APPROXIMATION AND REGULARIZATION OF LIPSCHITZ FUNCTIONS: CONVERGENCE OF THE GRADIENTS

MARC-OLIVIER CZARNECKI AND LUDOVIC RIFFORD

ABSTRACT. We examine the possible extensions to the Lipschitzian setting of the classical result on C^1-convergence: first (approximation), if a sequence (f_n) of functions of class C^1 from $\mathbb{R}^N$ to $\mathbb{R}$ converges uniformly to a function f of class C^1, then the gradient of f is a limit of gradients of f_n in the sense that $\mathrm{graph}(\nabla f) \subset \liminf_{n \to +\infty} \mathrm{graph}(\nabla f_n)$; second (regularization), the functions (f_n) can be chosen to be of class C^∞ and C^1-converging to f in the sense that $\lim_{n \to +\infty} \|f_n - f\|_\infty + \|\nabla f_n - \nabla f\|_\infty = 0$. In other words, the space of C^∞ functions is dense in the space of C^1 functions endowed with the C^1 pseudo-norm.

We first deepen the properties of Warga's counterexample (1981) for the extension of the approximation part to the Lipschitzian setting. This part cannot be extended, even if one restricts the approximation schemes to the classical convolution and the Lasry-Lions regularization. We thus make more precise various results in the literature on the convergence of subdifferentials.

We then show that the regularization part can be extended to the Lipschitzian setting, namely if $f : \mathbb{R}^N \to \mathbb{R}$ is a locally Lipschitz function, we build a sequence of smooth functions $(f_n)_{n \in \mathbb{N}}$ such that

$$\lim_{n \to +\infty} \|f_n - f\|_\infty = 0,$$
$$\lim_{n \to +\infty} d_{Haus}(\mathrm{graph}(\nabla f_n), \mathrm{graph}(\partial f)) = 0.$$

In other words, the space of C^∞ functions is dense in the space of locally Lipschitz functions endowed with an appropriate Lipschitz pseudo-distance. Up to now, Rockafellar and Wets (1998) have shown that the convolution procedure permits us to have the equality $\limsup_{n \to +\infty} \mathrm{graph}(\nabla f_n) = \mathrm{graph}(\partial f)$, which cannot provide the exactness of our result.

As a consequence, we obtain a similar result on the regularization of epi-Lipschitz sets. With both functional and set parts, we improve previous results in the literature on the regularization of functions and sets.

1. INTRODUCTION

The purpose of this paper is to examine the extension to the Lipschitzian setting of the following classical result on C^1-convergence, and thus to study the convergence properties of Clarke generalized gradients and Clarke normal cones, together with other related notions of (generalized) gradients and normal cones. In the smooth case, it is well known that if a C^1 function f is the limit of a sequence $(f_n)_{n \in \mathbb{N}}$ of C^1 functions, then every gradient of f can be seen as the limit of some

gradients of the f_n's. Furthermore every C^1 function can be approached by smooth functions for the C^1-convergence. We state the results in the next proposition.

Proposition A. *Let $f : \mathbb{R}^N \to \mathbb{R}$ be a function of class C^1:*
   (a) *For every sequence $(f_n)_{n \in \mathbb{N}}$ in $C^1(\mathbb{R}^N, \mathbb{R})$ converging uniformly to f, we have*

$$\operatorname{graph}(\nabla f) \subset \liminf_{n \to +\infty} \operatorname{graph}(\nabla f_n).$$

   (b) *(Regularization) There exists a sequence $(f_n)_{n \in \mathbb{N}}$ in $C^\infty(\mathbb{R}^N, \mathbb{R})$ such that*

$$\lim_{n \to +\infty} \|f_n - f\|_\infty + \|\nabla f_n - \nabla f\|_\infty = 0.$$

In particular

$$\operatorname{graph}(\nabla f) = \liminf_{n \to +\infty} \operatorname{graph}(\nabla f_n) = \limsup_{n \to +\infty} \operatorname{graph}(\nabla f_n).$$

   In fact, although they are apparently independent, part (b) is a consequence of (a) by a classical convolution-type argument. In the present paper, our aim is to understand how Proposition A can be extended to the case of locally Lipschitz functions. Of course, in this setting we have to specify what kind of "gradient" we consider.
   In the convex setting, i.e., when $(f_n)_{n \in \mathbb{N}}$ is a sequence of lower semi-continuous convex functions on $\mathbb{R}^N$ epi-converging (or Γ-converging) to a convex function f, Attouch's theorem (see [1] and [2]) provides a global extension to both parts (a) and (b). In fact, Attouch's theorem establishes in reflexive Banach space the general equivalence of Mosco convergence of the functions and set-convergence of the graph of their subdifferential. A rich literature has since been developed to extend and generalize Attouch's theorem; see for example Attouch and Beer [3], Poliquin [29], Levy, Poliquin and Thibault [25], Zolezzi [38] to quote only some of them. Note also that outside the convex setting, there is no chance of obtaining a similar statement to Attouch's theorem in its full generality. Indeed, it is easy to build a C^∞ approximation of a function f with gradients far from the (generalized) gradients of the initial function.
   This last phenomenon highlights the difference between our point and an approach *à la* Attouch's theorem, which is not strictly a regularization. In the Attouch's theorem direction, the question would rather be to identify a class of Lipschitz functions for which epi-convergence of the sequence (f_n) to f would be equivalent to the convergence of the graph of the gradients (see Zolezzi [38]).
   Let us now explain why we focus on the Lipschitzian setting for possible extensions of Proposition A. It is easy to prove that if a convex function is the uniform limit of a sequence of locally Lipschitz functions, then (a) holds for a Clarke generalized gradient (we will recall the notions of nonsmooth analysis that we use in the next section). As we shall see, this result has a geometric meaning. Note that the convex case permits us to solve our problem in the case of semiconvex (or semiconcave) functions. As a matter of fact since these functions are locally the sum of a convex function and of a smooth function, Proposition A combined with the result of the convex framework implies that if a sequence $(f_n)_{n \in \mathbb{N}}$ of locally Lipschitz functions on $\mathbb{R}^N$ converges to some semiconvex function, then (a) holds for Clarke generalized gradients. This result is also a consequence of the fact that Proposition A(a) remains true for Clarke gradients if we assume that f is C^1 and

that the f_n's are locally Lipschitz. It is in fact an immediate consequence of the following.

Proposition B. *Let $(M_n)_{n\in\mathbb{N}}$ be a sequence of closed subsets of $\mathbb{R}^N$, and let $M = \limsup_{n\to\infty} M_n = \liminf_{n\to\infty} M_n$. Then*

$$\mathrm{graph}\,(N_M^P) \subset \liminf_{n\to\infty} \mathrm{graph}\,(N_{M_n}^P).$$

From this result, one easily deduces that Proposition A(a) can be extended for lower semicontinuous functions and their proximal (also Dini's/Fréchet and limiting) subgradients. Lebourg [24] gave this extension for the Fréchet subgradients (of f). One easily sees that Proposition A(b) may not hold in this case, simply because the graph of the proximal subgradient may not be closed, and because of the nonconvexity of the limiting subgradient. Simple counterexamples also show that if either f or f_n is not assumed to be Lipschitz, Proposition A(a) may not hold for the Clarke subgradient. These considerations naturally lead to considering the Lipschitz function with Clarke's subgradient. Precisely,

Question 1. Let $f : \mathbb{R}^N \to \mathbb{R}$ be a locally Lipschitz function.
(a) For every sequence $(f_n)_{n\in\mathbb{N}}$ in $Lip_{loc}(\mathbb{R}^N, \mathbb{R})$ epi-converging (or Γ-converging) to f, do we have

$$\mathrm{graph}\,(\partial f) \subset \liminf_{n\to+\infty} \mathrm{graph}\,(\partial f_n)?$$

(b) (Regularization) Does there exist a sequence $(f_n)_{n\in\mathbb{N}}$ in $C^\infty(\mathbb{R}^N, \mathbb{R})$ converging uniformly to f, such that

$$\mathrm{graph}\,(\partial f) = \liminf_{n\to+\infty} \mathrm{graph}\,(\nabla f_n) = \limsup_{n\to+\infty} \mathrm{graph}\,(\nabla f_n)?$$

In terms of convergence of sets and their Clarke's normal cones, the formulation of Question 1 becomes:

Question 2. Let M be a closed epi-Lipschitz subset of $\mathbb{R}^N$.
(a) For every sequence $(M_n)_{n\in\mathbb{N}}$ of closed epi-Lipschitz subsets of $\mathbb{R}^N$ converging to M, do we have

$$\mathrm{graph}\,(N_M^C) \subset \liminf_{n\to\infty} \mathrm{graph}\,(N_{M_n}^C)?$$

(b) Does there exist a sequence $(M_n)_{n\in\mathbb{N}}$ of smooth sets converging to M for the Hausdorff distance, such that

$$\mathrm{graph}\,(N_M^C) = \liminf_{n\to+\infty} \mathrm{graph}\,(N_{M_n}^C) = \limsup_{n\to+\infty} \mathrm{graph}\,(N_{M_n}^C)?$$

These two questions (also only in terms of $\limsup$) were long-standing questions after the results of Ioffe [19], Mordukhovich [27], Kruger and Mordukhovich [21] on the convergence of functions and their gradients. Warga [37] provided an example of a function $f : \mathbb{R}^2 \to \mathbb{R}$ and a sequence $(f_n)_{n\in\mathbb{N}}$ of functions of class C^∞ which give a negative answer to Question 1(a) and its consequent Question 2(a) in terms of sets. However, one may wonder if Question 1(a) would have a positive answer with the Lasry-Lions regularization, in view of its many nice properties. This is not the case, and the function f exhibited by Warga again provides the counterexample.

Theorem A. *There exists a Lipschitz function $f : \mathbb{R}^2 \to \mathbb{R}$, such that*

$$\mathrm{graph}\,(\partial f) \not\subset \liminf_{\lambda,\mu\to 0} \mathrm{graph}\,(\nabla f_{\lambda,\mu}),$$

where $f_{\lambda,\mu}$ is the Lasry-Lions regularization of f, for $0 < \mu < \lambda$.

In fact, we show that the Lasry-Lions regularization of Warga's function f gives the smallest container of f, i.e., if (f_n) is a sequence of smooth functions which uniformly converges to this function f, then

$$\liminf_{\lambda,\mu\to 0} \text{graph}\,(\nabla f_{\lambda,\mu}) \subset \liminf_{n\to+\infty} \text{graph}\,(\nabla f_n).$$

Theorem A adds precision to Benoist [5] and Georgiev and Zlateva [18]. Let us also mention that Question 1(a) has a positive answer when $N = 1$ and also its consequent Question 2(a) in terms of sets in dimension 2. As a direct by-product of the example of Warga, we obtain an example of strict inclusion between two different kinds of normal cones, thus showing additional interest of [14, Theorem 2.2] (see section 3.6).

Furthermore, we give a positive answer to Question 1(b) with the following result.

Theorem B. *Let Ω be an open subset of $\mathbb{R}^N$ and let $f : \Omega \to \mathbb{R}$ be a locally Lipschitz function. Then there exists a sequence of smooth functions $(f_n)_{n\in\mathbb{N}}$ such that*

(1.1)
$$\lim_{n\to+\infty} \|f_n - f\|_\infty = 0,$$

(1.2)
$$\lim_{n\to+\infty} d_{Haus}(\text{graph}\,(\nabla f_n), \text{graph}\,(\partial f)) = 0.$$

In particular

$$\text{graph}\,(\partial f) = \liminf_{n\to+\infty} \text{graph}\,(\nabla f_n) = \limsup_{n\to+\infty} \text{graph}\,(\nabla f_n).$$

Theorem B can be viewed as a density result: endow the space $Lip_{loc}(\Omega, \mathbb{R})$ with the pseudo-distance defined by

$$d_{Lip}(f,g) := \|f - g\|_\infty + d_{Haus}(\text{graph}\,(\partial g), \text{graph}\,(\partial g)).$$

Then Theorem B can be expressed as the density of functions of class C^∞ for the corresponding topology:

$$Lip_{loc}(\Omega, \mathbb{R}) = \overline{C^\infty(\Omega, \mathbb{R})}.$$

The inclusion $\limsup_{n\to+\infty} \text{graph}\,(\nabla f_n) \subset \text{graph}\,(\partial f)$ can easily be obtained by the classical convolution procedure. Rockafellar and Wets [35, Theorem 9.67] show furthermore, up to a density argument (see Lemma 8.2 below), that one can have the equality $\limsup_{n\to+\infty} \text{graph}\,(\nabla f_n) = \text{graph}\,(\partial f)$ with the convolution procedure. They achieve this by taking convex combinations of mollifiers. However, one cannot deduce from their result information on the set $\liminf_{n\to+\infty} \text{graph}\,(\nabla f_n)$, and hence on the complete set convergence of the sets $\text{graph}\,(\nabla f_n)$.

As a consequence of our result, we improve (see Section 3.4) the convergence results of [15], hence of [11] and [16], where only one side of the convergence of the gradients was obtained. We can state Theorem B in terms of sets.

Theorem C. *Let M be a closed epi-Lipschitz subset of $\mathbb{R}^N$. Then there exists a sequence of smooth sets $(M_n)_{n\in\mathbb{N}}$ such that*

(1.3)
$$\lim_{n\to+\infty} d_{Haus}(M, M_n) = 0,$$

(1.4)
$$\lim_{n\to+\infty} d_{Haus}(\text{graph}\,(N^P_{M_n} \cap S), \text{graph}\,(N^C_M \cap S)) = 0.$$

In particular

$$\text{graph}\,(N^C_M) = \liminf_{n\to+\infty} \text{graph}\,(N^C_{M_n}) = \limsup_{n\to+\infty} \text{graph}\,(N^C_{M_n}).$$

As a consequence we improve the approximation results on epi-Lipschitz sets of [12]. We will also make more precise the approximation results on proximally nondegenerate sets of [14].

2. Main results

2.1. **Basic definitions in nonsmooth analysis.**[1] We recall the following definitions of the *Painlevé set-limit* associated with a net $(M_n)_{n\in\mathcal{N}}$ of subsets of $\mathbb{R}^N$, where $\mathcal{N}$ is a directed set (see, for example, Kuratowski [22]):

$$\liminf M_n := \{x \in \mathbb{R}^N | \exists (x_n)_{n\in\mathcal{N}} \subset \mathbb{R}^N, x_n \to x, x_n \in M_n \text{ for all } n \text{ large enough}\},$$

$$\limsup M_n := \{x \in \mathbb{R}^N | \exists (x_n)_{n\in\mathcal{N}} \subset \mathbb{R}^N, \exists \varphi \in \mathcal{I}, x_n \to x, x_n \in M_{\varphi(n)} \text{ for all } n\},$$

where $\mathcal{I}$ is the set of all maps $\varphi : \mathcal{N} \to \mathcal{N}$ such that $\varphi(n) \geq n$ for every $n \in \mathcal{N}$. When $\mathcal{N} = \mathbb{N}$, we write $\liminf_{n\to+\infty} M_n$ and $\limsup_{n\to+\infty} M_n$. Intimately related with the notion of the Painlevé set-limit is the *Hausdorff "distance"*. Let M and M' be two nonempty closed subsets of $\mathbb{R}^N$,

$$d_{Haus}(M, M') := \sup\{\sup_{x\in M} d_{M'}(x), \sup_{x\in M'} d_M(x)\}.$$

We briefly recall the definitions of the different nonsmooth tools that we use in this paper. Let the function $f : \Omega \to \mathbb{R}$ be given, where Ω is a subset of $\mathbb{R}^N$. Let $x \in \Omega$; we say that ζ is a *proximal subgradient* (see Rockafellar [34]) of f at x (written $\zeta \in \partial_P f(x)$) if and only if there exists $\rho, \sigma > 0$ such that for any $y \in x + \rho B \cap \Omega$,

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

We now define the *limiting subdifferential*[2] of f at x by

$$\partial_L f(x) := \limsup_{x'\to x} \partial_P f(x').$$

We define $\partial^P f$ and $\partial^L f$ by

$$\partial^P f(x) := -\partial_P(-f)(x) \text{ and } \partial^L f(x) := -\partial_L(-f)(x).$$

Finally, if the function f is Lipschitz on a neighborhood of x, we define *Clarke's generalized gradient* as follows: for any $x \in \mathbb{R}^N$ we set

$$\partial f(x) := \mathrm{co}\,\partial_L f(x) = \mathrm{co}\,\partial^L f(x)$$

---

[1] Throughout this paper, $\mathbb{R}$ denotes the set of real numbers. If $x = (x_1, ..., x_N)$ and $y = (y_1, ..., y_N)$ belong to $\mathbb{R}^N$, we denote by $\langle x, y \rangle = \sum_{i=1}^N x_i y_i$, the canonical scalar product of $\mathbb{R}^N$, $|x| = \sqrt{\langle x, x \rangle}$, the corresponding Euclidean norm. We denote by $B(x, r) = \{y \in \mathbb{R}^N | |x - y| < r\}$, $\overline{B}(x, r) = \{y \in \mathbb{R}^N | |x - y| \leq r\}$ and $S(x, r) = \{y \in \mathbb{R}^N | |x - y| = r\}$. We denote by $B = B(0, 1)$, $\overline{B} = \overline{B}(0, 1)$, and $S^{N-1} = S(0, 1)$. For $(x, y) \in \mathbb{R}^N \times \mathbb{R}^N$ and $(x', y') \in \mathbb{R}^N \times \mathbb{R}^N$, we denote by $d_\infty((x, y), (x', y')) = \sup\{|x - x'|, |y - y'|\}$ to avoid any confusion with $\|f\|_\infty$, and $\overline{B}_\infty((x, y), r) = \{(x', y') \in \mathbb{R}^N \times \mathbb{R}^N | d_\infty((x, y), (x', y')) \leq 1\}$. If $X \subset \mathbb{R}^N$, $Y \subset \mathbb{R}^N$, and $x \in \mathbb{R}^N$, we let $d_X(x) = \inf_{y\in X} |x - y|$, $X \setminus Y = \{x \in X | x \notin Y\}$ be the set-difference of the sets X and Y, $X + Y = \{x + y | x \in X, y \in Y\}$ the sum of the sets X and Y, $\mathrm{cl}\,X$ the closure of X, $\mathrm{int}\,X$ the interior of X, $\mathrm{bd}\,X = \mathrm{cl}\,X \setminus \mathrm{int}\,X$ the boundary of X, and $\mathrm{co}\,X$ the convex hull of X. Consider a function $f : X \to \mathbb{R}$; its epigraph, denoted by $\mathrm{epi}\,f$, is defined by $\mathrm{epi}\,f := \{(x, y) \in X \times \mathbb{R} | y \geq f(x)\}$. The function f is locally Lipschitz if, for every $x \in X$, there is $\varepsilon > 0$ and $K > 0$ such that $|f(x_1) - f(x_2)| \leq K|x_1 - x_2|$ for every x_1 and x_2 in $B(x, \varepsilon)$. If F is a correspondence (or multivalued map) from X to $\mathbb{R}^N$, its graph, denoted by $\mathrm{graph}(F)$, is defined by $\mathrm{graph}(F) = \{(x, y) \in X \times \mathbb{R}^N | y \in F(x)\}$.

