Ω_1, and in particular f_1 is measurable (hence so is f, since N has measure zero). Now, according to Lemma 2, if we take an arbitrary compact subset Ω_2 of Ω_1, it will be enough for us to check that the restriction f_2 of f_1 to Ω_2 has the Lusin property of class C^2.
Because (5) holds for every $x \in \Omega_2$ and this implies

$$\liminf_{y \to x} \frac{f_2(y) - f_2(x)}{|y - x|} > -\infty$$

for all $x \in \Omega_2$, given $\varepsilon > 0$, we may apply Theorem 1 to get a function $g \in C^1(\mathbb{R}^n)$ such that

$$\mathcal{L}^n(\{ x \in \Omega_2 \, : \, f_2(x) \neq g(x) \}) \leq \frac{\varepsilon}{4}.$$

Observe also that the set $A = \{ x \in \Omega_2 : f_2(x) = g(x) \}$ is measurable and bounded, and according to the preceding remark we have $\xi_x = \nabla g(x)$ for almost every $x \in A$, so we can find a compact subset Ω_3 of A such that $\mathcal{L}^n (A \setminus \Omega_3) \leq \varepsilon/4$ and $\xi_x = \nabla g(x)$ for all $x \in \Omega_3$. Then we have that

$$\liminf_{y \to x, y \in \Omega_3} \frac{g(y) - g(x) - \langle \nabla g(x), y - x \rangle}{|y - x|^2} > -\infty \tag{6}$$

for every $x \in \Omega_3$. Now let us define for each $j \in \mathbb{N}$

$$E_j := \left\{x \in \Omega_3 \,:\, g(y) - \langle \nabla g(x), y\rangle \geq g(x) - \langle \nabla g(x), x\rangle - j|y-x|^2 \text{ for all } y \in \Omega_3\right\},$$

and note that the sets E_j are measurable and increasing to Ω_3. There exists $j_0 \in \mathbb{N}$ such that

$$\mathcal{L}^n(\Omega_3 \setminus E_{j_0}) \leq \frac{\varepsilon}{4}.$$

It will be enough for us to prove the following:

Claim 7. *We have that*

$$\limsup_{y\to x,\, y\in E_{j_0}} \frac{|g(y)-g(x)-\langle \nabla g(x), y-x\rangle|}{|y-x|^2} < +\infty$$

for almost every $x \in E_{j_0}$.

Assume for a moment that the Claim is true, that is, the restriction of g to E_{j_0} has an approximate $(2-1)$-Taylor polynomial at every $x \in E_{j_0}$. By [26, Theorem 1] this is equivalent to saying that the restriction of g to E_{j_0} has the Lusin property of class C^2. So we may find a function $h \in C^2(\mathbb{R}^n;\mathbb{R})$ such that

$$\mathcal{L}^n(\{x \in E_{j_0}\,;\, g(x) \neq h(x)\}) \leq \frac{\varepsilon}{4},$$

and we easily conclude that

$$\mathcal{L}^n(\{x \in \Omega_2 :\, f_2(x) \neq h(x)\}) \leq \varepsilon,$$

as we wanted to show.

In order to prove Claim (7) we will borrow some ideas from [24]. We define new functions $\widetilde{g}:\mathbb{R}^n \to \mathbb{R}$ and $\hat{g}:\mathbb{R}^n\to\mathbb{R}$ by

$$\begin{aligned} &\widetilde{g}(x) = g(x) + j_0|x|^2, && x \in \mathbb{R}^n\\ &\hat{g}(x) = \sup\{p(x) :\, p \text{ affine and } p \leq \widetilde{g} \text{ on } \Omega_3\}, && x \in \mathbb{R}^n \end{aligned}$$

By definition of E_{j_0} we have $\widetilde{g}(y) \geq \widetilde{g}(x) + \langle \nabla\widetilde{g}(x), y-x\rangle$ for all $y \in \Omega_3, x \in E_{j_0}$, and by using this inequality it is easy to see that

$$\widetilde{g}(x) = \hat{g}(x)$$

for all $x \in E_{j_0}$. On the other hand, since Ω_3 is compact and g is continuous on Ω_3, it is easy to see that $\hat{g}$ is everywhere finite. Moreover, as a supremum of affine functions, $\hat{g}$ is convex. Therefore $\hat{g}$ is locally Lipschitz on Ω_3. Also g is of class C^1, hence so is $\widetilde{g}$. Since the functions $\widetilde{g}$ and $\hat{g}$ agree on E_{j_0}, we then also have that

$$\nabla\hat{g}(x) = \nabla\widetilde{g}(x)$$

for almost every $x \in E_{j_0}$ (see [14, Theorem 3.3(i)] for instance).