[2] See Mordukhovich [26] for the geometric presentation.

and we recall that

$$\partial f(x) = \mathrm{co}\,\{\zeta | \exists (x_n)_{n \in \mathbb{N}} \subset \mathrm{Dom}\,(\nabla f),\ x_n \to x, \nabla f(x_n) \to \zeta\},$$

which is nonempty by Rademacher's Theorem [32]. Let M be a subset of $\mathbb{R}^N$ and let $x \in R^N$. The *projection* of x on M is the set of points of M nearest to x:

$$\mathrm{proj}\,_M(x) := \{y \in \mathbb{R}^N | \|x - y\| = d_M(x)\}.$$

The *proximal normal cone* to M at x is the set of the *perpendicular vectors* to M at x (see Bourbaki [7]), precisely defined by

$$N_M^P(x) := \{v \in \mathbb{R}^N | \exists \alpha > 0, B(x + \alpha v, \alpha \|v\|) \cap M = \emptyset\}.^3$$

The *limiting normal cone* (see Mordukhovich [26]) $N_M^L(x)$ and the *Clarke normal cone* $N_M^C(x)$ are then defined by

$$
\begin{aligned}
N_M^L(x) &:= \limsup_{x' \to x} N_M^P(x'), \\
N_M^C(x) &:= \mathrm{cl\,co}\, N_M^L(x).
\end{aligned}
$$

The set M is said to be *epi-Lipschitz* at $x \in M$ if it can be written as the epigraph of a Lipschitz function on a neighborhood of x, up to an isometry. Equivalently, from Rockafellar [33], if $N_M^C(x) \cap -N_M^C(x) = \{0\}$. The relations between subdifferentials and generalized gradients of the function f at $x \in \Omega$ and normal cones to the epigraph of f are the following:

$$
\begin{aligned}
\zeta \in \partial_P f(x) &\quad \Leftrightarrow \quad (\zeta, -1) \in N_M^P((x, f(x))), \\
\zeta \in \partial_L f(x) &\quad \Leftrightarrow \quad (\zeta, -1) \in N_M^L((x, f(x))), \\
\zeta \in \partial_C f(x) &\quad \Leftrightarrow \quad (\zeta, -1) \in N_M^C((x, f(x))).
\end{aligned}
$$

Federer [17] introduces sets *of positive reach* as follows.[4] Let $M \subset \mathbb{R}^N$, define reach (M) as the radius of the largest tube around M on which there exists a unique projection on M, precisely

$$\mathrm{reach}\,(M) := \sup\{r | \forall x \in M + B(0, r), \mathrm{proj}\,_M(x) \text{ is a singleton}\}.$$

Then M is of positive reach if reach $(M) > 0$.

2.2. **Statements of the results.** Our first main result specifies Theorem A given in the Introduction and thus provides a negative answer to Question 1(a). We first recall the example of Warga [37] who exhibits a function $f : \mathbb{R}^2 \to \mathbb{R}$ and a sequence $(f_n)_{n \in \mathbb{N}}$ of its mollified functions such that $0 \in \partial f(0)$ and the gradient of the approximating function remains "far" from 0. It is clearly not possible to require this last property to hold for any convolution since one may build a convolution kernel ad hoc in order to obtain, for example, $0 \in \partial f_n(0)$ (see Rockafellar and Wets [35, Theorem 9.67]). However, one may wonder if Question 1(a) would have a positive answer with the Lasry-Lions regularization, in view of its many nice regularizing properties. This is not the case as shown by Theorem 2.1, part (b). The special case where $N = 1$ is examined in Section 2.4, Proposition 2.4.

---

[3]In other words, $N_M^P(x) = \{v \in \mathbb{R}^N | \exists \alpha > 0, x \in \mathrm{proj}\,_M(x + \alpha x)\}$.

[4]Sets of positive reach are an important notion and have been studied with various applications by many authors. They are called proximally smooth sets by Clarke, Stern and Wolenski [10], prox regular sets by Poliquin and Rockafellar [30], Poliquin, Rockafellar and Thibault [31], and sets with property ρ by Plaskacz [28].

Theorem 2.1. *The function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by (see Figure 1)*

$$(2.1) \qquad\qquad f(x_1, x_2) = \left| |x_1| + x_2 \right| + \frac{1}{2}x_1$$

is $\frac{\sqrt{13}}{2}$-Lipschitz and satisfies

$$0 \in \partial f(0).$$

Moreover:

(a) [37, Warga] *(Convolution) There exists a convolution kernel θ such that, for every $\lambda > 0$, the convoluted (or mollified) function $f * \theta_\lambda : \mathbb{R}^2 \to \mathbb{R}$ defined by*

$$f * \theta_\lambda(x) := \int_{\mathbb{R}^2} \theta_\lambda(y) f(x + y) dy, \text{ for every } x \in \mathbb{R}^2, \text{ where } \theta_\lambda(y) = \lambda^{-2}\theta\left(\frac{y}{\lambda}\right),$$

is of class C^∞, $\frac{\sqrt{13}}{2}$-Lipschitz, and moreover

$$(2.2) \qquad\qquad \lim_{\lambda \to 0} \|f - f * \theta_\lambda\|_\infty = 0,$$

$$(2.3) \qquad\qquad |\nabla(f * \theta_\lambda)| \geq \frac{1}{5},$$

hence in particular $0 \notin \limsup_{\lambda \to 0} \operatorname{graph}(\nabla f * \theta_\lambda)$.

(b) *(Lasry-Lions regularization) For every $0 < \mu < \lambda$, the function $f_{\lambda,\mu}$ defined by*

$$f_{\lambda,\mu}(x) = -\inf\left\{ -\inf\{f(z) + \frac{1}{2\lambda}|y - z|^2, z \in \mathbb{R}^2\} + \frac{1}{2\mu}|x - y|^2, y \in \mathbb{R}^2 \right\}$$

is of class C^1, $\frac{\sqrt{13}}{2}$-Lipschitz, and moreover

$$(2.4) \qquad\qquad \lim_{\lambda,\mu \to 0} \|f - f_{\lambda,\mu}\|_\infty = 0,$$

$$(2.5) \qquad\qquad |\nabla f_{\lambda,\mu}| \geq \frac{1}{\sqrt{2}}.$$

(c) *Moreover, for every $0 < \mu < \lambda$,*

$$\nabla f_{\lambda,\mu}(\mathbb{R}^2) = \operatorname{co}\left\{ \left(\frac{1}{2}, 0\right), \left(\frac{3}{2}, 1\right), \left(-\frac{1}{2}, 1\right) \right\}$$

$$\cup \operatorname{co}\left\{ \left(\frac{1}{2}, 0\right), \left(-\frac{1}{2}, -1\right), \left(\frac{3}{2}, -1\right) \right\},$$

and if (f_n) is a sequence of locally Lipschitz functions from $\mathbb{R}^N$ to $\mathbb{R}$ which converges uniformly to f on compact sets, then:

$$\{0\} \times \left(\operatorname{co}\left\{ \left(\frac{1}{2}, 0\right), \left(\frac{3}{2}, 1\right), \left(-\frac{1}{2}, 1\right) \right\} \cup \operatorname{co}\left\{ \left(\frac{1}{2}, 0\right), \left(-\frac{1}{2}, -1\right), \left(\frac{3}{2}, -1\right) \right\} \right)$$
$$\subset \liminf_{n \to +\infty} \operatorname{graph}(\partial f_n).$$

The proof of Theorem 2.1, parts (a), (b) and (c), is given in Section 6.

Remark 2.1. Let us briefly comment on Theorem 2.1. Though it is the "best" (smallest) convex generalized gradient for locally Lipschitz functions, it is well known that Clarke's generalized gradient may be too large. Here the fact that $0 \in \partial f(0)$ for Warga's function does not give much useful information on the function f. The element 0 is not really a critical point of the function f. On the contrary, part (c) shows that, at least for Warga's function, the gradients of the
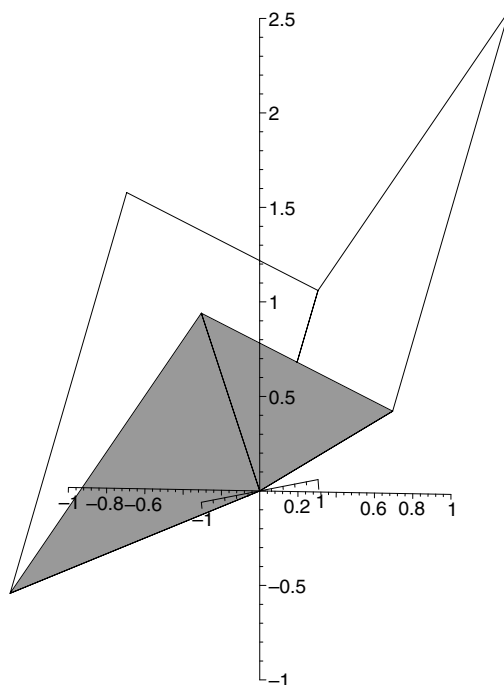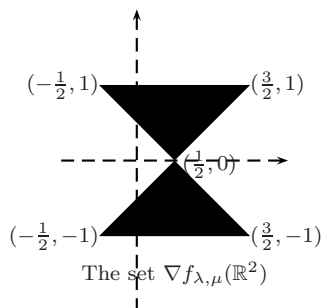
FIGURE 1. $f(x_1, x_2) = \left| |x_1| + x_2 \right| + \frac{1}{2} x_1$

The set $\nabla f_{\lambda,\mu}(\mathbb{R}^2)$

Lasry-Lions regularization capture all the useful "differential" information on the function f.

Remark 2.2. Theorem 2.1(b) shows that the converse inclusion of Benoist [5, Théorème 3] does not hold in general. The same implies that, in general, one cannot remove the convexification in the left-hand side of Georgiev and Zlateva [18, Theorem 5.1].

Remark 2.3. In view of part(c), we raise the following question: for every locally Lipschitzian function f (possibly defined on a Hilbert space), does its Lasry-Lions regularization give the smallest container, namely

$$\operatorname{graph}(\Lambda f) := \liminf_{\|g-f\|_\infty \to 0, g \text{ of class } C^1} \operatorname{graph} \nabla g = \liminf_{\lambda,\mu \to 0} \operatorname{graph}(\nabla f_{\lambda,\mu})?$$

As pointed out above for Warga's function, the answer to this question would be of great importance in order to capture the useful differential information on a locally Lipschitz function, between the (too large) Clarke generalized gradient on one side, and the proximal and limiting gradients on the other side, which obviously miss some good information before their convexification (as a mean value statement).

We now state our second main result, which is a more precise version of Theorem B given in the Introduction and thus gives a positive answer to Question 1(b). The local statement given below is important for applications where global estimates may not be sufficient (see for example, Sections 3.4 and 3.5).

Theorem 2.2. *Let Ω be an open subset of $\mathbb{R}^N$ and let $f : \Omega \to \mathbb{R}$ be a locally Lipschitz function. Then, for every continuous function $\varepsilon : \Omega \to \mathbb{R}_+ \setminus \{0\}$, there exists a function $f_\varepsilon : \Omega \to \mathbb{R}$ of class C^∞ such that for every $x \in \Omega$, we have*

$$|f_\varepsilon(x) - f(x)| \leq \varepsilon(x) \tag{2.6}$$

and

$$\nabla f_\varepsilon(x) \quad \in \quad \partial f(B(x, \varepsilon(x)) \cap \Omega) + B(0, \varepsilon(x)), \tag{2.7}$$
$$\partial f(x) \quad \subset \quad \nabla f_\varepsilon(B(x, \varepsilon(x)) \cap \Omega) + B(0, \varepsilon(x)). \tag{2.8}$$

Theorem 2.2 is proved in Section 4. Moreover, in view of the following easy corollary, Theorem B given in the Introduction is a consequence of Theorem 2.2.

Corollary 2.1. *Under the assumptions of Theorem 2.2, taking ε to be constant, then*

$$\|f_\varepsilon - f\|_\infty \leq \varepsilon,$$
$$d_{Haus}(\text{graph} \nabla f_\varepsilon, \text{graph} \partial f) \leq \varepsilon.$$

Remark 2.4. The sequence $(f_n)_{n \in \mathbb{N}}$ given by Theorem B clearly satisfies $\forall x \in \Omega, \forall \zeta \in \partial f(x), \exists (x_n)_{n \in \mathbb{N}}$ such that

$$\nabla f(x_n) \to_{n \to \infty} \zeta.$$

2.3. Convergence of sets. First note that Theorem 2.1 provides a negative answer to Question 2(a) in the Introduction, in dimension $N \geq 3$, that we now precisely state.

Proposition 2.1. *Let $M = \text{epi} f$ and $M_n = \text{epi} f_n$, where f and $(f_n)_{n \in \mathbb{N}}$ are given by Theorem 2.1. Then $\lim_{n \to +\infty} d_{Haus}(M, M_n) = 0$ and*

$$\left((0,0,0), (0,0,-1)\right) \in \text{graph} N_M^C \setminus \limsup_{n \to +\infty} \text{graph} N_{M_n}^C.$$

Remark 2.5. Thus, in general, one cannot convexify the left-hand side of Benoist [6, Lemma 6.2], one cannot remove the convexification in the right-hand side of Cornet and Czarnecki [12, Lemma 4.1] and one cannot replace the left-hand side of Jourani [20, Theorem 1.2] with Clarke's normal cone.

We now state the consequence of Theorem 2.2 in terms of sets.

Theorem 2.3. *Let M be a closed epi-Lipschitz subset of $\mathbb{R}^N$ and let $\varepsilon : \mathbb{R}^N \to \mathbb{R}_+ \setminus \{0\}$ be a continuous function. Then there exists a closed C^∞ submanifold of $\mathbb{R}^N$ with a boundary of full dimension M' such that, for every $x \in M$ and $x' \in M'$,*

$$(2.9) \qquad\qquad x' \in M + B\left(0, \varepsilon(x')\right),$$

$$(2.10) \qquad\qquad x \in M' + B\left(0, \varepsilon(x)\right),$$

$$(2.11) \qquad N^P_{M'}(x') \cap S^{N-1} \subset N^C_M\left(B(x', \varepsilon(x'))\right) \cap S^{N-1} + B\left(0, \varepsilon(x')\right),$$

$$(2.12) \qquad N^C_M(x) \cap S^{N-1} \subset N^P_{M'}\left(B(x, \varepsilon(x))\right) \cap S^{N-1} + B\left(0, \varepsilon(x)\right).$$

Note that (2.9) to (2.12) can be equivalently reformulated as follows, up to a change of ε:

$$\begin{aligned} N^P_{M'}(x') \cap \overline{B} &\subset N^C_M\left(B(x', \varepsilon(x'))\right) \cap \overline{B} + B\left(0, \varepsilon(x')\right), \\ N^C_M(x) \cap \overline{B} &\subset N^P_{M'}\left(B(x, \varepsilon(x))\right) \cap \overline{B} + B\left(0, \varepsilon(x)\right). \end{aligned}$$

Note that Theorem 2.3 holds true in the more general case where the space $\mathbb{R}^N$ is replaced by an open subset Ω of $\mathbb{R}^N$. Under this generalization, we have the equivalence between the set and functional formulations.

Theorem 2.4. *(Theorem 2.2) $\Leftrightarrow$ (Theorem 2.3).*

The proof of Theorem 2.3 is given in Section 5. In view of the following easy corollary, Theorem C given in the Introduction is a consequence of Theorem 2.3.

Corollary 2.2. *Under the assumptions of Theorem 2.3, taking ε to be constant, then*

$$\begin{aligned} d_{Haus}(M', M) &\leq \varepsilon, \\ d_{Haus}(\mathrm{graph}\left(N^P_{M'} \cap S\right), \mathrm{graph}\left(N^C_M \cap S\right)) &\leq \varepsilon. \end{aligned}$$

Remark 2.6. The sequence $(M_n)_{n \in \mathbb{N}}$ given by Theorem C is a smooth normal approximation of M in the sense of [12]. It thus enjoys additional properties; see [12, Theorem 2.2].

Remark 2.7. Theorem C improves [12, Theorem 2.1], on the smooth normal approximation of epi-Lipschitz sets, by replacing the inclusion $\limsup \mathrm{graph}\left(N^C_{M_n}\right) \subset \mathrm{graph}\left(N^C_M\right)$ by an equality, where M and M_n are defined in Theorem C. In fact, by a careful reading of the proofs of [12] and also [11], one deduces that, in [12, Theorem 2.1], $\limsup \mathrm{graph}\left(N^C_{M_n}\right)$ can be equal to the graph of a smaller cone, namely $\mathrm{graph}\left(\widetilde{N}_M\right)$, where $\widetilde{N}_M$ is defined in Section 3.5. In view of the forthcoming Proposition 3.4, the inclusion $\limsup \mathrm{graph}\left(N^C_{M_n}\right) \subset \mathrm{graph}\left(N^C_M\right)$ in [12, Theorem 2.1] may be strict.

Theorem 2.3 may clearly not hold for any choice of the sets M and M'. However, if the set M is of positive reach, we have the following result, which specifies Proposition B in the Introduction. Since the set M is of positive reach, the different normal cones $N^P_M(\cdot), N^L_M(\cdot)$ and $N^C_M(\cdot)$ coincide.

Proposition 2.2. *Let M be a closed subset of $\mathbb{R}^N$ of positive reach, $r \in \mathbb{R}_+ \setminus \{0\}$ such that $r \leq \operatorname{reach}(M)$, and let M' be a closed subset of $\mathbb{R}^N$. If $\varepsilon := d_{Haus}(M, M') < \frac{r}{2}$, then*

$$(2.13) \qquad \operatorname{graph}(N_M^P \cap S^{N-1}) \subset \operatorname{graph}(N_{M'}^P \cap S^{N-1}) + \frac{\sqrt{\varepsilon^2 + 4\varepsilon r}}{\min\{r/4, 1\}} \overline{B}_\infty$$

in the space $\mathbb{R}^N \times \mathbb{R}^N$.

Proposition 2.2 is proved in Section 5.3. It may clearly not be true if the set M is not of positive reach; take for example, in $\mathbb{R}^2$, $M = \operatorname{epi} f$ with $f(x) = x^2 \sin(1/x)$, $f(0) = 0$.

2.4. **The special cases of dimension 1 for functions and 2 for sets.** Contrary to what happens in dimension $N \geq 2$, we can state the following result for functions on the real line, which provides a positive answer to Question 1(a) in dimension 1.

Proposition 2.3. *Let $f : \mathbb{R} \to \mathbb{R}$ be a locally Lipschitz function and let $(f_n)_{n \in \mathbb{N}} : \mathbb{R} \to \mathbb{R}$ be a sequence of locally Lipschitz functions epi-converging (or Γ-converging) to f. Then*

$$\operatorname{graph}(\partial f) \subset \liminf_{n \to +\infty} \operatorname{graph}(\partial f_n).$$

We now state the set version of Proposition 2.3, which extends Proposition B in the Introduction to the case of Clarke's normal cone in dimension 2.

Proposition 2.4. *Let M be an epi-Lipschitz set in $\mathbb{R}^2$ and let $(M_n)_{n \in \mathbb{N}}$ be a sequence of epi-Lipschitz sets in $\mathbb{R}^2$, such that $M = \limsup_{n \to \infty} M_n = \liminf_{n \to \infty} M_n$. Then*

$$\operatorname{graph}(N_M^C) \subset \liminf_{n \to \infty} \operatorname{graph}(N_{M_n}^C).$$

Proposition 2.4 is proved in Section 7.6. In fact Proposition 2.3 is a consequence of Darboux's theorem, i.e., for every x, y in $\mathbb{R}$, the set $\partial f([x, y])$ is a segment. It can also be seen as a corollary of the following remarkable property.

Lemma 2.1. *Let $f : \mathbb{R} \to \mathbb{R}$ be a locally Lipschitz function. Then for any $(x, y) \in \mathbb{R} \times \mathbb{R}$,*

$$\partial f(x) = \partial_P f(x) \cup \partial^L f(x) = \partial_L f(x) \cup \partial^P f(x).$$

This lemma, which will be proved in Section 7.3, somehow provides a new vision of the Clarke generalized gradient in dimension 1. It can also be stated for epi-Lipschitz sets in dimension 2, a result which will easily imply Proposition 2.4.

Lemma 2.2. *Let M be an epi-Lipschitz closed set in $\mathbb{R}^2$. Then for any $x \in M$, we have*

$$N_M^C(x) = N_M^P(x) \cup N_{\mathbb{R}^2 \setminus \operatorname{int} M}^L(x) = N_M^L(x) \cup N_{\mathbb{R}^2 \setminus \operatorname{int} M}^P(x).$$

We will prove Lemma 2.2 in Section 7.4.

3. Related results, remarks and consequences

This section is devoted to the presentation of related results (Propositions 3.1 and 3.2) which motivated our study and of some direct consequences of our results (Corollaries 3.2 and 3.3, Proposition 3.4). Note that Propositions 3.1 and 3.2 are used to prove the results of Section 2.4.

3.1. **Stability of normal cones.** The next result, which yields Proposition B in the Introduction, is a slight generalization of Lemma 6.2 from Benoist [6] (see also Theorem 3 from Ioffe [19], Theorem 2.3 from Mordukhovich [27] and Theorem P.3 from Kruger and Mordukhovich [21]). Jourani [20] extended the first part of Proposition 3.1 to the Hilbert setting.

Proposition 3.1. *Let M be a nonempty closed subset of $\mathbb{R}^N$, and let $(M_n)_{n \in \mathbb{N}}$ be a sequence of closed subsets of $\mathbb{R}^N$, such that $M = \limsup_{n \to \infty} M_n$. Then the following hold:*

$$(3.1) \qquad \operatorname{graph}(N_M^P) \quad \subset \quad \limsup_{n \to \infty} \operatorname{graph}(N_{M_n}^P),$$

$$(3.2) \qquad \operatorname{graph}(N_M^L) \quad \subset \quad \limsup_{n \to \infty} \operatorname{graph}(N_{M_n}^P) \subset \limsup_{n \to \infty} \operatorname{graph}(N_{M_n}^L).$$

In addition if we assume that $M = \limsup_{n \to \infty} M_n = \liminf_{n \to \infty} M_n$, then

$$(3.3) \qquad \operatorname{graph}(N_M^P) \quad \subset \quad \liminf_{n \to \infty} \operatorname{graph}(N_{M_n}^P),$$

$$(3.4) \qquad \operatorname{graph}(N_M^L) \quad \subset \quad \liminf_{n \to \infty} \operatorname{graph}(N_{M_n}^P) \subset \liminf_{n \to \infty} \operatorname{graph}(N_{M_n}^L).$$

We give a short proof of Proposition 3.1 in Section 7.1. As we said in the Introduction, in general it does not hold for Clarke normal cones. However, since all the different normal cones coincide in the case of sets of positive reach, Proposition 3.1 can be stated for Clarke normal cones if the limit set M is assumed to be (locally) of positive reach. As an immediate consequence, the following result illustrates what can be done with Clarke's normal cone.

Corollary 3.1. *Let $(M_n)_{n \in \mathbb{N}}$ be a sequence of closed subsets of $\mathbb{R}^N$, and let $M = \limsup_{n \to \infty} M_n$. Then for every $x \in M$,*

$$N_M^C(x) \subset \operatorname{cl}\left(\operatorname{co} \limsup_{x' \to x, n \to \infty} N_{M_n}^L(x')\right).$$

If we additionally assume that the set $\operatorname{co} \limsup_{x' \to x, n \to \infty} N_{M_n}^L(x')$ is pointed, then we can suppress "cl" in the above assertion, and in particular the set M is epi-Lipschitz.

We refer to [12] for the proof of Corollary 3.1 and more details.

3.1.1. *Remarks on Proposition* 3.1.

Remark 3.1. The inclusion (3.1) may be strict. Consider $M_n = \mathbb{R} \times \mathbb{R}_- \cup \mathbb{R}_- \times (0, 1/n]$ $(n \geq 1)$, $M = \mathbb{R} \times \mathbb{R}_-$, $N_M^P(0) = \{0\} \times \mathbb{R}_+$, $\limsup_{x' \to 0, n \to \infty} N_{M_n}^P(x') = \mathbb{R}_+^2$.

Remark 3.2. The inclusion (3.1) may no longer be true if we replace the assumption $M = \limsup M_n$ by $M = \liminf M_n$. Consider $M_{2n} = \mathbb{R}_- \times \mathbb{R}$, $M_{2n+1} = \mathbb{R} \times \mathbb{R}_-$, $M = \mathbb{R}_-^2$, $N_M^P(0) = \mathbb{R}_+^2$, $\limsup_{x' \to 0, n \to \infty} N_{M_n}^P(x') = \mathbb{R}_+ \times \{0\} \cup \{0\} \times \mathbb{R}_+$.

Remark 3.3. The inclusion (3.3) may no longer be true if we only assume that $M = \liminf M_n$, resp. $M = \limsup M_n$ (without assuming $M = \limsup M_n$, resp. $M = \liminf M_n$). Consider $M = \mathbb{R}_-^2$, resp. $M = \mathbb{R}^2 \setminus \operatorname{int} \mathbb{R}_+^2$, and the sequence (M_n) defined in Remark 3.2.

3.1.2. *Remark on Corollary* 3.1.

Remark 3.4. In Corollary 3.1, one cannot exchange the convexification and the limsup operations, i.e., one cannot replace the set co $\limsup_{x' \to x, n \to \infty} N_{M_n}^L(x')$ by the set $\limsup_{x' \to x, n \to \infty} N_{M_n}^C(x')$. When both M and M_n are epi-Lipschitz, see Proposition 2.1. If either M or M_n is not epi-Lipschitz, then see the elementary counterexamples below to the inclusion $N_M^C(x) \subset \limsup_{x' \to x, n \to \infty} N_{M_n}^C(x')$.

Example 1. If M is not epi-Lipschitz, take $M_n = \Big([1/n, +\infty) \times [1/n, +\infty)\Big) \cup \Big((-\infty, -1/n] \times (-\infty, -1/n]\Big)$ $(n \geq 1)$, $M = \mathbb{R}_+^2 \cup \mathbb{R}_-^2$. Then $N_M(0) = \mathbb{R}^2$ and $\limsup_{x \to 0, n \to \infty} N_{M_n}(x) = \mathbb{R}_+^2 \cup \mathbb{R}_-^2$.

Example 2. Consider $M = \{(x, y) \in \mathbb{R}^2 | y \leq |x|\}$, $M_n = \{(x, y) \in \mathbb{R}^2 | y \leq |x| + \sqrt{|x|}/k\}$. Then $N_M(0) = \{(x, y) \in \mathbb{R}^2 | y \geq |x|\}$, $\limsup_{x \to 0, n \to \infty} N_{M_n}(x) = \mathbb{R} = \mathbb{R} \times \{0\}$. In this case M is epi-Lipschitz, $\delta(M \cap \overline{B}(0, p), M_n \cap \overline{B}(0, p))$ converges to 0 for every p, but M_n is not epi-Lipschitz, and $\limsup_{x \to 0, n \to \infty} N_{M_n}(x)$ is not pointed.

3.2. **Stability of subdifferentials.**
In terms of functions, Proposition 3.1 can be interpreted as follows. $\partial_D f$ denotes the Dini-subgradient of f (see for example [9]). Note that (3.7) corresponds to Lebourg [24].

Proposition 3.2. *Let* $f : \mathbb{R}^N \to \mathbb{R}$ *be a lower semicontinuous function and let* $(f_n)_{n \in \mathbb{N}}$ *be a sequence of lower semicontinuous functions from* $\mathbb{R}^N$ *into* $\mathbb{R}$ *which epi-converges (or* Γ*-converges) to* f. *Then the following properties hold:*

$$(3.5) \qquad \text{graph}(\partial_P f) \quad \subset \quad \liminf_{n \to +\infty} \text{graph}(\partial_P f_n),$$

$$(3.6) \qquad \text{graph}(\partial_L f) \quad \subset \quad \liminf_{n \to +\infty} \text{graph}(\partial_P f_n) \subset \liminf_{n \to +\infty} \text{graph}(\partial_L f_n),$$

$$(3.7) \qquad \text{graph}(\partial_D f) \quad \subset \quad \liminf_{n \to +\infty} \text{graph}(\partial_P f_n) \subset \liminf_{n \to +\infty} \text{graph}(\partial_D f_n).$$

In fact, one easily sees that assertions (3.5), (3.6) and (3.7) are equivalent. The short proof of Proposition 3.2 is given in Section 7.2. In view of Theorem 2.1, the corresponding property does not hold for Clarke generalized gradients in the case of locally Lipschitz functions.

3.3. **The topological viewpoint.**
The regularization part of Proposition A in the Introduction can be viewed as a density result: endow the space $C^1(\Omega, \mathbb{R})$ with the C^1 pseudo-norm ($\|f\|_{C^1} := \|f\|_\infty + \|\nabla f\|_\infty$) Then

$$C^1(\Omega, \mathbb{R}) = \overline{C^\infty(\Omega, \mathbb{R})}^{\|\cdot\|_{C^1}}.$$

As we explained in the Introduction, we want to extend the result to the Lipschitzian setting. The C^1 pseudo-norm can be naturally extended into the *Lip* pseudo-norm to the space $Lip_{loc}(\Omega, \mathbb{R})$ ($\|f\|_{Lip} := \|f\|_\infty + Lip(f)$, where $Lip(f)$ is the Lipschitz constant of f. In view of Rademacher's theorem, which asserts that f is differentiable almost everywhere in $\mathbb{R}^N$, we have $\|f\|_{Lip} = \|f\|_\infty + \|\nabla f\|_\infty$). But one easily sees that the subspace $C^1(\Omega, \mathbb{R})$ is closed for the corresponding topology $\mathcal{T}^{\|\cdot\|_{Lip}}$. For our purpose of approximating Lipschitz functions by C^∞ functions, the topology $\mathcal{T}^{\|\cdot\|_{Lip}}$ is too fine: for example, $x \mapsto |x|$ and $x \mapsto |x - v|$ are "far",

even for small v, since $\|\quad|.|-|.-v|\quad\|_{Lip} \geq 2$. The next step is to introduce the pseudo-distance d_{Lip}, defined by

$$d_{Lip}(f,g) := \|f-g\|_\infty + d_{Haus}(\text{graph}\,(\partial g), \text{graph}\,(\partial g)),$$

and we obtain

$$Lip_{loc}(\Omega, \mathbb{R}) = \overline{C^\infty(\Omega, \mathbb{R})}^{d_{Lip}}.$$

Note that

$$d_{Lip}(f,g) \leq \|f-g\|_{Lip},$$

hence $\mathcal{T}^{d_{Lip}} \subset \mathcal{T}^{\|\cdot\|_{Lip}}$. Also, d_{Lip} cannot be induced by a norm (it could only be the pseudo-norm $\|\cdot\|_{Lip}...$), but it yields a structure of topological vector space (either directly on Lipschitz functions, with a prescribed Lipschitz constant, or defining on $Lip_{loc}(\Omega, \mathbb{R})$ the distance $d(f,g) := \sum_{n\in\mathbb{N}} \frac{1}{2^n} \min\{1, d_{Lip}(f|_{K_n}, g|_{K_n})\}$, where (K_n) is a sequence of compact subsets of Ω such that $\Omega = \bigcup_{n\in\mathbb{N}} K_n$).

3.4. **Quasi-smoothing of Lipschitz functions.** The following result improves the smoothing results of [15] and also [14, Theorem 4.1], [16, Chapitre 1, Lemma 4.2], [11, Lemma 4.2]. In particular some properties (see below) in the above results were obtained in the special case where $f = d_M$, the distance function to a nonempty (and different from the whole space $\mathbb{R}^N$) set M, and also $f = d_M - d_{\mathbb{R}^N \setminus M}$, the signed distance function to M. The proofs of these properties resulted from a geometrical observation on the distance function ([16, Chapitre 1, Proposition 4.4]). In the following results, these properties, stated in Proposition 3.3, are extended to any locally Lipschitz function $f : \Omega \to \mathbb{R}$. We recall that the smoothing results mentioned above are the key tools for the approximation results developed in [12] and [14] and also for the study of the existence of (generalized) equilibria developed in [13] and [14]. The precision brought by Corollary 3.2 is exactly assertion (3.11). Assertions (3.8), (3.9) and (3.10) are shown in [15] et al. with a convolution-type argument.

Corollary 3.2. *Let U be an open subset of $\mathbb{R}^N$, let $f : U \to \mathbb{R}$ be a locally Lipschitz function, and let $\delta : U \to \mathbb{R}_+$ be a lower semicontinuous function. Then there is a function $f_\delta : U \to \mathbb{R}$ such that:*

(3.8)   *f_δ is locally Lipschitz on U, and C^∞ on $\{x \in U | \delta(x) > 0\}$,*

(3.9)   *$|f_\delta(x) - f(x)| \leq \delta(x), \forall x \in U$,*

(3.10)   *$\partial f_\delta(x) \subset \partial f\left(\overline{B}(x, \delta(x)) \cap U\right) + \overline{B}(0, \delta(x)), \forall x \in U$,*

(3.11)   *$\partial f(x) \subset \partial f_\delta\left(\overline{B}(x, \delta(x)) \cap U\right) + \overline{B}(0, \delta(x)), \forall x \in U$.*

The proof of Corollary 3.2 is given in Section 8.1. The following property can immediately be deduced from Corollary 3.2 (from assertions (3.10) and (3.11)).

Proposition 3.3. *Under the assumptions of Corollary 3.2, let $x \in U$ such that $\delta(x) = 0$. Then*

$$\partial f_\delta(x) = \partial f(x) \ \text{if} \ \delta(x) = 0.$$

Moreover, if additionally $\delta \leq d_{\delta^{-1}(\{0\})}$ on a neighborhood of x, then

$$\limsup_{x'\to x, \delta(x')>0} \partial f_\delta(x') = \limsup_{x'\to x, \delta(x')>0} \partial f(x').$$

The first part was proved in the special cases $f = \Delta_M$, respectively $f = d_M$, in [16, Chapitre 1, Lemma 4.2], respectively [15], by using [16, Chapitre 1, Proposition 4.4]. The second part, shown for the distance function d_M in [15] (again by using [16, Chapitre 1, Proposition 4.4]) is a key tool for [14, Theorem 2.1 and Theorem 2.2].

3.5. **Smooth normal approximation of proximally nondegenerate sets.** Before stating the result, we recall the essential definitions introduced in [14]. Let M be a closed subset of $\mathbb{R}^N$. For $x \in M$, the cone $\widetilde{N}_M(x)$ is defined by

$$\widetilde{N}_M(x) = \bigcup_{\lambda \geq 0} \lambda \limsup_{x' \to x, d_M(x') > 0} \partial d_M(x').$$

The cone $\widetilde{N}_M(x)$ contains the limiting normal cone $N_M^L(x)$ and is included in Clarke's normal cone $N_M^C(x)$. It may be nonconvex. The set M is said to be proximally nondegenerate if

$$(3.12) \qquad 0 \notin \limsup_{x' \to x, d_M(x') > 0} \partial d_M(x') \text{ for every } x \in M.$$

Closed epi-Lipschitz sets and sets of positive reach are proximally nondegenerate. The following result improves [14, Theorem 2.1], replacing an inclusion by an equality in assertion (cn).

Corollary 3.3. *Let M be a compact proximally nondegenerate subset of $\mathbb{R}^N$. Then the set M admits a smooth normal approximation (M_n) in the sense that:*

  (i) *for every n, M_n is a compact and smooth subset of $\mathbb{R}^N$, i.e., is a closed C^∞ submanifold with a boundary of $\mathbb{R}^N$ of full dimension;*
  (ii) *for every n, $M_{n+1} \subset M_n \subset B(M, 1)$, and $M = \bigcap_{n \in \mathbb{N}} M_n$;*
  (ret) *for every n, M is a deformation retract of M_n;*
  (cn) *$\mathrm{graph}(\widetilde{N}_M) = \limsup_{n \to \infty} \mathrm{graph}(N_{M_n})$.*

The proof of Corollary 3.3 is given in Section 8.2.

3.6. **Strict inclusion $\widetilde{N}_M \subset N_M^C$.** In [14], there is a simple example showing that the cone $\widetilde{N}_M$ may be strictly smaller than Clarke's normal cone N_M^C if the set M is assumed to be proximally nondegenerate. It remained an open question in the epi-Lipschitz case. The counterexample given in Theorem 2.1 shows that the cone $\widetilde{N}_M$ may be strictly smaller than Clarke's normal cone N_M^C even if the set M is epi-Lipschitz.

Proposition 3.4. *Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by $f(x_1, x_2) = \left| |x_1| + x_2 \right| + \frac{1}{2}x_1$. Then*

$$N_{\mathrm{epi}\,f}^C(0) = \bigcup_{\lambda \geq 0} \lambda \partial f(0) \times \{-1\}$$

$$= \bigcup_{\lambda \geq 0} \lambda \left(\mathrm{co}\left\{ \left(\frac{3}{2}, 1\right), \left(-\frac{1}{2}, 1\right), \left(-\frac{1}{2}, -1\right), \left(\frac{3}{2}, -1\right) \right\} \times \{-1\} \right),$$

$$\widetilde{N}_{\mathrm{epi}\,f}(0) = \bigcup_{\lambda \geq 0} \lambda \left(\mathrm{co}\left\{ \left(\frac{1}{2}, 0\right), \left(\frac{3}{2}, 1\right), \left(-\frac{1}{2}, 1\right) \right\} \right.$$

$$\left. \cup \mathrm{co}\left\{ \left(\frac{1}{2}, 0\right), \left(-\frac{1}{2}, -1\right), \left(\frac{3}{2}, -1\right) \right\} \times \{-1\} \right),$$

which is strictly smaller than Clarke's cone $N_{\mathrm{epi}\,f}^C(0)$.

The computation of Clarke's normal cone $N_{\mathrm{epi}\,f}^{C}(0)$ amounts to the knowledge of $\partial f(0)$ and is straightforward. The remainder of the proof of Proposition 3.4 is given in Section 8.3. As a consequence of Proposition 3.4, [14, Theorem 2.2] on the existence of equilibria is strictly stronger than the previous results in the literature, even in the epi-Lipschitz case. Also, it implies that Corollary 3.3 and Theorem 2.3 are not comparable.

4. Regularization of functions: Proof of Theorem 2.2

The main idea of the proof is the following. With traditional approximation schemes, such as the convolution or the Lasry-Lions regularization, we may lose some elements of the subdifferential ∂f as shown by Theorem 2.1. On the other hand, two categories of functions do not have this drawback, namely the C^1 functions and the convex functions. If we were able to approach f by C^1 functions which keep all the values of the subdifferential ∂f, the proof of Theorem 2.2 would be finished. In view of Attouch's Theorem ([1] and [2] Theorem 3.66), if f is convex and (f_n) is a sequence of functions which Mosco converges to f (which here amounts to the epi-convergence of (f_n) to f), then $\operatorname{graph}\partial f = \liminf \operatorname{graph}\partial f_n = \limsup \operatorname{graph}\partial f_n$. So the main idea of the proof of Theorem 2.2 is to locally replace the function f by a convex[5] function close to f, whose subdifferential has almost the same values than ∂f. This is achieved by using the support function σ_K for K close to ∂f in a neighborhood of a point $x \in \Omega$.

The proof of Theorem 2.2 is organized as follows. In Section 4.1, we first prove an easier version of Theorem 2.2 involving the proximal subgradient, which in particular gives Theorem 2.2 in the lower and upper C^2 cases. In Section 4.2, in the general case, we apply the idea that we just described to obtain a smooth approximation of f satisfying the conclusions of Theorem 2.2 in a neighborhood of a given element $x \in \Omega$. In Section 4.3, we put together the local approximations in order to obtain a global upper C^2 approximation. We finally conclude in Section 4.4.

4.1. Involving the P-subgradient (regularization in the lower and upper C^2 cases).

Proposition 4.1. *Let Ω be an open subset of $\mathbb{R}^N$. Let $f : \Omega \to \mathbb{R}$ be a locally Lipschitz function. For every continuous function $\varepsilon : \Omega \to \mathbb{R}_+ \setminus \{0\}$, there exists a function $\tilde{f} : \Omega \to \mathbb{R}$ of class C^∞ such that for every $x \in \Omega$, we have*

$$(4.1) \qquad |\tilde{f}(x) - f(x)| \leq \varepsilon(x),$$

$$(4.2) \qquad \nabla \tilde{f}(x) \subset \partial f(x + \varepsilon(x)B \cap \Omega) + \varepsilon(x)B,$$

$$(4.3) \qquad \partial_P f(x) \cup \partial^P f(x) \subset \nabla \tilde{f}(x + \varepsilon(x)B \cap \Omega) + \varepsilon(x)B.$$

Moreover, if the function f is convex, then the function $\tilde{f}$ can be taken to be convex as well on a given compact subset of Ω and

$$(4.4) \qquad \nabla \tilde{f}(x) \subset \operatorname{co}\partial f(x + \alpha B).$$

Remark 4.1. Poliquin's extension of Attouch's theorem [29, Theorem 2.1] is not of immediate help for the proof of Proposition 4.1, since [29] requires an equi-semiconvexity (equi-primal lower nice) assumption on the approximation. This could clearly not be satisfied in our case.

---

[5]Concave functions would of course do as well. The point is to eliminate the effect of the saddle-type behavior such as the function defined in Theorem 2.1.

4.1.1. *Preliminary results.* Before beginning the proof of Proposition 4.1, we need some preliminary results.

Lemma 4.1. *Let M and M' be two nonempty closed subsets of $\mathbb{R}^N$. Let $x \in M$ and $v \in \mathbb{R}^N$ such that $x \in \operatorname{proj}_M(x+v)$. Then, for every $x' \in \operatorname{proj}_{M'}\left(x+\frac{v}{2}\right)$,*

$$(4.5) \qquad |x-x'|^2 \leq d_{Haus}(M,M')^2 + 4d_{Haus}(M,M')|v|.$$

The majoration in (4.5) is sharp. For example, in $\mathbb{R}^2$ take $M = \{(0,0),(0,1)\}$, $M' = \{(0,-1/4),(0,3/4)\}$, $x = (0,0)$, $v = (0,1/2)$, $x' = (0,3/4)$.