Next, by applying Alexandroff's theorem [3] (see also [10] in dimension 2) with the convex function $\hat{g}$, we obtain that $\hat{g}$ is twice differentiable almost everywhere in Ω_3. This implies that

$$\limsup_{y\to x,\, y\in E_{j_0}} \frac{|\widetilde{g}(y)-\widetilde{g}(x)-\langle \nabla\widetilde{g}(x), y-x\rangle|}{|y-x|^2} = \tag{7}$$

$$= \limsup_{y\to x,\, y\in E_{j_0}} \frac{|\hat{g}(y)-\hat{g}(x)-\langle \nabla\hat{g}(x), y-x\rangle|}{|y-x|^2} < +\infty$$

for almost every $x \in E_{j_0}$. However, by the definition of $\widetilde{g}(x) = g(x) + j_0|x|^2$, we have

$$\frac{|g(y) - g(x) - \langle \nabla g(x), y - x\rangle|}{|y-x|^2} \leq$$
$$\leq \frac{|g(y) - g(x) - \langle \nabla g(x), y - x\rangle + (j_0(|y|^2 + |x|^2 - 2\langle x, y\rangle)|}{|y-x|^2} + j_0 =$$
$$= \frac{|\widetilde{g}(y) - \widetilde{g}(x) - \langle \nabla \widetilde{g}(x), y - x\rangle|}{|y-x|^2} + j_0,$$

and by combining with (7) we immediately obtain Claim (7). □

Corollary 8. *Let $\Omega \subset \mathbb{R}^n$ be a Lebesgue measurable set, and $f : \Omega \to \mathbb{R}$ be a function such that for almost every $x \in \Omega$ there exists a vector $\xi_x \in \mathbb{R}^n$ such that*

$$\limsup_{y \to x} \frac{f(y) - f(x) - \langle \xi_x, y - x\rangle}{|y-x|^2} < +\infty. \tag{8}$$

Then for every $\varepsilon > 0$ there exists a function $g \in C^2(\mathbb{R}^n)$ such that

$$\mathcal{L}^n\left(\{x \in \Omega : f(x) \neq g(x)\}\right) \leq \varepsilon.$$

This is of course an immediate consequence of Theorem 6 applied to $-f$.
According to Remark 4, we also have that

$$\xi_x = \nabla g(x)$$

for almost every $x \in \Omega$ with $f(x) = g(x)$.

Corollary 9. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a measurable function, and define $\Omega = \{x \in \mathbb{R}^n : \partial_P f(x) \neq \emptyset\}$. Then for every $\varepsilon > 0$ there exists a function $g \in C^2(\mathbb{R}^n)$ such that*

$$\mathcal{L}^n\left(\{x \in \Omega : f(x) \neq g(x)\}\right) \leq \varepsilon.$$

Here $\partial_P f(x)$ denotes the proximal subdifferential of f at x, which is defined as the set of all $\zeta \in \mathbb{R}^n$ for which there exist $\sigma, \eta > 0$ such that

$$f(y) \geq f(x) + \langle \zeta, y - x\rangle - \sigma|y-x|^2$$

for all $y \in B(x, \eta)$. The set $\partial_P f(x)$ coincides with $\{\zeta \in \mathbb{R}^n : \zeta = \nabla\varphi(x), \varphi \in C^2(\mathbb{R}^n), f - \varphi$ attains a minimum at $x\}$, so every function f for which the viscosity subdifferential of second order is nonempty at x also has a nonempty proximal subdifferential at x. The set $\partial_P f(x)$ can also be equivalently defined as the set of vectors $\zeta \in \mathbb{R}^n$ such that

$$\liminf_{h \to 0} \frac{f(x+h) - f(x) - \langle \zeta, h\rangle}{|h|^2} > -\infty,$$

so it is clear that the above Corollary is an immediate consequence of Theorem 6. Notice also that this corollary allows us to recover, with a different proof, the mentioned result for convex functions established independently by Alberti [2] and Imonkulov [21].

Let us finally present two examples. The first one concerns the following matter: one could erroneously think that if a function f satisfies (5) then f will automatically satisfy

$$\limsup_{y \to x} \frac{f(y) - f(x) - \langle \xi_x, y - x\rangle}{|y-x|^2} < +\infty \tag{9}$$

for almost every $x \in \Omega$ as well, and then one could immediately apply Liu-Tai's theorem [26] to conclude the proof of Theorem 6. This is not feasible.

Example 10. Let us first consider a Cantor set of positive measure, $C \subset [0,1]$. More precisely,

$$C = [0,1] \setminus \bigcup_n J_n$$

where each J_n is the union of 2^{n-1} disjoint intervals of length $\frac{1}{4^n}$ and $J_n \cap J_m = \emptyset$ for $n \neq m$.

$$J_n = \bigcup_{k=1}^{2^{n-1}} (a_n^k, b_n^k),$$

where $b_n^k < a_n^{k+1}$ for $k < 2^{n-1}$. Let us inductively construct the sets J_n. Setting $J_1 = (\frac{3}{8}, \frac{5}{8})$, if $n \geq 1$, we assume that $J_1, \ldots, J_n$ satisfy that

$$[0,1] \setminus \bigcup_{k=1}^{n} J_k$$

consists in 2^n disjoint intervals of length $\frac{1}{2^{n+1}} + \frac{1}{2^{2n+1}}$, because

$$\mathcal{L}\big([0,1] \setminus \bigcup_{k=1}^{n} J_k\big) = 1 - \sum_{k=1}^{n} \frac{2^{k-1}}{4^k} = 1 - \frac{1}{2}(1 - \frac{1}{2^n}) = \frac{1}{2} + \frac{1}{2^{n+1}}.$$

For each of these intervals composing $[0,1] \setminus \bigcup_{k=1}^{n} J_k$, we consider a subinterval, centered at the corresponding middle point, of length $\frac{1}{4^{n+1}}$. Then J_{n+1} will be the union of these subintervals. It is clear that $\mathcal{L}(C) = \frac{1}{2}$.

Now let us define a function f in the following way: we set

$$f(x) = 0 \quad \text{for every} \quad x \in C,$$

while for every $n \in \mathbb{N}$ and $k = 1, \ldots, 2^{n-1}$, $f : [a_n^k, b_n^k] \to \mathbb{R}$ will be a non negative continuous function such that $f : (a_n^k, b_n^k) \to \mathbb{R}$ is C^∞,

$$\max_{x \in I_n^k} f(x) = f(a_n^k + \frac{1}{2}(b_n^k - a_n^k)) = \frac{1}{2^n},$$

and such that f, as well as all its one-sided derivatives, equal 0 at a_n^k and at b_n^k. It is clear that f is continuous. Let us denote

$$\Delta_x(y) = \frac{f(y) - f(x) - \xi_x(y-x)}{|y-x|^2}.$$

If $x \notin C$ then, taking $\xi_x = f'(x)$, we have $\lim_{y \to x} \Delta_x(y) = \frac{1}{2} f''(x)$. If $x \in C$, then

$$\frac{f(y) - f(x)}{|y-x|^2} \geq 0.$$

Hence for every x there exists ξ_x such that

$$\liminf_{y \to x} \Delta_x(y) > -\infty.$$

Let us observe that f also satisfies conditions of the form

$$\liminf_{y\to x}\frac{f(y)-P(y-x)}{|y-x|^k}>-\infty,$$

where P is a polynomial of degree $k-1$ for every k.
Now let $\tilde{C}=C\setminus(\{0,1\}\cup\{a_n^k,b_n^k\}_{n,k})$. We claim that

$$\limsup_{y\to x}\Delta_x(y)=+\infty$$

for every $x\in\tilde{C}$ and every ξ_x. Let us prove this. If $x\in\tilde{C}$ there exist subsequences $\{a_{m_j}^{r_j}\}_j$ and $\{b_{n_j}^{k_j}\}_j$, decreasing and increasing respectively, such that