Proof of Lemma 4.1. Consider $x' \in \operatorname{proj}_{M'}(x+\frac{v}{2})$. Set $\delta := d_{Haus}(M,M')$. There exists $s' \in M'$ such that $|x-s'| \leq \delta$. Hence we deduce

$$|x+\frac{v}{2}-x'| = d_{M'}\left(x+\frac{v}{2}\right) \quad \leq \quad |x+\frac{v}{2}-s'|$$
$$\leq \quad |x+\frac{v}{2}-x| + |x-s'|$$
$$(4.6) \qquad\qquad\qquad\qquad\qquad \leq \quad \frac{1}{2}|v| + \delta.$$

But we can compute

$$d_{M'}\left(x+\frac{v}{2}\right)^2 = |x-x'|^2 + \frac{1}{4}|v|^2 + \langle v, x-x'\rangle.$$

This equality combined with (4.6) gives

$$(4.7) \qquad |x-x'|^2 + \langle v, x-x'\rangle \leq \delta^2 + \delta|v|.$$

On the other hand, there exists $s \in M$ such that $|x'-s| \leq \delta$. Thus the fact that $x \in \operatorname{proj}_M(x+v)$ implies that

$$|v| = |x+v-x| = d_M(x+v) \leq |x+v-s| \quad \leq \quad |x+v-x'| + |x'-s|$$
$$(4.8) \qquad\qquad\qquad\qquad\qquad\qquad\qquad \leq \quad |x+v-x'| + \delta.$$

Assume now $|v| \geq \delta$. From (4.8), we deduce

$$(|v|-\delta)^2 \leq |x-x'|^2 + |v|^2 + 2\langle v, x-x'\rangle,$$

which implies

$$\delta^2 - 2\delta|v| \leq |x-x'|^2 + 2\langle v, x-x'\rangle.$$

Combined with (4.7), we deduce

$$|x-x'|^2 \leq \delta^2 + 4\delta|v|.$$

Consider now the case $|v| < \delta$. Since $|x-x'| \leq |x+\frac{v}{2}-x'| + \frac{1}{2}|v|$, then (4.6) implies

$$|x-x'|^2 \leq \delta^2 + 2\delta|v| + |v|^2.$$

Since $|v| < \delta$, then $|v|^2 \leq \delta|v|$, and from above $|x-x'|^2 \leq \delta^2 + 3\delta|v|$, which proves (4.5) as well. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Lemma 4.2. *For every continuous function $\varepsilon : \Omega \to \mathbb{R}_+ \setminus \{0\}$, there exists a continuous map $\alpha : \Omega \to \mathbb{R}_+ \setminus \{0\}$ such that, for every $x \in \Omega$, $B(x,\alpha(x)) \subset \Omega$ and*

$$(4.9) \qquad \operatorname{co}\partial f(x+\alpha(x)B) \subset \partial f(B(x,\varepsilon(x)) \cap \Omega) + B(0,\varepsilon(x)).$$

Proof of Lemma 4.2. Since the multivalued map $x \mapsto \partial f(x)$ is upper semicontinuous on Ω, for every $x \in \Omega$ there exists $\alpha_x \in (0, \varepsilon(x))$ such that $B(x, \alpha_x) \subset \Omega$ and

$$\forall y \in B(x, \alpha_x), \partial f(y) \subset \partial f(x) + \frac{\varepsilon(x)}{2} B.$$

Since the right-hand side of the inclusion above is convex, we can deduce that

(4.10) $$\mathrm{co}\, \partial f(y + \alpha_x B) \subset \partial f(x) + \frac{\varepsilon(x)}{2} B.$$

Furthermore by continuity of $\varepsilon(\cdot)$, for every $x \in \Omega$ there exists a constant $\beta_x \in (0, \alpha_x)$ such that

(4.11) $$y \in B(x, \beta_x) \implies \varepsilon(y) \leq 2\varepsilon(x).$$

The family $(B(x, \frac{\beta_x}{2}))_{x \in \Omega}$ is an open covering of Ω, thus it admits a locally finite refinement $(B(y, \frac{\beta_y}{2}))_{y \in Y}$, where Y is a countable subset of Ω, and associate with it a subordinated C^∞ partition of unity $(\lambda_y)_{y \in Y}$. Define the function $\alpha : \Omega \mapsto (0, \infty)$ by

$$\alpha(x) := \frac{1}{2} \sum_{y \in Y} \lambda_y(x) \beta_y, \forall x \in \Omega.$$

Let us now prove that (4.9) is satisfied. Fix $x \in \Omega$. By definition of $\alpha(x)$ there exists $y \in Y$ such that

(4.12) $$\alpha(x) \leq \frac{\beta_y}{2} \text{ and } x \in B\left(y, \frac{\beta_y}{2}\right).$$

Thus we deduce that

$$\partial f(x + \alpha(x)B) \subset \partial f\left(y + \left(\frac{\beta_y}{2} + \alpha(x)\right) B\right) \subset \partial f(y + \beta_y B).$$

Hence by (4.10) and since $\beta_y < \alpha_y$, this implies

$$
\begin{aligned}
\mathrm{co}\,[\partial f(x + \alpha(x)B)] \quad &\subset \quad \partial f(y) + \frac{\varepsilon(y)}{2} B \\
&\subset \quad \partial f(x + \frac{\beta_y}{2}B) + \frac{\varepsilon(y)}{2} B \quad \text{(by (4.12))} \\
&\subset \quad \partial f(x + \varepsilon(x)B) + \varepsilon(x)B \quad \text{(by (4.11))}.
\end{aligned}
$$

The proof is complete. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

4.1.2. *Proof of Proposition* 4.1. For every $x \in \Omega$, we set

$$\delta(x) := \min\left\{\frac{1}{2}d(\mathbb{R}^N \setminus \Omega, x), 1\right\}.$$

Since the set Ω is open, then $\delta(x) > 0$ and clearly $\overline{B}(x, \delta(x)) \subset \Omega$. Furthermore, for every $x \in \Omega$ the function f is globally Lipschitz on the compact ball $\overline{B}(x, \delta(x))$. Let us denote by L_x^f the Lipschitz constant of f on this ball, i.e., there is no constant $L < L_x^f$ such that f is L-Lipschitz on $\overline{B}(x, \delta(x))$. The function $x \mapsto L_x^f$ may not be continuous, but by construction it is locally bounded. Thus there exists a continuous function $Lip(f, \cdot) : \Omega \to [1, +\infty)$ such that for every $x \in \Omega, L_x^f \leq Lip(f, x)$ and hence such that

for every x', x'' in $\overline{B}(x, \delta(x)), |f(x') - f(x'')| \leq Lip(f, x)|x' - x''|$.

In order to obtain a C^∞ approximation, we shall use a classical convolution argument rather than a Lasry-Lions regularization. By Lemma 4.2, there exists a function $\alpha : \Omega \to \mathbb{R}_+ \setminus \{0\}$ such that, for every $x \in \Omega$,

$$(4.13) \qquad \mathrm{co}\, \partial f(x + \alpha(x)B) \subset \partial f(B(x, \varepsilon(x)) \cap \Omega) + B\left(0, \frac{\varepsilon(x)}{2}\right).$$

For every $x \in \Omega$, we set

$$(4.14) \qquad \delta_1(x) := \frac{1}{9 Lip(f, x)} \min\left\{\delta(x)^2, \frac{\varepsilon(x)^2}{Lip(f, x)^4}\right\}.$$

As an immediate consequence of [11, Lemma 4.3], there exists a function $\rho : \Omega \to \mathbb{R}_+ \setminus \{0\}$ of class C^∞ such that for every $x \in \Omega$,

$$(4.15) \qquad \rho(x) \le \min\left\{\delta(x), \alpha(x), \frac{\varepsilon(x)}{Lip(f, x)}, \frac{\min\{\delta_1(x'), x' \in \overline{B}(x, \delta(x))\}}{Lip(f, x)}\right\},$$

$$(4.16) \quad |\nabla \rho(x)| \le \frac{\varepsilon(x)}{2 Lip(f, x)}.$$

Note that if we only want an approximation of f on a compact subset of Ω, we can take ρ to be constant. We define the function $\tilde{f} : \mathbb{R}^N \to \mathbb{R}$ by

$$\tilde{f}(x) := \int_{\overline{B}} \theta(t) f(x - \rho(x)t)dt, \forall x \in \mathbb{R}^N,$$

where θ is a mollifier.[6] The function $\tilde{f}$ is of class C^∞, and it satisfies

$$\tilde{f}(x) - f(x) = \int_{\overline{B}} \theta(t) f(x - \rho(x)t) - f(x)dt.$$

Since $\rho(x) \le \delta(x)$, then $x - \rho(x)t \in \overline{B}(x, \delta(x))$ for every $t \in \overline{B}$. Since the function f is L_x^f-Lipschitz on $\overline{B}(x, \delta(x))$, $|f(x - \rho(x)t) - f(x)| \le L_x^f \rho(x)|t|$. Hence

$$(4.17) \qquad |\tilde{f}(x) - f(x)| \le L_x^f \rho(x) \le \varepsilon(x)$$

and (4.1) is satisfied. From Rademacher's Theorem [32], since f is locally Lipschitz, it is differentiable almost everywhere in $\mathbb{R}^N$, and for every point x where it is differentiable, we have

$$(4.18) \qquad \nabla f(x) \subset \partial f(x).$$

On the other hand, since the function f is locally Lipschitz and the function ρ is of class C^1, from Lebesgue's Theorem we deduce that for every $x \in \mathbb{R}^N$,

$$\begin{aligned}(4.19) \qquad \nabla \tilde{f}(x) &= \int_{\overline{B}} \theta(t) \nabla_x f(x - \rho(x)t)dt \\ &= \int_{\overline{B}} \theta(t)\Big(\nabla f(x - \rho(x)t) - \langle \nabla f(x - \rho(x)t), t\rangle \nabla \rho(x)\Big)dt.\end{aligned}$$

For every $t \in \overline{B}$, since the function f is L_x^f-Lipschitz on $\overline{B}(x, \delta(x))$ and $x - \rho(x)t \in \overline{B}(x, \delta(x))$, $|\nabla f(x - \rho(x)t)| \le L_x^f$ for a.e. t, and we deduce from (4.18) and (4.19)

$$(4.20) \qquad \nabla \tilde{f}(x) \subset \mathrm{co}\left(\partial f(x + \rho(x)B)\right) + \overline{B}(0, L_x^f |\nabla \rho(x)|).$$

---

[6]I.e., $\theta : \mathbb{R}^N \to \mathbb{R}$ is a C^∞ function satisfying the following three conditions: (i) for all $x \in \mathbb{R}^N$, $\theta(x) \ge 0$, (ii) $\mathrm{supp}\, \theta \subset B$, and (iii) $\int_{\mathbb{R}^N} \theta(t)dt = 1$.

Since $\rho(x) \leq \alpha(x)$ and $|\nabla \rho(x)| \leq \frac{\varepsilon(x)}{2Lip(f,x)}$, in view of (4.13), we obtain

$$\nabla \tilde{f}(x) \subset \partial f(B(x, \varepsilon(x)) \cap \Omega) + B(0, \varepsilon(x)),$$

and the assertion (4.2) is satisfied. Let us now prove the property (4.3). Consider $x \in \Omega$ and $\zeta \in \partial_P f(x)$. Then

$$\zeta \in \partial_P f(x) \Longleftrightarrow (\zeta, -1) \in N_{\mathrm{epi}(f)}^P(x, f(x)).$$

In order to apply Lemma 4.1 to the sets

$$
\begin{aligned}
M &:= \operatorname{epi} f \cap \overline{B}\left(x, \frac{\delta(x)}{2}\right) \times \mathbb{R}, \\
M' &:= \operatorname{epi} \tilde{f} \cap \overline{B}\left(x, \frac{\delta(x)}{2}\right) \times \mathbb{R},
\end{aligned}
$$

we first claim

$$d_{Haus}(M, M') \leq \delta_1(x).$$

Indeed, if $(x_1, y_1) \in M$, then in view of (4.17), $y_1 \geq f(x_1) \geq \tilde{f}(x_1) - L_{x_1}^f \rho(x_1)$. Since $(x_1, y_1 + L_{x_1}^f \rho(x_1)) \in M'$, we have $d_{M'}(x_1, y_1) \leq L_{x_1}^f \rho(x_1)$. If $(x_1, y_1) \in M'$, we also get $d_M(x_1, y_1) \leq L_{x_1}^f \rho(x_1)$. We obtain the bound on $d_{Haus}(M, M')$ by showing

$$(4.21) \qquad \forall x_1 \in \overline{B}\left(x, \frac{\delta(x)}{2}\right), L_{x_1}^f \rho(x_1) \leq \delta_1(x).$$

Indeed, since $x_1 \in \overline{B}(x, \frac{\delta(x)}{2})$, we have $B(x_1, \delta(x)) \subset B(x, 2\delta(x)) \subset \Omega$, and hence $d(\mathbb{R}^N \setminus \Omega, x_1) \geq \delta(x)$ and $\delta(x_1) \geq \frac{\delta(x)}{2}$. In particular, $x \in \overline{B}(x_1, \delta(x_1))$. In view of (4.15), we deduce $L_{x_1}^f \rho(x_1) \leq \delta_1(x)$. Note also that

$$(x, f(x)) \in \operatorname{proj}_M\left((x, f(x)) + (\zeta, -1)\right).$$

Take $(x', y') \in \operatorname{proj}_{M'}\left((x, f(x)) + \frac{1}{2}(\zeta, -1)\right)$. From Lemma 4.1

$$|(x, f(x)) - (x', y')| \leq (\delta_1(x)^2 + 4\delta_1(x)\sqrt{|\zeta|^2 + 1})^{1/2}.$$

Since $\zeta \in \partial f(x)$, $|\zeta| \leq L_x^f \leq Lip(f, x)$. Recalling that $1 \leq Lip(f, x)$ and noting that $\delta_1(x) \leq 1$, we have

$$
\begin{aligned}
\left(\delta_1(x)^2 + 4\delta_1(x)\sqrt{|\zeta|^2 + 1}\right)^{1/2} &\leq \left((1 + 4\sqrt{2})Lip(f, x)\delta_1(x)\right)^{1/2} \\
&\leq \min\left\{\delta(x), \frac{\varepsilon(x)}{Lip(f, x)^2}\right\},
\end{aligned}
$$

hence

$$|(x, f(x)) - (x', y')| \leq \min\left\{\delta(x), \frac{\varepsilon(x)}{Lip(f, x)^2}\right\}.$$

Since $|x' - x| < \delta(x)$, then $(x', y') \in \operatorname{epi} \tilde{f}$. On the other hand, $(x, f(x)) + \frac{1}{2}(\zeta, -1) - (x', y') \in N_M^P(x', y')$, hence $(x', y') \in \operatorname{bd} \operatorname{epi} \tilde{f}$, i.e., $y' = f(x')$. We thus deduce

$$(\zeta', -1) := \left(\frac{x - x' + \frac{1}{2}\zeta}{\frac{1}{2} - f(x) + \tilde{f}(x')}, -1\right) \in N_{\mathrm{epi}\, \tilde{f}}^P(x', f(x')),$$

which implies $\zeta' = \nabla \tilde{f}(x')$ since $\tilde{f}$ is of class C^1. Since $|x' - x| < \frac{\varepsilon(x)}{Lip(f,x)^2}$ and in view of (4.17), we have $|f(x) - \tilde{f}(x')| \leq |f(x) - f(x')| + |f(x') - \tilde{f}(x')| \leq L_x^f \frac{\varepsilon(x)}{Lip(f,x)^2} + L_{x'}^f \rho(x')$. In view of (4.21), $L_{x'}^f \rho(x') \leq \delta_1(x) \leq \frac{\varepsilon(x)}{Lip(f,x)}$, assuming

without any loss of generality that $\varepsilon \leq 1$. Hence $|f(x) - \tilde{f}(x')| \leq 2\frac{\varepsilon(x)}{Lip(f,x)} \leq \frac{1}{4}$ (assuming without any loss of generality $Lip(f,x) \geq 8\varepsilon(x)$).

$$
\begin{aligned}
|\zeta' - \zeta| &= \frac{|x - x' - \zeta(-f(x) + \tilde{f}(x'))|}{|\frac{1}{2} - f(x) + \tilde{f}(x')|} \\
&\leq \frac{1}{4}(|x - x'| + |\zeta||\tilde{f}(x') - f(x)|) \\
&\leq \frac{1}{4}\left(\frac{\varepsilon(x)}{Lip(f,x)^2} + 2L_x^f\frac{\varepsilon(x)}{Lip(f,x)}\right) \\
&\leq \varepsilon(x).
\end{aligned}
$$

Since additionally $|x' - x| < \varepsilon(x)$, and recalling $\zeta' = \nabla\tilde{f}(x')$, we conclude that $\zeta \in \nabla\tilde{f}(x + \varepsilon(x)B) + \varepsilon(x)B$. Hence $\partial_P f(x) \subset \nabla\tilde{f}(x + \varepsilon(x)B \cap \Omega) + \varepsilon(x)B$. By considering the hypographs of f and $\tilde{f}$ for vectors in $\partial^P f(x)$, we obtain $\partial^P f(x) \subset \nabla\tilde{f}(x + \varepsilon(x)B \cap \Omega) + \varepsilon(x)B$, which gives (4.3). $\qquad\square$

Remark 4.2. Note that if ρ is constant and if f is convex, the function $\tilde{f}$ is also convex on $\mathbb{R}^N$. As a matter of fact, in this case we can verify that for any $x, y \in \mathbb{R}^N$ and $\lambda \in [0, 1]$,

$$
\begin{aligned}
\tilde{f}(\lambda x + (1 - \lambda)y) &= \int_{\overline{B}} \theta(t)f(\lambda x + (1 - \lambda)y - \rho t)dt \\
&\leq \int_{\overline{B}} \theta(t)\Big(\lambda f(x - \rho t) + (1 - \lambda)f(y - \rho t)\Big)dt \\
&\leq \lambda\tilde{f}(x) + (1 - \lambda)\tilde{f}(y).
\end{aligned}
$$

Since in this case $\nabla\rho = 0$, we obtain (4.4) from (4.20).