$$\lim_j a_{m_j}^{r_j}=\lim_j b_{n_j}^{k_j}=x.$$

More precisely, we chose $a_{m_j}^{r_j}$ such that

$$0<a_{m_j}^{r_j}-x\le\frac{1}{2^{m_j+1}}+\frac{1}{2^{2m_j+1}},$$

and $b_{n_j}^{k_j}$ such that

$$0<x-b_{n_j}^{k_j}\le\frac{1}{2^{n_j+1}}+\frac{1}{2^{2n_j+1}}.$$

Let us consider the case that $\xi_x\ge 0$. We take $y_j=b_{n_j}^{k_j}-\frac{1}{2}(b_{n_j}^{k_j}-a_{n_j}^{k_j})$. We have

$$\Delta_x(y_j)\ge\frac{f(y_j)}{|y_j-x|^2}=\frac{1}{2^{n_j}}\frac{1}{|y_j-x|^2}\ge 2^{n_j}$$

since $|y_j-x|\le\frac{1}{2^{n_j}}$. In particular we obtain that $\limsup_{y\to x}\Delta_x(y)=+\infty$.
The case $\xi_x\le 0$ can be dealt with similarly by considering $y_j=a_{m_j}^{r_j}+\frac{1}{2}(b_{m_j}^{r_j}-a_{m_j}^{r_j})$. $\square$

Our second example shows that there are no analogues of Theorem 6 for higher order of differentiability.

Example 11. Let $f:\mathbb{R}\to\mathbb{R}$ be the function given by

$$f(x)=\frac{1}{\pi^2}\sum_{n=1}^{\infty}2^{-3n}\cos\left(2^n\pi x\right).$$

This is a C^2 function such that f'' is not differentiable at any point (see [19]) and

$$\limsup_{|y|\to 0}\frac{|f''(x+y)+f''(x-y)-2f''(x)|}{|y|}<+\infty$$

for every $x\in\mathbb{R}$ (see [Stein(1970), p. 148]). By [26, Theorem 4] f'' is not approximately differentiable on a set of positive measure.
For every x, we have that

$$\lim_{y\to x}\frac{f(y)-f(x)-f'(x)(y-x)-\frac{1}{2}f''(x)(y-x)^2}{|y-x|^2}=0$$

If $a > 0$ we have
$$\liminf_{y\to x} \frac{f(y)-f(x)-f'(x)(y-x)-(\frac{1}{2}f''(x)-a)(y-x)^2}{|y-x|^2} > 0,$$
hence
$$\liminf_{y\to x} \frac{f(y)-f(x)-f'(x)(y-x)-(\frac{1}{2}f''(x)-a)(y-x)^2}{|y-x|^k} = \infty > -\infty$$
for every $k > 2$. If an analogue of Theorem 6 for some order $k > 2$ were true for this function f, then, according to Liu-Tai's characterization of Lusin properties and approximate differentiability of higher order [26], we would have that f is approximately differentiable of order k. However, in [26, p. 194] it is shown that the coefficients of order j of the Taylor expansion of an approximately differentiable function of order k coincide, up to sets of arbitrarily small measure, with derivatives of order j of C^k functions; in particular those coefficients have the Lusin property of class C^{k-j} and therefore, again by [26, Theorem 1], they are almost everywhere approximately differentiable of order $k-j$. This would imply that f'' is approximately differentiable almost everywhere, which we know to be false.
Another example can be given by taking $g : \mathbb{R} \to \mathbb{R}$ to be a continuous function which is nowhere approximately differentiable (see [23, Chapter 6]), setting
$$f(x) = \int_0^x \left(\int_0^t g(s)ds\right) dt,$$
and repeating the preceding argument word by word. One could also use as g the Takagi function of Example 5, which by [9, Theorem 2] and [26] is not approximately differentiable on any set of positive measure.

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] 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.

[4] D. Azagra and P. Hajłasz *Lusin-type properties of convex functions*, preprint, 2017.

[5] T. Bagby and W.P. Ziemer, *Pointwise differentiability and absolute continuity*, Trans. Amer. Math. Soc. 191 (1974), 129–148.

[6] B. Bojarski and P. Hajłasz, *Pointwise inequalities for Sobolev functions*, Studia Math. 106 (1993), 77–92.

[7] B. Bojarski, P. Hajłasz, 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.

[8] 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.

[9] J.B. Brown and G. Kozlowski, *Smooth interpolation, HÃűlder continuity, and the TakagiâĂŞvan der Waerden function.* Amer. Math. Monthly 110 (2003), no. 2, 142–147.