4.2. Local approximation.

Lemma 4.3. *Let Ω be an open subset of $\mathbb{R}^N$, let $f : \Omega \to \mathbb{R}$ be a locally Lipschitz function, and let Let $\overline{x} \in \Omega$ and $\varepsilon > 0$ such that $\overline{B}(\overline{x}, \varepsilon) \subset \Omega$. Denote by L the Lipschitz constant of f on the set $\overline{B}(\overline{x}, \varepsilon)$. For every $\alpha > 0$, there exists a function $\tilde{g} : B(\overline{x}, \varepsilon) \to \mathbb{R}$ of class C^∞ such that, for every $x \in B(\overline{x}, \varepsilon)$,*

$$
\begin{aligned}
(4.22) && \tilde{g}(x) &\geq f(x) + \varepsilon|x - \overline{x}| - \alpha, \\
(4.23) && \tilde{g}(x) &\leq f(x) + (2L + \varepsilon)|x - \overline{x}| + \alpha, \\
(4.24) && \nabla\tilde{g}(x) &\in \partial f(\overline{x} + \varepsilon B) + \varepsilon B, \\
(4.25) && \partial f(x) + \varepsilon B &\subset \nabla\tilde{g}(\overline{x} + \alpha B) + \alpha B.
\end{aligned}
$$

Proof of Lemma 4.3. Set

$$
K := \partial f(\overline{x} + \varepsilon\overline{B}) + \varepsilon\overline{B}
$$

and define the function $g : \mathbb{R}^N \mapsto \mathbb{R}$ by, for every $x \in \mathbb{R}^N$,

$$
(4.26) \qquad g(x) := f(\overline{x}) + \sigma_K(x - \overline{x}),
$$

where σ_K is the support function of the convex set K, precisely defined by

$$
\sigma_K(x) := \max\{\langle x, p \rangle : p \in K\}.
$$

We now claim that the function g is convex and satisfies, for every $x \in \mathbb{R}^N$,

$$(4.27) \qquad g(x) \geq f(x) + \varepsilon|x - \overline{x}|,$$

$$(4.28) \qquad g(x) \leq f(x) + (2L + \varepsilon)|x - \overline{x}|,$$

$$(4.29) \qquad \partial g(x) \subset K,$$

$$(4.30) \qquad \partial g(\overline{x}) = K.$$

Indeed, the convexity of g, and assertions (4.29) and (4.30) are direct consequences of the properties of the support function σ_K of the convex set K.[7] $\square$

Lemma 4.4. *For any $q \in \mathbb{R}^N, \partial\sigma_K(q) \subset K$ and in particular, $\partial\sigma_K(0) = K$.*

Let us now prove (4.27) and (4.28). Let $x \in B(\overline{x}, \varepsilon)$, we can compute

$$\begin{aligned} g(x) - f(x) &= f(\overline{x}) - f(x) + \sigma_K(x - \overline{x}) \\ (4.31) \qquad &= \langle \zeta, \overline{x} - x \rangle + \sigma_K(x - \overline{x}), \end{aligned}$$

for some $\zeta \in \partial f(z)$ and $z \in [\overline{x}, x]$, in view of Lebourg's theorem (see for example [8]). Since $[\overline{x}, x] \subset \overline{x} + \varepsilon\overline{B}$, we have $\zeta \in \partial f(\overline{x} + \varepsilon\overline{B})$ and we deduce

$$\zeta + \varepsilon\frac{x - \overline{x}}{|x - \overline{x}|} \in \partial f(\overline{x} + \varepsilon\overline{B}) + \varepsilon\overline{B} = K.$$

From the definition of σ_K, we have $\sigma_K(x - \overline{x}) \geq \langle x - \overline{x}, \zeta + \varepsilon\frac{x-\overline{x}}{|x-\overline{x}|} \rangle$, and then (4.31) implies

$$(4.32) \qquad g(x) - f(x) \geq \langle \zeta, \overline{x} - x \rangle + \langle \zeta + \varepsilon\frac{x - \overline{x}}{|x - \overline{x}|}, x - \overline{x} \rangle = \varepsilon|x - \overline{x}|,$$

which proves (4.27). In view of Cauchy-Schwarz inequality

$$\sigma_K(x - \overline{x}) \leq \max_{k \in K}|k||x - \overline{x}|.$$

Since $K = \partial f(\overline{x} + \varepsilon\overline{B}) + \varepsilon\overline{B}$ and since f is L-Lipschitz on the set $\overline{B}(\overline{x}, \varepsilon)$, we deduce that

$$\sigma_K(x - \overline{x}) \leq (L + \varepsilon)|x - \overline{x}|.$$

From (4.26), and since $f(\overline{x}) \leq f(x) + L|x - \overline{x}|$, we obtain

$$(4.33) \qquad g(x) \leq f(x) + (2L + \varepsilon)|x - \overline{x}|,$$

which proves (4.28). In view of Proposition 4.1, we can regularize the convex function g and obtain a smooth function $\tilde{g} : B(\overline{x}, \varepsilon) \to \mathbb{R}$ such that for every $x \in B(\overline{x}, \varepsilon)$,

$$(4.34) \qquad |\tilde{g}(x) - g(x)| \leq \alpha,$$

$$(4.35) \qquad \nabla\tilde{g}(x) \subset \operatorname{co}\partial g(x + \alpha B),$$

$$(4.36) \qquad \partial g(x) \subset \nabla\tilde{g}(x + \alpha B) + \alpha B.$$

By definition, the function $\tilde{g}$ is of class C^∞. Assertions (4.22) and (4.23) are immediate consequences of (4.27) and (4.28) combined with (4.34). Let us now verify the properties concerning the gradients. Since $\nabla\tilde{g}(x) \in \operatorname{co}\partial g(x + \alpha B)$ and $\partial g(x) \subset K$, we obtain $\nabla\tilde{g}(x) \in \partial f(\overline{x} + \varepsilon\overline{B}) + \varepsilon\overline{B}$, which proves the inclusion (4.24). Since $\partial f(x) \subset K$, $K = \partial g(\overline{x})$, and $\partial g(\overline{x}) \subset \nabla\tilde{g}(\overline{x} + \alpha B) + \alpha B$, we deduce (4.25).

---

[7] We refer to [4] for the following lemma.

4.3. Global upper C^2 approximation. In this section, we complete the proof of Theorem 2.2. Recall that $f : \Omega \to \mathbb{R}$ is a locally Lipschitz function and $\varepsilon : \Omega \to \mathbb{R}_+ \setminus \{0\}$ is a continuous function. We shall construct an upper C^2 approximation of f on the whole set Ω by combining together the local approximations given by the previous section. So we first need to define the points around which we will consider local approximations (given by Lemma 4.3), and the associated constants (named ε and α in Lemma 4.3). For every $x \in \Omega$, we set

$$\delta(x) := \min\left\{\frac{1}{2}d(\mathbb{R}^N \setminus \Omega, x), 1\right\}.$$

Since the set Ω is open, then $\delta(x) > 0$ and clearly $\overline{B}(x, \delta(x)) \subset \Omega$. We recall (see the proof of Proposition 4.1) that there exists a continuous function $Lip(f, \cdot) : \Omega \to [1, +\infty)$ such that for every $x \in \Omega$, f is $Lip(f, x)$-Lipschitz on $\overline{B}(x, \delta(x))$. For every $x \in \Omega$, we set

$$(4.37) \qquad \varepsilon_1(x) \quad := \quad \min\left\{\frac{1}{2}\min\{\varepsilon(x'), x' \in \overline{B}(x, \delta(x))\}, \delta(x)\right\},$$

$$(4.38) \qquad \varepsilon_2(x) \quad := \quad \min\{\varepsilon_1(x'), x' \in \overline{B}(x, \delta(x))\}.$$

Since the functions $x \mapsto \varepsilon(x)$ and $x \mapsto \delta(x)$ are continuous, so are the functions $x \mapsto \varepsilon_1(x)$ and $x \mapsto \varepsilon_2(x)$. Moreover as an immediate consequence of [11, Lemma 4.3], there exists a function $\rho : \Omega \to R_+ \setminus \{0\}$ of class C^∞ such that for every $x \in \Omega$,

$$(4.39) \qquad \rho(x) \quad \leq \quad \min\left\{\frac{\delta(x)}{3}, \frac{\varepsilon_2(x)}{24Lip(f, x)}\right\},$$

$$(4.40) \qquad |\nabla\rho(x)| \quad \leq \quad \frac{1}{4}.$$

Since the function ρ has positive values, the family $(B(x, \rho^2(x)))_{x \in \Omega}$ is an open covering of Ω. Thus it admits a locally finite refinement $(B(y, \rho^2(y)))_{y \in Y}$. Precisely, there exists a locally finite set $Y \subset \Omega$ such that

$$\Omega \subset \bigcup_{y \in Y} B\left(y, \rho^2(y)\right).$$

By construction of the function ρ, for every $y \in \Omega$ the ball $\overline{B}(y, 3\rho(y))$ is included in Ω. Hence we can define for every $y \in Y$, the set $Z(y)$ by

$$Z(y) := \{y' \in Y, y' \in \overline{B}(y, 3\rho(y)) \text{ and } y' \neq y\}.$$

Since the set Y is locally finite and since, for every $y \in Y$, the set $\overline{B}(y, 3\rho(y))$ is compact and the set $Z(y)$ is finite. For every $y \in Y$, we define[8]

$$(4.41) \quad \alpha_y := \varepsilon_2(y) . \min\left\{\frac{\min\{\rho(y'), y' \in Z(y) \cup \{y\}\}}{16}, \frac{\min\{|y - y'|, y' \in Z(y)\}}{6Lip(f, y) + 4}\right\}.$$

For every $y \in Y$, note that by construction, $\alpha_y > 0$, $\varepsilon_1(y) > 0$, $B(y, \varepsilon_1(y)) \subset B(y, \delta(y)) \subset \Omega$. We apply Lemma 4.3 with the function f and with the constants $\varepsilon = \varepsilon_1(y)$ and $\alpha = \alpha_y$. Thus for every $y \in Y$, we obtain a function $\tilde{g}_y : B(y, \varepsilon_1(y)) \to \mathbb{R}$ which is of class C^∞ on $B(y, \varepsilon_1(y))$ and such that (4.27)-(4.30) are satisfied. For every $x \in \Omega$, since $\varepsilon_2(x) \leq \varepsilon_1(x)$ and $1 \leq Lip(f, x)$, and from the definition of the function ρ, we have

$$\rho(x) \leq \varepsilon_1(x).$$

---

[8]We take the convention $\min \emptyset = +\infty$. But in view of the forthcoming Lemma 4.6, of (4.3), and since $\rho(y)^2 < \rho(y)$, we can deduce that $Z(y) \neq \emptyset$.

Hence, for every $y \in Y$, the function $\tilde{g}_y$ is defined on $B(y, \rho(y))$. Besides, note that since $\rho^2 \leq \rho \leq 1$, the family $(B(y, \rho(y)))_{y \in Y}$ defines an open covering of Ω. Hence we can define on Ω the map $h : \Omega \to \mathbb{R}$ by

$$h(x) := \inf\{\tilde{g}_y(x), y \in Y \text{ such that } x \in B(y, \rho(y))\}.$$

Lemma 4.5. *The function h is locally a minimum of a finite number of functions of class C^∞, hence locally Lipschitz and upper C^2, and it satisfies for every $x \in \Omega$,*

$$(4.42) \qquad |f(x) - h(x)| \leq \varepsilon(x),$$

and

$$(4.43) \qquad \partial h(x) \quad \in \quad \operatorname{co}\partial f(x + \varepsilon(x)B \cap \Omega) + \varepsilon(x)B,$$

$$(4.44) \qquad \partial f(x) \quad \subset \quad \partial h(x + \varepsilon(x)B \cap \Omega) + \varepsilon(x)B.$$

Proof of Lemma 4.5. Let us prove that the function h is locally a minimum of a finite number of functions of class C^∞. In order to do that, we first compare $\tilde{g}_y(x)$ and $\tilde{g}_{y'}(x)$ for $x \in \Omega$, $y \neq y'$ in Y such that $x \in B(y, \rho(y)) \cap B(y', \rho(y'))$. From Lemma 4.3, we write

$$\tilde{g}_{y'}(x) - \tilde{g}_y(x) \geq \varepsilon_1(y')|x - y'| - (2Lip(f, y) + \varepsilon_1(y))|x - y| - \alpha_y - \alpha_{y'}.$$

Since $\varepsilon_1(y) \leq 1 \leq Lip(f, y)$, we deduce

$$(4.45) \qquad \tilde{g}_{y'}(x) - \tilde{g}_y(x) \geq \varepsilon_1(y')|x - y'| - 3Lip(f, y)|x - y| - \alpha_y - \alpha_{y'}.$$

For every $x \in \Omega$, set

$$I(x) := \left\{ y \in Y, x \in B\left(y, \frac{\rho(y)}{2}\right) \right\} \text{ and } J(x) := \left\{ y \in Y, x \in B\left(y, \frac{2\rho(y)}{3}\right) \right\}.$$

We claim that the sets $I(x)$ and $J(x)$ are finite. Indeed, let us first prove the inclusions

$$(4.46) \qquad I(x) \subset J(x) \subset \{y \in Y, x \in B(y, \rho(y))\} \subset \{y \in Y, y \in \overline{B}(x, \delta(x))\}.$$

The first two inclusions are immediate. Consider now $y \in Y$ such that $x \in B(y, \rho(y))$. From (4.39), $\rho(y) \leq \delta(y)$, hence $x \in \overline{B}(y, \delta(y))$. From (4.38) and (4.37), we deduce $\varepsilon_2(y) \leq \varepsilon_1(x) \leq \delta(x)$. Hence by (4.39) this implies

$$\rho(y) \leq \frac{\varepsilon_2(y)}{24Lip(f, y)} \leq \varepsilon_1(x) \leq \delta(x).$$

In conclusion we deduce

$$y \in \overline{B}(x, \delta(x))$$

which proves the last inclusion. Since the covering $(B(y, \rho^2(y)))_{y \in Y}$ is locally finite and since the set $\overline{B}(x, \delta(x))$ is a compact subset of Ω, the set

$$\{y \in Y, \overline{B}(x, \delta(x)) \cap B(y, \rho^2(y)) \neq \emptyset\}$$

is finite and nonempty. Moreover, this set contains the set $\{y \in Y, y \in \overline{B}(x, \delta(x))\}$. In view of (4.46), we conclude that the sets $I(x)$ and $J(x)$ are finite.

We now claim the following properties on the function h:

$$(4.47) \qquad h(x) \quad = \quad \min\{\tilde{g}_y(x), y \in I(x)\}, \forall x \in \Omega,$$

$$(4.48) \qquad h(x') \quad = \quad \min\{\tilde{g}_y(x'), y \in J(x)\}, \forall x \in \Omega, \forall x' \in B(x, \frac{\rho(x)}{8}),$$

$$(4.49) \qquad h(x) \quad = \quad \tilde{g}_y(x), \forall y \in Y, \forall x \in B(y, \alpha_y).$$

Let us first prove the first equation. If $y' \in Y \setminus I(x)$ (i.e., $x \notin B(y', \frac{\rho(y')}{2})$) and $x \in B(y', \rho(y'))$, take $y \in Y$ such that $x \in B(y, \rho(y)^2)$. From (4.45) we have

$$\tilde{g}_{y'}(x) - \tilde{g}_y(x) \geq \varepsilon_1(y')\frac{\rho(y')}{2} - 3Lip(f,y)\rho(y)^2 - \alpha_y - \alpha_{y'}.$$

$\square$

Lemma 4.6. *Let y and y' be in Y such that $B(y, \rho(y)) \cap B(y', \rho(y')) \neq \emptyset$. Then*

$$\frac{1}{2}\rho(y) \leq \rho(y') \ and \ y' \in B(y, 3\rho(y)).$$

Proof of Lemma 4.6. If $\rho(y) \leq \rho(y')$, the result is clear. If $\rho(y') < \rho(y)$, since $B(y, \rho(y)) \cap B(y', \rho(y')) \neq \emptyset$, we have $|y - y'| \leq \rho(y) + \rho(y') < 2\rho(y)$. From the mean-value theorem, we deduce

$$\rho(y) - \rho(y') \leq \sup\{|\nabla\rho(z)|, z \in B(y, 2\rho(y))\}.2\rho(y).$$

From (4.40), $|\nabla\rho| \leq \frac{1}{4}$, hence $\rho(y) - \rho(y') \leq \frac{1}{2}\rho(y)$, hence

$$\frac{1}{2}\rho(y) \leq \rho(y'). \qquad \square$$

We now return to the proof of Lemma 4.5. Since $B(y, \rho(y)) \cap B(y', \rho(y')) \neq \emptyset$, from Lemma 4.6, $\frac{1}{2}\rho(y) \leq \rho(y')$ and we deduce

$$\tilde{g}_{y'}(x) - \tilde{g}_y(x) \geq \varepsilon_1(y')\frac{\rho(y)}{4} - 3Lip(f,y)\rho(y)^2 - \alpha_y - \alpha_{y'}.$$

From Lemma 4.6, $y' \in B(y, 3\rho(y))$. In view of (4.39) $\rho(y) \leq \frac{\delta(y)}{3}$, hence $y' \in B(y, \delta(y))$. Hence, from (4.37), $\varepsilon_2(y) \leq \varepsilon_1(y')$. Since from (4.39), $\rho(y) \leq \frac{\varepsilon_2(y)}{32Lip(f,y)}$, we have, $\rho(y) \leq \frac{\varepsilon_1(y')}{24Lip(f,y)}$. Hence

$$\tilde{g}_{y'}(x) - \tilde{g}_y(x) \quad \geq \quad \frac{\rho(y)}{8}\Big(\varepsilon_1(y') - 24Lip(f,y)\rho(y)\Big) + \varepsilon_1(y')\frac{\rho(y)}{8} - \alpha_y - \alpha_{y'}$$

$$(4.50) \qquad\qquad \geq \quad \varepsilon_1(y')\frac{\rho(y)}{8} - \alpha_y - \alpha_{y'}.$$

From (4.41), since $\varepsilon_2(y) \leq \varepsilon_1(y')$ and clearly $y \in Z(y) \cap \{y\}$, we deduce

$$\alpha_y \leq \varepsilon_1(y')\frac{\rho(y)}{16}.$$

From Lemma 4.6, $y \in B(y', 3\rho(y'))$, hence $y \in Z(y') \cap \{y'\}$. Hence from (4.41), since clearly $\varepsilon_2(y') \leq \varepsilon_1(y')$, we deduce

$$\alpha_{y'} \leq \varepsilon_1(y')\frac{\rho(y)}{16}.$$

In view of (4.50), we deduce

$$\tilde{g}_{y'}(x) - \tilde{g}_y(x) \geq 0,$$

which ends the proof of the first equation.

We now prove the second equation. Let $x' \in B(x, \frac{\rho(x)}{8})$, and let $y \in I(x')$. Then $|x' - y| \leq \frac{\rho(y)}{2}$ and $|x - y| < \frac{\rho(x)}{8}$, hence

$$(4.51) \qquad\qquad\qquad |x - y| \leq \frac{\rho(y)}{2} + \frac{\rho(x)}{8}.$$

We distinguish the two cases $\rho(x) \leq \rho(y)$ and $\rho(x) > \rho(y)$. If $\rho(x) \leq \rho(y)$, then from (4.51), $|x - y| \leq \frac{5}{8}\rho(y) \leq \frac{2}{3}\rho(y)$, hence $y \in J(x)$. If $\rho(x) > \rho(y)$, then $|x - y| < \frac{5}{8}\rho(x)$. From (4.40), $|\nabla\rho| \leq \frac{1}{4}$, hence from the mean-value theorem, we deduce $\rho(x) - \rho(y) \leq \frac{1}{4}|x - y| \leq \frac{5}{16}\rho(x)$. Hence

$$\frac{9}{16}\rho(x) \leq \rho(y).$$

Then from (4.51), $|x - y| \leq \frac{73}{128}\rho(y) \leq \frac{2}{3}\rho(y)$, hence $y \in J(x)$. We proved the inclusion $I(x') \subset J(x)$, which clearly implies the second equation.

For the third equation, take $y \in Y$, $x \in B(y, \alpha_y)$ and $y' \neq y$. From (4.45),

$$\tilde{g}_{y'}(x) - \tilde{g}_y(x) \geq \varepsilon_1(y')|y - y'| - (3Lip(f, y) + \varepsilon_1(y'))|x - y| - \alpha_y - \alpha_{y'}$$
$$(4.52) \qquad\qquad \geq \varepsilon_1(y')|y - y'| - (3Lip(f, y) + 2)\alpha_y - \alpha_{y'},$$

recalling that $\varepsilon_1(y') \leq 1$. From Lemma 4.6, $y' \in B(y, 3\rho(y))$ and $y \in B(y', 3\rho(y'))$, hence $y' \in Z(y)$ and $y \in Z(y')$. From (4.41), since $\varepsilon_2(y) \leq \varepsilon_1(y')$, we deduce

$$\alpha_y \leq \varepsilon_1(y')\frac{|y - y'|}{6Lip(f, y) + 4}.$$

Again from (4.41), since $y \in Z(y')$ and clearly $\varepsilon_2(y') \leq \varepsilon_1(y')$, we deduce

$$\alpha_{y'} \leq \varepsilon_1(y')\frac{|y - y'|}{6Lip(f, y') + 4} \leq \varepsilon_1(y')\frac{|y - y'|}{2}.$$

In view of (4.52), we deduce

$$\tilde{g}_{y'}(x) - \tilde{g}_y(x) \geq \varepsilon_1(y')|y - y'|\left(1 - \frac{1}{2} - \frac{1}{2}\right) \geq 0,$$

which ends the proof of the third equation.