[10] H. Busemann and W. Feller, *KrÃijmmungseigenschaften Konvexer FlÃd'chen.* (German) Acta Math. 66 (1936), no. 1, 1–47.

[11] A. P. Calderón and A. Zygmund, *Local properties of solutions of elliptic partial differential equations*, Studia Math. 20 (1961), 171–225.

[12] F.H. Clarke, Yu.S. Ledyaev, R.J. Stern, and P.R. Wolenski, *Nonsmooth Analysis and Control Theory.* Grad. Texts in Math. 178, Springer, 1998.

[13] M.G. Crandall, H. Ishii, P.-L. Lions, *User's guide to viscosity solutions of second order partial differential equations*, Bull. Amer. Math. Soc. 27 (1992) 1–67.
[14] L.C. Evans, R.F. Gariepy, *Measure theory and fine properties of functions*. Revised edition. Textbooks in Mathematics. CRC Press, Boca Raton, FL, 2015.
[15] H. Federer, *Surface area. II*, Trans. Amer. Math. Soc. 55, (1944), 438–456.
[16] J. Ferrera, *An introduction to nonsmooth analysis*. Elsevier/Academic Press, Amsterdam, 2014.
[17] J. Ferrera and J. Gomez-Gil, *Generalized Takagi-Van der Waerden functions and their subdifferentials*, J. Convex Anal. 25 (4) (2018) (to appear).
[18] P. Gora and R.J. Stern, *Subdifferential analysis of the Van der Waerden function*, J. Convex Anal. 18 (3) (2011), 699-705.
[19] G.H. Hardy, *Weierstrass's non-differentiable function.* Trans. Amer. Math. Soc. 17 (1916), no. 3, 301–325.
[20] M.W. Hirsch, *Differential topology*. Graduate Texts in Mathematics, 33. Springer-Verlag, New York, 1976.
[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] A.B. Kharazishvili, *Strange functions in real analysis*. Second edition. Pure and Applied Mathematics (Boca Raton), 272. Chapman & Hall/CRC, Boca Raton, FL, 2006.
[24] M. Kocan and X.-J. Wang, *On the generalized Stepanov theorem*, Proc. Amer. Math. Soc. 125 (1997), 2347–2352.
[25] Fon-Che Liu, *A Luzin type property of Sobolev functions*, Indiana Univ. Math. J. 26 (1977), 645–651.
[26] F.-C. Liu and W.-S. Tai, *Approximate Taylor polynomials and differentiation of functions*, Topol. Methods Nonlinear Anal. 3 (1994), no. 1, 189–196.
[27] N. Lusin, *Sur les propiÃl'tÃl's des fonctions measurables*, Comptes Rendus Acad. Sci. Paris 154 (1912), 1688–1690.
[28] J. Michael and W.P. Ziemer, *A Lusin type approximation of Sobolev functions by smooth functions*, Contemp. Math. 42 (1985), 135–167.
[Stein(1970)] E.M. Stein, *Singular integrals and differentiability properties of functions*, Princeton Mathematical Series, No. 30, Princeton University Press, Princeton, N.J., 1970.
[29] T. Takagi, *A simple example of the continuous function without derivative*, Proc. Phys. Math. Soc. Tokio Ser. II 1 (1903), 176-177.
[30] H. Whitney, *Analytic extensions of differentiable functions defined in closed sets*, Trans. Amer. Math. Soc. 36 (1934), 63–89.
[31] H. Whitney, *On totally differentiable and smooth functions*, Pacific J. Math. 1, (1951). 143–159.
[32] W. P. Ziemer, *Weakly Differentiable Functions*, Springer–Verlag, 1989.

ICMAT (CSIC-UAM-UC3-UCM), Departamento de Análisis Matemático, Facultad Ciencias Matemáticas, Universidad Complutense, 28040, Madrid, Spain
E-mail address: azagra@mat.ucm.es

IMI, Departamento de Análisis Matemático, Facultad Ciencias Matemáticas, Universidad Complutense, 28040, Madrid, Spain
E-mail address: ferrera@mat.ucm.es

ICMAT (CSIC-UAM-UC3-UCM), Calle Nicolás Cabrera 13-15. 28049 Madrid, Spain
E-mail address: miguel.garcia@icmat.es

Departamento de Análisis Matemático, Facultad Ciencias Matemáticas, Universidad Complutense, 28040, Madrid, Spain
E-mail address: gomezgil@mat.ucm.es