Let us now verify the conclusions of Lemma 4.5. From (4.47), the function h is locally a minimum of a finite number of functions of class C^∞. For every $x \in \Omega$, from the definition of the function h, there exists $y \in Y$ such that $x \in B(y, \rho(y))$ and $h(x) = \tilde{g}_y(x)$. On the other hand, by Lemma 4.3 we have

$$f(x) + \varepsilon_1(y)|x - y| - \alpha_y \leq \tilde{g}_y(x) \leq f(x) + (2Lip(f, y) + \varepsilon_1(y))|x - y| + \alpha_y.$$

From (4.41), and since clearly $y \in Z(y) \cup \{y\}$, we obtain

$$\alpha_y \leq \varepsilon_1(y)\frac{\rho(y)}{16} \leq \frac{\varepsilon_1(y)}{16}.$$

In addition, (4.39) yields

$$\rho(y) \leq \frac{\varepsilon_1(y)}{24Lip(f, y)}.$$

Since $|x - y| < \rho(y)$ and $h(x) = \tilde{g}_y(x)$, we deduce

$$f(x) - \varepsilon_1(y) \leq h(x) \leq f(x) + \varepsilon_1(y).$$

Recall that $x \in B(y, \rho(y))$ implies $x \in \overline{B}(y, \delta(y))$. Hence we obtain $\varepsilon_1(y) \leq \frac{\varepsilon(x)}{2}$ which implies (4.42). Since the function h is locally a minimum of a finite number of C^∞ functions $\tilde{g}_y$, one elementarily shows (see for example [9, ex. 11.17, p. 48, ex. 4.6, p. 83])

$$(4.53) \qquad\qquad \partial^P h(x) = \partial h(x) = \mathrm{co} \bigcup_{y \in Y, \tilde{g}_y(x) = h(x)} \{\nabla\tilde{g}_y(x)\}.$$

By (4.48) we can write,

$$\partial h(x) \subset \text{co} \bigcup_{y \in J(x)} \{\nabla \tilde{g}_y(x)\} \subset \text{co} \bigcup_{y \in J(x)} \partial f(y + \varepsilon_1(y)B) + \varepsilon_1(y)B.$$

Since $\rho(y) \leq \varepsilon_1(y) \leq \frac{\varepsilon(x)}{2}$, we deduce that $y + \varepsilon_1(y)B \subset B(x, \varepsilon(x))$, hence

$$\partial h(x) \subset \text{co}\, \partial f(x + \varepsilon(x)B) + \varepsilon(x)B,$$

which proves (4.43).

Let us now prove (4.44). Consider $x \in \Omega$ and $y \in Y$ such that $x \in B(y, \rho(y)^2)$. From Lemma 4.3, $\partial f(x) \subset \nabla \tilde{g}_y(y + \alpha_y B) + \alpha_y B$. But from (4.49), for every $x' \in B(y, \alpha_y)$, $\nabla h(x') = \nabla \tilde{g}_y(x')$. Hence

$$\partial f(x) \subset \nabla h(y + \alpha_y B) + \alpha_y B.$$

Since $\rho(y)^2 \leq \rho(y)\delta(y) \leq 1$ we have $x \in B(y, \delta(y))$, hence $\varepsilon_1(y) \leq \varepsilon(x)$. In addition, from (4.41), $\alpha_y \leq \frac{\varepsilon_2(y)}{16}$, and it is easy to verify that $\alpha_y \leq \varepsilon(x)$ and $y + \alpha_y B \subset x + \varepsilon(x)B$. In conclusion, for every $x \in \Omega$, $\partial f(x) \subset \partial h(x + \varepsilon(x)B) + \varepsilon(x)B$, and (4.44) is proved. $\square$

4.4. Conclusion.

We now apply Proposition 4.1 to the map h defined in the previous section, and we obtain a function $f_\varepsilon : \Omega \to \mathbb{R}$ of class C^∞ such that for every $x \in \Omega$

$$\begin{aligned}
&|f_\varepsilon(x) - h(x)| \leq \varepsilon(x), \\
&\nabla f_\varepsilon(x) \subset \partial h(x + \varepsilon(x)B \cap \Omega) + \varepsilon(x)B, \\
&\partial^P h(x) \subset \nabla f_\varepsilon(x + \varepsilon(x)B \cap \Omega) + \varepsilon(x)B.
\end{aligned}$$

From (4.53), $\partial h(x) = \partial^P h(x)$. In view of Lemma 4.5, and assuming without any loss of generality that we considered the function $\varepsilon/2$ instead of the function ε, the function f_ε satisfies the conclusions of Theorem 2.2.

5. CONVERGENCE OF SETS: PROOF OF THEOREM 2.3 AND PROPOSITION 2.2

Going from the functional point of view to the set point of view should be quite straightforward, but it usually proves longer than first expected. In our case, epi-Lipschitz sets can be locally written, up to an isometry, as the epigraph of a Lipschitz function, to which the previous regularization (Theorem 2.2) can be applied. This is the scheme of our proof, and the difficulty lies mainly in gluing together the regularized like we do, but so introducuing an obnoxious step (related to (5.11) below). Also, one can work more accurately, using Corollary 3.2, but thus obtaining only a C^1 approximation to be regularized itself. A tempting approach is to use a representation of the set M, for example the distance function d_M or the signed distance function, and to regularize it with Theorem 2.2. But through this approach, one is not able to retrieve information on the precise location of the gradients of the regularized representation. At present, this point prevents us from obtaining Theorem 2.2 for a wider class of sets (for example proximally nondegenerate sets). Also, proving Theorem 2.3 without using functions by copying the technique of the proof of Theorem 2.2 would quite surely amount to a much longer proof.

5.1. **Local representations and their approximations.** For every $x \in \mathbb{R}^N$, we set

$$(5.1) \qquad \varepsilon'(x) := \frac{1}{3} \min \left\{ \min \left\{ \varepsilon(x'), x' \in \overline{B}(x,1) \right\}, 1 \right\}.$$

Since the set M is epi-Lipschitz, the correspondence $x \mapsto N_M^C(x) \cap S$ is upper semicontinuous with compact values, and there exists a sequence $(x_n)_{n \in \mathbb{N}}$ in M and a sequence $(r_n)_{n \in \mathbb{N}}$ of positive real numbers such that $(B(x_n, \frac{r_n^2}{2}))_{n \in \mathbb{N}}$ is a locally finite covering of M and, for every n,

$$(5.2) \qquad N_M^C \left(B(x_n, 2r_n) \cap M \right) \cap S \subset N_M^C(x_n) \cap S + B\left(0, \varepsilon'(x_n)\right),$$

$$(5.3) \qquad r_n \leq \varepsilon'(x_n).$$

Take $e_n \in \operatorname{int} T_M^C(x_n) \cap S$. From (5.2), $e_n \in \operatorname{int} T_M^C(B(x_n, 2r_n) \cap M)$, and in view of Rockafellar [33], we may assume without any loss of generality that

$$(5.4) \qquad y + tw \in M \text{ for every } (y, w, t) \in \overline{B}(x_n, 2r_n) \times \overline{B}(e_n, 2r_n) \times [0, 2r_n].$$

For every n, we define the function $\lambda_n : B(x_n, r_n^2) \to \mathbb{R}$ by

$$(5.5) \qquad \lambda_n(x) := \inf \left\{ t \in \mathbb{R} | x + te_n \in M \cap B(x_n, r_n^2) \right\}.$$

The fact that the function λ_n is well defined (together with the following assertions) is detailed in [16, Chapitre 1, Appendix] and partially stated in [11, Proposition 4.4]:

$$(5.6) \qquad \text{The function } \lambda_n \text{ is } \frac{1}{r_n} - \text{Lipschitz,}$$

$$(5.7) \qquad M \cap B(x_n, r_n^2) = \{x \in B(x_n, r_n^2) | \lambda_n(x) \leq 0\},$$

$$(5.8) \qquad \operatorname{bd} M \cap B(x_n, r_n^2) = \{x \in B(x_n, r_n^2) | \lambda_n(x) = 0\},$$

$$(5.9) \qquad \lambda_n(x + te_n) = \lambda_n(x) - t \text{ for every } x \in B(x_n, r_n^2) \text{ and } t \in \mathbb{R}$$
$$\text{such that } x + te_n \in B(x_n, r_n^2),$$

$$(5.10) \qquad \partial \lambda_n(x) = N_M(x + \lambda_n(x)e_n) \cap \{v \in \mathbb{R}^N | \langle v, e_n \rangle = -1\}.$$

Equation (5.9) highlights the fact that the function λ_n is built by a projection on the hyperplane $e_n^\perp$. Precisely (see [16, Chapitre 1, Appendix, claim 5.5]), let $\pi_1 : \mathbb{R}^{N-1} \times \mathbb{R} \to \mathbb{R}^{N-1}$, respectively $\pi_2 : \mathbb{R}^{N-1} \times \mathbb{R} \to \mathbb{R}$, be the projection defined by $\pi_1(x_1, x_2) = x_1$, respectively $\pi_2(x_1, x_2) = x_2$. We define the linear map $A_n : \mathbb{R}^N \to \mathbb{R}^{N-1} \times \mathbb{R}$ by $A_n(u + \lambda e_n) = (L_n(u), \lambda)$ for all $(u, \lambda) \in e_n^\perp \times \mathbb{R}$, and if L_n is an isometry from $e_n^\perp$ to $\mathbb{R}^{N-1}$. Let $U_n = B(x_n, r_n^2)$ and let $V_n = \pi_1(A_n(U_n))$. If $x \in V_n$, and if $(y, z) \in \mathbb{R}^2$ are such that $(x, y) \in A_n(U_n)$ and $(x, z) \in A_n(U_n)$, then clearly $A_n^{-1}(x, y) - yv = A_n^{-1}(x, z) - zv$, hence $\lambda_n(A_n^{-1}(x, y)) + y = \lambda_n(A_n^{-1}(x, z)) + z$. Hence the function $\varphi_n : V_n \to \mathbb{R}$ defined by $\varphi_n(x) = \lambda_n(A_n^{-1}(x, y)) + y$ if $(x, y) \in A_n(U_n)$ is well defined. The map φ_n is clearly Lipschitzian, and $\lambda_n = \varphi_n \circ \pi_1 \circ A_n|_{U_n} - \pi_2 \circ A|_{U_n}$.

Now consider a sequence of positive real numbers $(r_n')_{n \in \mathbb{N}}$ such that $r_n' \in (0, \frac{r_n^2}{2})$ and

$$\overline{B}(x_n, r_n') \cap \bigcup_{p \neq n} \overline{B}(x_p, r_p') = \emptyset.$$

Let $(\alpha_n)_{n \in \mathbb{N}}$ be a C^∞ partition of unity subordinated to the covering $(B(x_n, \frac{r_n^2}{2}))_{n \in \mathbb{N}}$ and such that

$$\alpha_n|_{\overline{B}(x_n, r_n')} = 1.$$

Remark that the function $f := \sum_{n \in \mathbb{N}} \alpha_k \lambda_k$ is "almost" a Lipschitzian inequality representation of M on $\mathbb{R}^N$, in the sense that it satisfies assertions (i) to (iv) of [11,

Definition 2.1]. I.e., $M = \{x \in \mathbb{R}^N | f(x) \leq 0\}$ and $N_M(x) = \bigcup_{\lambda \geq 0} \lambda \partial f(x)$ together with other nice properties. Our proof consists in smoothing the functions λ_n rather than the function f itself, for reasons that are announced at the beginning of the section and that will precisely appear below. For every $x \in \mathrm{bd}\, M$, $\sum_{k \in \mathbb{N}} \alpha_k(x) \partial \lambda_k(x)$ is a nonempty compact subset of $N_M(x) \setminus \{0\}$ (see [16]), and we set

$$\delta(x) := d\left(\sum_{k \in \mathbb{N}} \alpha_k(x) \partial \lambda_k(x), 0 \right).$$

The correspondences $\partial \lambda_k$ are upper semicontinuous with compact values, the functions α_k are continuous, and locally we take the sum of a finite number of terms, hence there exists a sequence $(y_n)_{n \in \mathbb{N}}$ in $\mathrm{bd}\, M$, and a sequence $(\rho_n)_{n \in \mathbb{N}}$ of positive real numbers in $(0, 1)$ such that $(B(y_n, \rho_n))_{n \in \mathbb{N}}$ is a locally finite covering of $\mathrm{bd}\, M$ and, for every n,

(5.11)

$$\sum_{k \in \mathbb{N}} \alpha_k(B(y_n, \rho_n)) \partial \lambda_k(B(y_n, \rho_n)) \subset \sum_{k \in \mathbb{N}} \alpha_k(y_n) \partial \lambda_k(y_n) + B\left(0, \frac{\delta(y_n)}{3} \varepsilon'(y_n) \right).$$

Since the family $(y_n)_{n \in \mathbb{N}}$, respectively $(x_n)_{n \in \mathbb{N}}$, is locally finite and since, for every $n \in \mathbb{N}$, the set $\overline{B}(x_n, 1)$ is compact, the set $\{p | y_p \in \overline{B}(x_n, 1)\}$, respectively $\{k | x_k \in \overline{B}(x_n, 1)\}$, is finite. We set, for every $n \in \mathbb{N}$,

(5.12) $\quad \varepsilon_n := \min \left\{ \varepsilon'(x_n), r_n' r_n, \frac{1}{2} \min\{r_k^2, x_k \in \overline{B}(x_n, 1)\}, \right.$

$$\frac{\varepsilon'(x_n)}{\mathrm{card}\, \{k | x_k \in \overline{B}(x_n, 1)\}} \min \left\{ \frac{r_k}{2\|\nabla \alpha_k\|_\infty}, x_k \in \overline{B}(x_n, 1) \right\}$$

$$\left. \cdot \min \left\{ \frac{\delta(y_p)}{3}, y_p \in \overline{B}(x_n, 1) \right\} \right\}.$$

From Theorem 2.2, there exists a C^∞ regularization φ_n' of the function φ_n associated to the map λ_n, which yields a C^∞ regularization $\lambda_n' : B(x_n, r_n^2) \to \mathbb{R}$ of λ_n by setting $\lambda_n' = \varphi_n' \circ \pi_1 \circ A_n|_{U_n} - \pi_2 \circ A|_{U_n}$. The function λ_n' satisfies, for every $x \in B(x_n, r_n^2)$,

(5.13) $\qquad\qquad |\lambda_n'(x) - \lambda_n(x)| \quad \leq \quad \varepsilon_n,$

(5.14) $\qquad\qquad\qquad \nabla \lambda_n'(x) \quad \in \quad \partial \lambda_n(B(x, \varepsilon_n)) + B(0, \varepsilon_n),$

(5.15) $\qquad\qquad\qquad \partial \lambda_n(x) \quad \subset \quad \nabla \lambda_n'(B(x, \varepsilon_n)) + B(0, \varepsilon_n),$

and additionally

(5.16) $\qquad \lambda_n'(x + te_n) \quad = \quad \lambda_n'(x) - t$ for every $x \in B(x_n, r_n^2)$

$\qquad\qquad\qquad\qquad\qquad$ and $t \in \mathbb{R}$ such that $x + te_n \in B(x_n, r_n^2)$,

(5.17) $\qquad \nabla \lambda_n'(x + te_n) \quad = \quad \nabla \lambda_n'(x) - t$ for every $x \in B(x_n, r_n^2)$

$\qquad\qquad\qquad\qquad\qquad$ and $t \in \mathbb{R}$ such that $x + te_n \in B(x_n, r_n^2)$.

5.2. **The regularized set and proof of Theorem 2.3.** We set

(5.18) $\qquad\qquad\qquad f' := \sum_{k \in \mathbb{N}} \alpha_k \lambda_k'$ and $M' = \{x \in \Omega | f'(x) \leq 0\},$

with $\Omega = \bigcup_{n \in \mathbb{N}} B(x_n, \frac{r_n^2}{2})$, and we now show that the set M' satisfies the conclusions of Theorem 2.3. We first prove that the set M' is smooth and afterwards the estimations and inclusions (2.9) to (2.12).

5.2.1. *The set M' is smooth.* The function f' is clearly of class C^∞, and we now prove that $\nabla f'(x) \neq 0$ for every $x \in \Omega$ such that $f'(x) = 0$, which implies by definition that the set M' is a closed C^∞ submanifold of $\mathbb{R}^N$ with a boundary of full dimension. For every $x \in \Omega$, we set

$$I(x) := \{n \in \mathbb{N} | \alpha_n(x) > 0\}$$

and we note that, from the definition of the partition $(\alpha_n)_{n \in \mathbb{N}}$, we have the inclusion $I(x) \subset \{n \in \mathbb{N} | x \in B(x_n, \frac{r_n^2}{2})\}$. We now claim the following.

Claim 5.1. *For every x such that $f'(x) = 0$,*

$$d(\mathrm{bd}\, M, x) \leq \max\{\varepsilon_n, n \in I(x)\},$$

and, for every $n \in I(x)$

$$|\lambda_n'(x)| \leq \frac{1}{3\|\nabla\alpha_n\|_\infty \mathrm{card}\, I(x)} \min\left\{\frac{\delta(y_q)}{3}, y_q \in \overline{B}\left(x, \frac{2}{3}\right)\right\} \min\{\varepsilon(x), 1\}.$$

Proof of Claim 5.1. Since $f'(x) = \sum_{k \in \mathbb{N}} \alpha_k(x)\lambda_k'(x) = 0$, there exist p and q in $I(x)$ such that $\lambda_p'(x) \leq 0$ and $\lambda_q'(x) \geq 0$, hence $\lambda_p(x) \leq \varepsilon_p$ and $\lambda_q(x) \geq -\varepsilon_q$. Since $\varepsilon_p \leq \frac{r_p^2}{2}$ and $\varepsilon_q \leq \frac{r_q^2}{2}$, $x + \varepsilon_p e_p \in B(x_p, r_p^2)$ and $\lambda_p(x + \varepsilon_p e_p) = \lambda_p(x) - \varepsilon_p \leq 0$, $x - \varepsilon_q e_q \in (x_q, r_q^2)$ and $\lambda_q(x - \varepsilon_q e_q) = \lambda_q(x) + \varepsilon_q \geq 0$. In view of (5.7) and (5.8), $x + \varepsilon_p e_p \in M$ and $x - \varepsilon_q e_q \notin \mathrm{int}\, M$. Hence there exists $\theta \in [0,1]$ such that $x + \theta\varepsilon_p e_p - (1-\theta)\varepsilon_q e_q \in \mathrm{bd}\, M$. Then

$$d(\mathrm{bd}\, M, x) \leq |\theta\varepsilon_p e_p - (1-\theta)\varepsilon_q e_q| \leq \theta\varepsilon_p + (1-\theta)\varepsilon_q \leq \max\{\varepsilon_n, n \in I(x)\}.$$

Take $n \in I(x)$ and $k_0 \in I(x)$ such that $\varepsilon_{k_0} = \max\{\varepsilon_k, k \in I(x)\}$. Then $x \in B(x_n, \frac{r_n^2}{2}) \cap B(x_{k_0}, \frac{r_{k_0}^2}{2})$, hence $|x_n - x_{k_0}| < \frac{r_n^2}{2} + \frac{r_{k_0}^2}{2} \leq 1$, i.e., $x_n \in \overline{B}(x_{k_0}, 1)$. From (5.12), we have $\varepsilon_{k_0} \leq \frac{r_n^2}{2}$. Then

$$|\theta\varepsilon_p e_p - (1-\theta)\varepsilon_q e_q| \leq \max\{\varepsilon_k, k \in I(x)\} \leq \varepsilon_{k_0} \leq \frac{r_n^2}{2},$$

hence $x + \theta\varepsilon_p e_p - (1-\theta)\varepsilon_q e_q \in B(x_n, r_n^2)$ and $\lambda_n(x + \theta\varepsilon_p e_p - (1-\theta)\varepsilon_q e_q) = 0$. Since the function λ_n is $\frac{1}{r_n}$-Lipschitz, $|\lambda_n(x)| \leq \frac{\varepsilon_{k_0}}{r_n}$ and $|\lambda_n'(x)| \leq \frac{\varepsilon_{k_0}}{r_n} + \varepsilon_n$. Since $r_n \leq 1$ and $\varepsilon_n \leq \varepsilon_{k_0}$, we have $|\lambda_n'(x)| \leq \frac{2\varepsilon_{k_0}}{r_n}$. On the other side, from (5.12) and since $x_n \in \overline{B}(x_{k_0}, 1)$,

$$\varepsilon_{k_0} \leq \frac{\varepsilon'(x_{k_0})}{\mathrm{card}\,\{k | x_k \in \overline{B}(x_{k_0}, 1)\}} \frac{r_n}{2\|\nabla\alpha_n\|_\infty} \min\left\{\frac{\delta(y_p)}{3}, y_p \in \overline{B}(x_{k_0}, 1)\right\}.$$

Noting that $k \in I(x) \Rightarrow x_k \in \overline{B}(x_{k_0}, 1)$, we deduce $I(x) \subset \{k | x_k \in \overline{B}(x_{k_0}, 1)\}$ and the corresponding inequality on the cardinals. Also, we remark that $y_p \in \overline{B}(x, \frac{2}{3}) \Rightarrow y_p \in \overline{B}(x_{k_0}, 1)$ and $\varepsilon'(x_{k_0}) \leq \frac{1}{3}\min\{\varepsilon(x), 1\}$. We deduce

$$\varepsilon_{k_0} \leq \frac{\min\{\varepsilon(x), 1\}}{3\mathrm{card}\, I(x)} \frac{r_n}{2\|\nabla\alpha_n\|_\infty} \min\left\{\frac{\delta(y_p)}{3}, y_p \in \overline{B}(x, \frac{2}{3})\right\},$$

which proves Claim 5.1. $\square$

We now come back to the smoothness of M'. Consider $x \in \Omega$ such that $f'(x) = 0$. Noting that

$$\nabla f'(x) = \sum_{k \in \mathbb{N}} \alpha_k(x) \nabla \lambda'_k(x) + \sum_{k \in \mathbb{N}} \nabla \alpha_k(x) \lambda'_k(x),$$

we estimate separately each term. For every $n \in I(x)$, $\nabla \lambda'_n(x) \in \partial \lambda_n(B(x, \varepsilon_n)) + B(0, \varepsilon_n)$, hence

$$\sum_{k \in \mathbb{N}} \alpha_k(x) \nabla \lambda'_k(x) \in \sum_{k \in \mathbb{N}} \alpha_k(x) \partial \lambda_k(B(x, \varepsilon_k)) + \sum_{k \in \mathbb{N}} \alpha_k(x) B(0, \varepsilon_k).$$

From Claim 5.1, $d(\mathrm{bd}\, M, x) \leq \max\{\varepsilon_n, n \in I(x)\}$. Take $n \in I(x)$ such that ε_n realizes the maximum. Then there is $y \in \mathrm{bd}\, M$ such that $|x - y| \leq \varepsilon_n$. Since $\mathrm{bd}\, M \subset \bigcup_{p \in \mathbb{N}} B(y_p, \frac{\rho_p}{3})$, there exists $p \in \mathbb{N}$ such that $y \in B(y_p, \frac{\rho_p}{3})$. Then $|x_n - y_p| \leq |x_n - x| + |x - y| + |y - y_p| \leq 1$. Hence $y_p \in \overline{B}(x_n, 1)$ and, in view of (5.12), $\varepsilon_n \leq \frac{\rho_p}{3} \varepsilon'(x_n)$. In particular, $|x - y| \leq \frac{\rho_p}{3}$, hence $x \in B(y_p, \frac{2\rho_p}{3})$. Consider $k \in I(x)$, then $\varepsilon_k \leq \varepsilon_n$ and $B(x, \varepsilon_k) \subset B(y_p, \rho_p)$. From the definition of ρ_p, we have

$$\sum_{k \in \mathbb{N}} \alpha_k(x) \partial \lambda_k(B(x, \varepsilon_k)) \subset \sum_{k \in \mathbb{N}} \alpha_k(B(y_p, \rho_p)) \partial \lambda_k(B(y_p, \rho_p))$$
$$\subset \sum_{k \in \mathbb{N}} \alpha_k(y_p) \partial \lambda_k(y_p) + B\left(0, \frac{\delta(y_p)}{3} \varepsilon'(y_p)\right),$$

hence

$$\sum_{k \in \mathbb{N}} \alpha_k(x) \nabla \lambda'_k(x) \in \sum_{k \in I(x)} \alpha_k(y_p) \partial \lambda_k(y_p) + B\left(0, \sum_{k \in I(x)} \alpha_k(x) \varepsilon_k + \frac{\delta(y_p)}{3} \varepsilon'(y_p)\right).$$

But $\sum_{k \in I(x)} \alpha_k(x) \varepsilon_k \leq \max\{\varepsilon_k, k \in I(x)\} \leq \frac{\delta(y_p)}{3} \varepsilon'(x_n)$. Since $x \in \overline{B}(x_n, 1)$ and $x \in \overline{B}(y_p, 1)$, then $\varepsilon'(x_n) \leq \frac{1}{3} \min\{\varepsilon(x), 1\}$ and $\varepsilon'(y_p) \leq \frac{1}{3} \min\{\varepsilon(x), 1\}$. Hence

$$(5.19) \quad \sum_{k \in \mathbb{N}} \alpha_k(x) \nabla \lambda'_k(x) \in \sum_{k \in I(x)} \alpha_k(y_p) \partial \lambda_k(y_p) + B\left(0, \frac{2\delta(y_p)}{9} \min\{\varepsilon(x), 1\}\right).$$

Let us now estimate the other term. From Claim 5.1, for every $k \in I(x)$, $\lambda'_k(x) \leq \frac{1}{3\|\nabla \alpha_k\|_\infty \mathrm{card}\, I(x)} \min\left\{\frac{\delta(y_q)}{3}, y_q \in \overline{B}(x, \frac{2}{3})\right\} \min\{\varepsilon(x), 1\}$. But $|x - y_p| \leq \frac{2\rho_p}{3} \leq \frac{2}{3}$, hence $y_p \in \overline{B}(x, \frac{2}{3})$. Hence $\lambda'_k(x) \leq \frac{1}{3\|\nabla \alpha_k\|_\infty \mathrm{card}\, I(x)} \frac{\delta(y_p)}{3} \min\{\varepsilon(x), 1\}$ and

$$(5.20) \quad \left|\sum_{k \in \mathbb{N}} \nabla \alpha_k(x) \lambda'_k(x)\right| = \left|\sum_{k \in I(x)} \nabla \alpha_k(x) \lambda'_k(x)\right| \leq \frac{\delta(y_p)}{9} \min\{\varepsilon(x), 1\}.$$

In view of (5.19) and (5.20), we deduce

$$(5.21) \quad \nabla f'(x) \in \sum_{k \in I(x)} \alpha_k(y_p) \partial \lambda_k(y_p) + B\left(0, \frac{\delta(y_p)}{3} \min\{\varepsilon(x), 1\}\right).$$

In particular, from the definition of the function δ, we deduce that

$$|\nabla f'(x)| \geq \frac{2\delta(y_p)}{3} > 0.$$

5.2.2. *Proof of* (2.9) *and* (2.10). Take $x \in M'$. Then $f'(x) \leq 0$, hence there exists $n \in I(x)$ such that $\lambda'_n(x) \leq 0$, $\lambda_n(x) \leq \varepsilon_n$, $\lambda_n(x + \varepsilon_n e_n) \leq 0$, $x + \varepsilon_n e_n \in M$ and $d(M, x) \leq \varepsilon_n$ Since $\varepsilon_n \leq \varepsilon'(x_n) \leq \varepsilon(x)$, we deduce (2.9). Now take $x \in M$. Then there exists $n \in \mathbb{N}$ such that $x \in B(x_n, r_n)$. Then $\lambda_n(x_n) \leq 0$, $\lambda_n(x_n + \varepsilon_n e_n) \leq -\varepsilon_n$ and $\lambda'_n(x_n + \varepsilon_n e_n) \leq 0$. But $x_n + \varepsilon_n e_n \in B(x_n, r'_n)$, hence $I(x_n + \varepsilon_n e_n) = \{n\}$ and $f'(x_n + \varepsilon_n e_n) = \lambda'_n(x_n + \varepsilon_n e_n) \leq 0$, hence $x_n + \varepsilon_n e_n \in M'$ and $d(M', x) \leq \varepsilon_n + r_n$. Since $\varepsilon_n \leq r_n$ and $2r_n \leq \varepsilon'(x_n) \leq \varepsilon(x)$, we deduce (2.10). □

5.2.3. *Proof of* (2.11). Take $x \in \operatorname{bd} M'$ and $v \in N^C_{M'}(x) \cap S$. Then $v = \frac{\nabla f'(x)}{|\nabla f'(x)|}$. In view of (5.21), recall that there exists $p \in \mathbb{N}$ such that $x \in B(y_p, \frac{2\rho_p}{3})$ and $\nabla f'(x) \in \sum_{k \in I(x)} \alpha_k(y_p) \partial \lambda_k(y_p) + B\left(0, \frac{\delta(y_p)}{3} \min\{\varepsilon(x), 1\}\right)$. Since $\sum_{k \in I(x)} \alpha_k(x) \partial \lambda_k(y_p) \subset N^C_M(y_p) \setminus \{0\}$, we have

$$\frac{\nabla f'(x)}{|\nabla f'(x)|} \in N^C_M(y_p) + B\left(0, \frac{\delta(y_p)}{3|\nabla f'(x)|} \min\{\varepsilon(x), 1\}\right).$$

From above $|\nabla f'(x)| \geq \frac{2\delta(y_p)}{3}$, hence $\frac{\nabla f'(x)}{|\nabla f'(x)|} \in N^C_M(y_p) + B(0, \frac{1}{2} \min\{\varepsilon(x), 1\}$. Take $v \in N^C_M(y_p)$ such that

$$\left| \frac{\nabla f'(x)}{|\nabla f'(x)|} - v \right| \leq \frac{1}{2} \min\{\varepsilon(x), 1\}.$$

Then $v \neq 0$ and

$$
\begin{aligned}
\left| \frac{\nabla f'(x)}{|\nabla f'(x)|} - \frac{v}{|v|} \right| &= \frac{1}{|v|} \left| |v| \frac{\nabla f'(x)}{|\nabla f'(x)|} - |v|v + |v|v - v \right| \\
&\leq \left| \frac{\nabla f'(x)}{|\nabla f'(x)|} - v \right| + ||v| - 1| \\
&\leq 2 \left| \frac{\nabla f'(x)}{|\nabla f'(x)|} - v \right| \leq \varepsilon(x).
\end{aligned}
$$

On the other hand, $|x - y_p| \leq \rho_p \leq \varepsilon(x)$, and we deduce (2.11). □

5.2.4. *Proof of* (2.12). Take $x \in \operatorname{bd} M$ and $v \in N^C_M(x) \cap S$. Then there exists $n \in \mathbb{N}$ such that $x \in B(x_n, r_n)$ and $v \in N^C_M(x_n) \cap S + B(0, \varepsilon'(x_n))$. In view of (5.10), $N^C_M(x_n) = \bigcup_{\lambda > 0} \lambda \partial \lambda_n(x_n)$, hence there exists $\zeta \in \partial \lambda_n(x_n)$ such that

$$\left| v - \frac{\zeta}{|\zeta|} \right| < \varepsilon'(x_n).$$

From (5.15), $\partial \lambda_n(x_n) \subset \nabla \lambda'_n(B(x_n, \varepsilon_n)) + B(0, \varepsilon_n)$, hence there exists $x' \in B(x_n, \varepsilon_n)$ such that

$$|\zeta - \nabla \lambda'_n(x')| < \varepsilon_n.$$

Since $N^C_M(x_n) \neq \{0\}$, $x_n \in \operatorname{bd} M$ and $\lambda_n(x_n) = 0$. The function λ_n being $\frac{1}{r_n}$-Lipschitz, $|\lambda_n(x')| \leq \frac{\varepsilon_n}{r_n}$, hence $|\lambda'_n(x')| \leq \frac{\varepsilon_n}{r_n} + \varepsilon_n$. Then the two points $x' + \varepsilon_n(1 + \frac{1}{r_n})e_n$ and $x' - \varepsilon_n(1 + \frac{1}{r_n})e_n$ belong to the ball $B(x_n, \varepsilon_n(2 + \frac{1}{r_n}))$. In view of (5.12), $\varepsilon_n(2 + \frac{1}{r_n})) \leq r'_n$, and the two points belong to the ball $B(x_n, r'_n)$, which is contained in $B(x_n, r_n^2)$. Then $\lambda'_n(x' + \varepsilon_n(1 + \frac{1}{r_n})e_n) \leq 0$ and $\lambda'_n(x' - \varepsilon_n(1 + \frac{1}{r_n})e_n) \geq 0$. Hence there exists $\theta \in [-\varepsilon_n(1 + \frac{1}{r_n}), \varepsilon_n(1 + \frac{1}{r_n})]$ such that $\lambda'_n(x' + \theta e_n) = 0$. Also $x' + \theta e_n \in B(x_n, r'_n)$ and $I(x' + \theta e_n) = \{n\}$. We thus deduce that $f'(x' + \theta e_n) =$

$\lambda'_n(x' + \theta e_n) = 0$, $x' + \theta e_n \in \mathrm{bd}\, M$ and $\nabla f'(x' + \theta e_n) = \nabla \lambda'_n(x' + \theta e_n) = \nabla \lambda'_n(x')$ in view of (5.17). Then $v' := \frac{\nabla f'(x' + \theta e_n)}{|\nabla f'(x' + \theta e_n)|} \in N^C_{M'}(x' + \theta e_n) \cap S$. On one hand,

$$|x' + \theta e_n - x| \leq \varepsilon_n + |\theta| \leq r'_n \leq r_n \leq \varepsilon(x).$$

On the other hand,

$$(5.22) \qquad |v' - v| \;=\; \left| v' - \frac{\zeta}{|\zeta|} + \frac{\zeta}{|\zeta|} - v \right|$$

$$(5.23) \qquad \leq \; \left| \frac{\nabla \lambda'_n(x')}{|\nabla \lambda'_n(x')|} - \frac{\zeta}{|\zeta|} \right| + \varepsilon'(x_n).$$

But

$$\left| \frac{\nabla \lambda'_n(x')}{|\nabla \lambda'_n(x')|} - \frac{\zeta}{|\zeta|} \right| \;=\; \frac{1}{|\nabla \lambda'_n(x')||\zeta|} \Big| |\zeta| \nabla \lambda'_n(x') - |\nabla \lambda'_n(x')| \nabla \lambda'_n(x')$$

$$+ |\nabla \lambda'_n(x')| \nabla \lambda'_n(x') - |\nabla \lambda'_n(x')| \zeta \Big|$$

$$\leq \; \frac{1}{|\zeta|} \Big| |\zeta| - |\nabla \lambda'_n(x')| \Big| + \frac{1}{|\zeta|} |\nabla \lambda'_n(x') - \zeta|$$

$$\leq \; 2|\zeta - \nabla \lambda'_n(x')| < 2\varepsilon_n.$$

Hence $|v' - v| \leq 2\varepsilon_n + \varepsilon'(x_n) \leq 3\varepsilon'(x_n) \leq \varepsilon(x)$, and we deduce (2.12). $\qquad \square$

5.3. **Proof of Proposition 2.2.** Consider $x \in M$ and $v \in N^P_M(x) \cap S^{N-1}$. Since $r \leq \mathrm{reach}\,(M)$, in view of [17, Theorem 4.8-(12)],

$$x \in \mathrm{proj}\,_M(x + rv).$$

Take $x' \in \mathrm{proj}\,_{M'}(x + \frac{r}{2} v)$. Recalling that $\varepsilon := d_{Haus}(M, M')$, from Lemma 4.1

$$|x - x'|^2 \leq \varepsilon^2 + 4\varepsilon r.$$

Let $v' = \frac{2}{r}(x + \frac{r}{2} v - x')$ (assuming that $r \neq 0$). From the definition of x', we have $v' \in N^P_{M'}(x')$. Then

$$|v - v'| = \frac{2}{r}|x - x'|.$$

On the other hand, note that $d(x + \frac{r}{2} v, M) = \frac{r}{2}$. Since $d_{Haus}(M, M') < \frac{r}{2}$, then $x + \frac{r}{2} v \notin M'$ and $x' \neq x + \frac{r}{2} v$, which implies $v' \neq 0$. Then $\frac{v'}{|v'|} \in N^P_{M'}(x') \cap S$ and

$$\left| v - \frac{v'}{|v'|} \right| \;\leq\; |v - v'| + \left| v' - \frac{v'}{|v'|} \right|$$

$$\leq \; |v - v'| + |v'| \left| 1 - \frac{1}{|v'|} \right|$$

$$\leq \; |v - v'| + \big| |v'| - 1 \big|$$

$$\leq \; |v - v'| + \big| |v'| - |v| \big|$$

$$\leq \; 2|v - v'| \leq 4 \frac{\sqrt{\varepsilon^2 + 4\varepsilon r}}{r}.$$

Hence $(x', \frac{v'}{|v'|}) \in \mathrm{graph}\, N^P_{M'} \cap S$ and

$$d_\infty \left((x, v), \left(x', \frac{v'}{|v'|} \right) \right) \leq \frac{\sqrt{\varepsilon^2 + 4\varepsilon r}}{\min\{r/4, 1\}}.$$
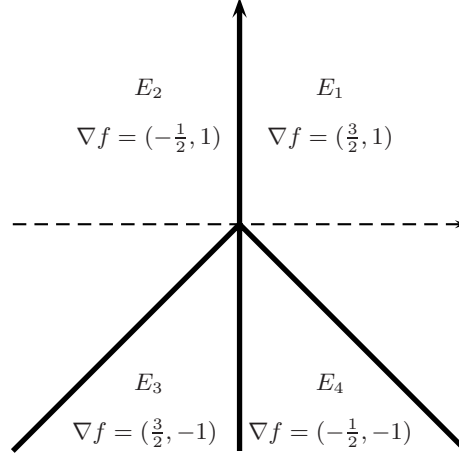
$\qquad \square$

6. Proof of Theorem 2.1

Clearly

$$
\begin{array}{rclcllcl}
f(x_1, x_2) &=& \tfrac{3}{2}x_1 &+& x_2 & \text{on} & E_1 &:=& \{(x_1, x_2) \in \mathbb{R}^2 | x_1 \geq 0, x_2 \geq -x_1\}, \\
f(x_1, x_2) &=& -\tfrac{1}{2}x_1 &+& x_2 & \text{on} & E_2 &:=& \{(x_1, x_2) \in \mathbb{R}^2 | x_1 \leq 0, x_2 \geq x_1\}, \\
f(x_1, x_2) &=& \tfrac{3}{2}x_1 &-& x_2 & \text{on} & E_3 &:=& \{(x_1, x_2) \in \mathbb{R}^2 | x_1 \leq 0, x_2 \leq x_1\}, \\
f(x_1, x_2) &=& -\tfrac{1}{2}x_1 &-& x_2 & \text{on} & E_4 &:=& \{(x_1, x_2) \in \mathbb{R}^2 | x_1 \geq 0, x_2 \leq -x_1\}.
\end{array}
$$

Since the gradient of f equals $(\tfrac{3}{2}, 1)$ on $\operatorname{int} E_1$, $(-\tfrac{1}{2}, 1)$ on $\operatorname{int} E_2$, $(\tfrac{3}{2}, -1)$ on $\operatorname{int} E_3$, and $(-\tfrac{1}{2}, -1)$ on $\operatorname{int} E_4$, we deduce that f is $\tfrac{\sqrt{13}}{2}$-Lipschitz and that

$$
0 \in \partial f(0) = \operatorname{co} \left\{ \left(\tfrac{3}{2}, 1 \right), \left(-\tfrac{1}{2}, 1 \right), \left(-\tfrac{1}{2}, -1 \right), \left(\tfrac{3}{2}, -1 \right) \right\}
$$

The proofs of part (a) and part (b) are independent and cannot be deduced one from the other. They are given in next two sections. In the third section, we prove part (c).

6.1. **Proof of part (a).** It is briefly outlined in Warga [37, Examples 3.3 and 2.4] and extensively developed here for the sake of completeness, moreover with a precise bound. For $\alpha > 0$ small enough, define two squares C and C_α as follows:

$$
\begin{array}{rcl}
C &:=& \operatorname{co} \{(1, 0), (0, 1), (-1, 0), (0, -1)\}, \\
C_\alpha &:=& \operatorname{co} \{(1 + \alpha, 0), (0, 1 + \alpha), (-1 - \alpha, 0), (0, -1 - \alpha)\}.
\end{array}
$$

Let $\theta : \mathbb{R}^2 \to \mathbb{R}$ be a C^∞ function satisfying the four following conditions:

$$
(6.1) \qquad \forall x \in \mathbb{R}^2, \ \theta(x) \in [0, \tfrac{1 - \alpha}{2}],
$$

$$
(6.2) \qquad \forall x \in C, \ \theta(x) = \tfrac{1 - \alpha}{2},
$$

$$
(6.3) \qquad \operatorname{supp} \theta \subset C_\alpha,
$$

$$
(6.4) \qquad \int_{\mathbb{R}^2} \theta(t) dt = 1.
$$

For every $\lambda > 0$, we define the function $f_\lambda : \mathbb{R}^2 \to \mathbb{R}$ by

$$f_\lambda(x) := \int_{\mathbb{R}^2} \theta(t) f(x + \lambda t) dt, \text{ for every } x \in \mathbb{R}^2.$$

It is C^∞, since it is a convolution of a C^∞ function and a Lipschitz function. On the other hand, since f is $\frac{\sqrt{13}}{2}$–Lipschitz, we have for every $x \in \mathbb{R}^2$,

$$|f(x) - f_\lambda(x)| \leq \int_{\mathbb{R}^2} \theta(t) \, |f(x + \lambda t) - f(x)| dt \leq \lambda \int_{\mathbb{R}^2} \theta(t) \frac{\sqrt{13}}{2} \, |t| dt,$$

which proves the uniform convergence of f_λ to f as λ tends to zero, i.e., assertion (2.2).

We now prove (2.3). First note that, since the function f is differentiable almost everywhere (by Rademacher's Theorem), we have for every $x \in \mathbb{R}^2$,

$$\nabla f_\lambda(x) = \int_{\mathbb{R}^2} \theta(t) \nabla f(x + \lambda t) dt.$$

Hence $|\nabla f_\lambda(x)| \leq \frac{\sqrt{13}}{2}$ and moreover

$$\nabla f_\lambda(x) = \int_{C_\alpha} \theta(t) \nabla f(x + \lambda t) dt + \int_{C_\alpha \setminus C} \theta(t) \nabla f(x + \lambda t) dt.$$

Then

$$\left| \int_{C_\alpha \setminus C} \theta(t) \nabla f(x + \lambda t) dt \right| \leq \text{meas} \, (C_\alpha \setminus C) \, \|\nabla f\|_\infty = 2(2\alpha + \alpha^2) \frac{\sqrt{13}}{2}.$$

Since $\theta(t) = \frac{1-\alpha}{2}$ for every $t \in C$,

$$\int_C \theta(t) \nabla f(x + \lambda t) dt$$

$$= \frac{1-\alpha}{2} \int_C \nabla f(x + \lambda t) dt$$

$$= \frac{1-\alpha}{2\lambda^{-2}} \int_{x+\lambda C} \nabla f(t) dt$$

$$= \frac{1-\alpha}{2\lambda^{-2}} \Big| \text{meas} \, (E_1 \cap x + \lambda C)(\frac{3}{2}, 1) + \text{meas} \, (E_2 \cap x + \lambda C)(-\frac{1}{2}, 1)$$

$$+ \text{meas} \, (E_3 \cap x + \lambda C)(\frac{3}{2}, -1) + \text{meas} \, (E_4 \cap x + \lambda C)(-\frac{1}{2}, -1) \Big|.$$

In view of the following lemma, if we choose $\alpha > 0$ small enough, we obtain

$$|\nabla f_\lambda(x)| \geq \frac{1}{5}.$$

Lemma 6.1. *For every $\lambda > 0$ and every $x \in \mathbb{R}^2$, we have*

$$\lambda^{-2} \Big| \text{meas} \, (E_1 \cap x + \lambda C) \left(\frac{3}{2}, 1 \right) + \text{meas} \, (E_2 \cap x + \lambda C) \left(-\frac{1}{2}, 1 \right)$$

$$+ \text{meas} \, (E_3 \cap x + \lambda C) \left(\frac{3}{2}, -1 \right) + \text{meas} \, (E_4 \cap x + \lambda C) \left(-\frac{1}{2}, -1 \right) \Big|$$

$$\geq \left(-1 + (\sqrt{2} + \sqrt{3})^{\frac{1}{3}} (3\sqrt{3} - 9\sqrt{2}) + (\sqrt{2} + \sqrt{3})^{\frac{2}{3}} (-9 + 6\sqrt{2}\sqrt{3}) \right)^{\frac{1}{2}}$$

with possible equality in the above equation.

Proof of Lemma 6.1. The proof of Lemma 6.1 is elementary and is achieved by considering 8 different cases (not necessarily disjoints). Note that one can also study the corresponding optimization problem, but the proof would not be shorter. We set $x := (x_1, x_2)$, $\mathbf{E_1} := \operatorname{meas}(E_1 \cap x + \lambda C)$, $\mathbf{E_2} := \operatorname{meas}(E_2 \cap x + \lambda C)$, $\mathbf{E_3} := \operatorname{meas}(E_3 \cap x + \lambda C)$, $\mathbf{E_4} := \operatorname{meas}(E_4 \cap x + \lambda C)$, and

$$
\begin{aligned}
g_\lambda(x) \;\; &:= \;\; \lambda^{-2} \left| \mathbf{E_1}\left(\frac{3}{2}, 1\right) + \mathbf{E_2}\left(-\frac{1}{2}, 1\right) + \mathbf{E_3}\left(\frac{3}{2}, -1\right) + \mathbf{E_4}\left(-\frac{1}{2}, -1\right) \right| \\
&= \;\; \lambda^{-2} \left(\left(\frac{3}{2}\mathbf{E_1} - \frac{1}{2}\mathbf{E_2} + \frac{3}{2}\mathbf{E_3} - \frac{1}{2}\mathbf{E_4}\right)^2 + (\mathbf{E_1} + \mathbf{E_2} - \mathbf{E_3} - \mathbf{E_4})^2 \right)^{1/2}.
\end{aligned}
$$

Note that $\mathbf{E_1} + \mathbf{E_2} + \mathbf{E_3} + \mathbf{E_4} = 2\lambda^2$, hence we have

$$
g_\lambda(x) = \lambda^{-2} \left((2(\mathbf{E_1} + \mathbf{E_3}) - \lambda^2)^2 + (2(\mathbf{E_1} + \mathbf{E_2}) - 2\lambda^2)^2 \right)^{1/2}.
$$

Without any loss of generality, noting that $g_\lambda(\lambda x) = g_1(x)$, we now assume that $\lambda = 1$. We now distinguish 8 (not necessarily disjoint) cases in order to give a minoration of $g_1(x)$:

Case 1: $1 \leq x_1$. Then $\mathbf{E_2} = \mathbf{E_3} = 0$, $\mathbf{E_1} + \mathbf{E_4} = 2$, hence

$$
g_1(x) = \left(8\mathbf{E_1^2} - 12\mathbf{E_1} + 5 \right)^{1/2} \geq \frac{1}{\sqrt{2}},
$$

with possible equality in the above equation.

Case 2: $x_1 \leq -1$. Then $\mathbf{E_1} = \mathbf{E_4} = 0$, $\mathbf{E_2} + \mathbf{E_3} = 2$, hence

$$
g_1(x) = \left(8\mathbf{E_2^2} - 20\mathbf{E_2} + 13 \right)^{1/2} \geq \frac{1}{\sqrt{2}}.
$$

Case 3: $1 + x_1 \leq x_2$ *or* $1 - x_1 \leq x_2$. Then $\mathbf{E_3} = \mathbf{E_4} = 0$, $\mathbf{E_1} + \mathbf{E_2} = 2$, hence

$$
g_1(x) = \left((2\mathbf{E_1} - 1)^2 + 4 \right)^{1/2} \geq 2.
$$

Case 4: $x_2 \leq -1 + x_1$ *and* $x_2 \leq -1 - x_1$. Then $\mathbf{E_1} = \mathbf{E_2} = 0$, $\mathbf{E_3} + \mathbf{E_4} = 2$, hence

$$
g_1(x) = \left((2\mathbf{E_3} - 1)^2 + 4 \right)^{1/2} \geq 2.
$$

Case 5: $x_2 \leq -1 + x_1$ *and* $-1 - x_1 \leq x_2$. Then $\mathbf{E_2} = 0$, $\mathbf{E_1} \leq 2$, hence

$$
\begin{aligned}
g_1(x) \;\; &= \;\; \left(8\mathbf{E_1^2} + 4\mathbf{E_3^2} + 8\mathbf{E_1}\mathbf{E_3} - 12\mathbf{E_1} - 4\mathbf{E_3} + 5 \right)^{1/2} \\
&= \;\; \left((2\mathbf{E_1} + 2\mathbf{E_3} - 1)^2 + 4\mathbf{E_1^2} - 8\mathbf{E_1} + 4 \right)^{1/2} \geq 2.
\end{aligned}
$$

Case 6: $-1 + x_1 \leq x_2$ *and* $x_2 \leq -1 - x_1$. Then $\mathbf{E_1} = 0$,

$$
g_1(x) = \left((2\mathbf{E_3} - 1)^2 + (2\mathbf{E_2} - 2)^2 \right)^{1/2}.
$$

If $x_2 \leq x_1$, then $\mathbf{E_3} \geq 3/4$ and $g_1(x) \geq 1/2$. If $x_1 + 1 \leq x_2$, then Case 3 applies and $g_1(x) \geq 2$. If $x_1 + 1 \geq x_2 \geq x_1$, then one easily shows that

$$
g_1(x) \geq g_1\left(\left(x_1 + \frac{-x_1 - x_2 - 1}{2}, x_2 + \frac{-x_1 - x_2 - 1}{2} \right) \right),
$$

with possible equality. Hence, without any loss of generality, we assume $-x_1 - x_2 - 1 = 0$. Then $\mathbf{E_2} = 1 + x_2 - x_1$ and $\mathbf{E_3} = \frac{1}{4}(3 + x_2 - x_1)(1 - (x_2 - x_1))$ and

$$g_1(x) = \frac{1}{2}((x_2 - x_1)^4 + 4(x_2 - x_1)^3 + 18(x_2 - x_1)^2 - 4(x_2 - x_1) + 1)^{1/2}.$$

One easily deduces $g_1(x) \geq \frac{1}{2}(18(x_2 - x_1)^2 - 4(x_2 - x_1) + 1)^{1/2} \geq \frac{\sqrt{7}}{6}$. More accurately, the polynomial $P(X) = X^4 + 4X^3 + 18X^2 - 4X + 1$ attains its minimum over the interval $[0, 1]$ at the point

$$-1 + (4 + 2\sqrt{6})^{\frac{1}{3}} + (4 + 2\sqrt{6})^{\frac{2}{3}}\left(1 - \frac{\sqrt{6}}{2}\right).$$

Indeed, it is the unique real root of P', and moreover $P'(0) = -4 < 0$ and $P'(1) = 48 > 0$, hence it is the unique candidate for the minimization of P over $[0, 1]$. The corresponding minimum of g_1 is

$$\left(-1 + (\sqrt{2} + \sqrt{3})^{\frac{1}{3}}(3\sqrt{3} - 9\sqrt{2}) + (\sqrt{2} + \sqrt{3})^{\frac{2}{3}}(-9 + 6\sqrt{2}\sqrt{3})\right)^{\frac{1}{2}}$$

and is attained for the corresponding values of x_1 and x_2.

Case 7: $0 \leq x_1$, $-1 + x_1 \leq x_2$, $x_2 \leq 1 - x_1$. Set $a := d(\mathbb{R}(1, 1), x + (0, -1)) = \frac{x_1 - x_2 + 1}{\sqrt{2}}$ and $b := d(\mathbb{R}(1, -1), x + (0, -1)) = \frac{-x_1 - x_2 + 1}{\sqrt{2}}$. Then

$$0 \leq b \leq a \leq \sqrt{2},$$
$$\mathbf{E_1} = 2 - \sqrt{2}b - (\sqrt{2} - a)^2/2,$$
$$\mathbf{E_1} + \mathbf{E_2} = 2 - ab,$$
$$\mathbf{E_3} = b^2/2.$$

Then $\mathbf{E_1} + \mathbf{E_3} \geq 2 - \sqrt{2}b - (\sqrt{2} - b)^2/2 + b^2/2 = 1$ and $g_1(x) \geq 1$.

Case 8: $x_1 \leq 0$, $x_2 \leq 1 + x_1$, $-1 - x_1 \leq x_2$. Set $a := d(\mathbb{R}(1, 1), x + (0, -1)) = \frac{x_1 - x_2 + 1}{\sqrt{2}}$ and $b := d(\mathbb{R}(1, -1), x + (0, -1)) = \frac{-x_1 - x_2 + 1}{\sqrt{2}}$. Then

$$0 \leq a \leq b \leq \sqrt{2},$$
$$\mathbf{E_1} = a(\sqrt{2} - b) + (\sqrt{2} - b)^2/2,$$
$$\mathbf{E_1} + \mathbf{E_2} = 2 - ab,$$
$$\mathbf{E_3} = ab - a^2/2.$$

Note that if $a \leq 2\sqrt{2} - \sqrt{5}$, then $\mathbf{E_1} + \mathbf{E_2} \geq 2 - \sqrt{2}a$ and $g_1(x) \geq 2\sqrt{10} - 6$. If $\sqrt{2} \geq a \geq 2\sqrt{2} - \sqrt{5}$, then $\mathbf{E_1} + \mathbf{E_3} \geq \sqrt{2}a - a^2/2 \geq \sqrt{10} - 5/2$ and $g_1(x) \geq 2\sqrt{10} - 6$. More accurately, $g_1(x) = P(a, b)$, where

$$P(a, b) = a^4 + 2a^2b^2 - 4\sqrt{2}a^3 + 4\sqrt{2}a^2b + 4\sqrt{2}ab^2 - 4\sqrt{2}b^3 + 6a^2 - 24ab + 10b^2$$
$$+ 4\sqrt{2}a - 4\sqrt{2}b + 5.$$

The solution of the minimization problem $\min\{P(a, b), 0 \leq a \leq b \leq \sqrt{2}\}$ is now classical. In order to solve the first order necessary condition, we compute the gradient of P:

$$\frac{dP}{da} = 4a^3 + 4ab^2 - 12\sqrt{2}a^2 + 8\sqrt{2}ab + 4\sqrt{2}b^2 + 12a - 24b + 4\sqrt{2},$$
$$\frac{dP}{db} = 4b^3 + 4a^2b - 12\sqrt{2}b^2 + 4\sqrt{2}a^2 + 8\sqrt{2}ab + 20b - 24a - 4\sqrt{2}.$$

In order to show that the gradient is nonzero on the domain, we compute the resultants in a and b,

$$\text{Resultant}\left(\frac{dP}{da}, \frac{dP}{db}, a\right) = 2^{-12}(36a^7 - 126\sqrt{2}a^6 + 203a^5 + 83\sqrt{2}a^4$$
$$- 287a^3 + 97\sqrt{2}a^2 - 11a - 3\sqrt{2}),$$

$$\text{Resultant}\left(\frac{dP}{da}, \frac{dP}{db}, b\right) = 2^{-12}(36b^7 - 138\sqrt{2}b^6 + 359b^5 - 117\sqrt{2}b^4$$
$$- 221b^3 + 179\sqrt{2}b^2 - 79b + 5\sqrt{2}).$$

A direct study shows that there is only one possible candidate (a, b) with $0 \leq a \leq b \leq \sqrt{2}$ for the equation

$$\begin{cases} \text{Resultant}\left(\frac{dP}{da}, \frac{dP}{db}, a\right) & = 0, \\ \text{Resultant}\left(\frac{dP}{da}, \frac{dP}{db}, b\right) & = 0. \end{cases}$$

Moreover, it satisfies $a \in [0.52,\ 0.58]$ and $b \in [0.72,\ 0.78]$. But

$$\frac{dP}{db}([0.52,\ 0.58] \times [0.72,\ 0.78]) \subset (-\infty, -1].$$

By a direct estimation of $P(a, b)$, or using the first order necessary condition, one easily shows that the minimum is not attained for $a = 0$ or $a = b$. For $b = \sqrt{2}$, $0 < a < \sqrt{2}$ the only real root of $\frac{dP}{da}(a, \sqrt{2})$ is $a = \sqrt{2} - (\sqrt{(2)} + \sqrt{(3)})^{(1/3)} + (\sqrt{(2)} + \sqrt{(3)})^{(2/3)}(\sqrt{(3)} - \sqrt{(2)})$. In fact we can observe that we obtain the same point as in Case 6, hence the same optimal value of g_1. This concludes the proof of Lemma 6.1. $\qquad\square$

6.2. **Proof of part (b).** It is in fact deduced from the exact computation of the inf sup convoluted function of f. In the next section, we compute the inf convoluted function f_λ of f, and in the following section, we compute the inf sup convoluted function $f_{\lambda,\mu}$.

6.2.1. *Calculus of f_λ, the inf convoluted function of f.* For $\lambda > 0$, we define the sets Λ_i on which we give an analytic formula of the inf convoluted function f_λ which is given in the following lemma:

$\Lambda_1 = \{(x_1, x_2) \in \mathbb{R}^2,\ x_1 - \frac{3\lambda}{2} \geq 0,\ x_1 + x_2 - \frac{5\lambda}{2} \geq 0\},$

$\Lambda_2 = \{(x_1, x_2) \in \mathbb{R}^2,\ x_1 + \frac{\lambda}{2} \leq 0,\ x_1 - x_2 + \frac{3\lambda}{2} \leq 0\},$

$\Lambda_3 = \{(x_1, x_2) \in \mathbb{R}^2,\ x_1 - \frac{\lambda}{2} \leq 0,\ -x_1 + x_2 + \frac{5\lambda}{2} \leq 0\},$

$\Lambda_4 = \{(x_1, x_2) \in \mathbb{R}^2,\ x_1 - \frac{\lambda}{2} \geq 0,\ x_1 + x_2 + \frac{3\lambda}{2} \leq 0\},$

$\Lambda_{1,2} = \{(x_1, x_2) \in \mathbb{R}^2,\ x_1 \in [-\frac{\lambda}{2}, \frac{3\lambda}{2}],\ x_2 \geq \lambda\},$

$\Lambda_{2,3} = \{(x_1, x_2) \in \mathbb{R}^2,\ x_1 + x_2 - \frac{\lambda}{2} \leq 0,\ x_1 - x_2 + \frac{3\lambda}{2} \geq 0,\ -x_1 + x_2 + \frac{5\lambda}{2} \geq 0,$
$\qquad\qquad x_1 - \frac{\lambda}{2} \leq 0\},$

$\Lambda_{4,1} = \{(x_1, x_2) \in \mathbb{R}^2,\ x_1 - x_2 - \frac{\lambda}{2} \geq 0,\ x_1 + x_2 - \frac{5\lambda}{2} \leq 0,\ x_1 + x_2 + \frac{3\lambda}{2} \geq 0,$
$\qquad\qquad x_1 - \frac{\lambda}{2} \geq 0\},$

$\Lambda_0 = \{(x_1, x_2) \in \mathbb{R}^2,\ x_1 + x_2 - \frac{\lambda}{2} \geq 0,\ x_1 - x_2 - \frac{\lambda}{2} \leq 0,\ x_2 \leq \lambda\}.$

Lemma 6.2. *Let $\lambda > 0$ and let $x \in \mathbb{R}^2$. Then*

$$
\begin{array}{rclll}
f_\lambda(x) & = & \frac{3}{2}x_1 \ + \ x_2 \ -\frac{13\lambda}{8} & \text{if} & x \in \Lambda_1, \\
f_\lambda(x) & = & -\frac{1}{2}x_1 \ + \ x_2 \ -\frac{5\lambda}{8} & \text{if} & x \in \Lambda_2, \\
f_\lambda(x) & = & \frac{3}{2}x_1 \ - \ x_2 \ -\frac{13\lambda}{8} & \text{if} & x \in \Lambda_3, \\
f_\lambda(x) & = & -\frac{1}{2}x_1 \ - \ x_2 \ -\frac{5\lambda}{8} & \text{if} & x \in \Lambda_4, \\
f_\lambda(x) & = & \frac{1}{2\lambda}x_1^2 \ + \ x_2 \ -\frac{\lambda}{2} & \text{if} & x \in \Lambda_{1,2}, \\
f_\lambda(x) & = & \frac{1}{4\lambda}x_1^2 \ + \ \frac{1}{4\lambda}x_2^2 \ -\frac{1}{2\lambda}x_1x_2 + \frac{1}{4}x_1 + \frac{1}{4}x_2 - \frac{\lambda}{16} & \text{if} & x \in \Lambda_{2,3}, \\
f_\lambda(x) & = & \frac{1}{4\lambda}x_1^2 \ + \ \frac{1}{4\lambda}x_2^2 \ +\frac{1}{2\lambda}x_1x_2 + \frac{1}{4}x_1 - \frac{1}{4}x_2 - \frac{\lambda}{16} & \text{if} & x \in \Lambda_{4,1}, \\
f_\lambda(x) & = & \frac{1}{2\lambda}x_1^2 \ + \ \frac{1}{2\lambda}x_2^2 & \text{if} & x \in \Lambda_0.
\end{array}
$$
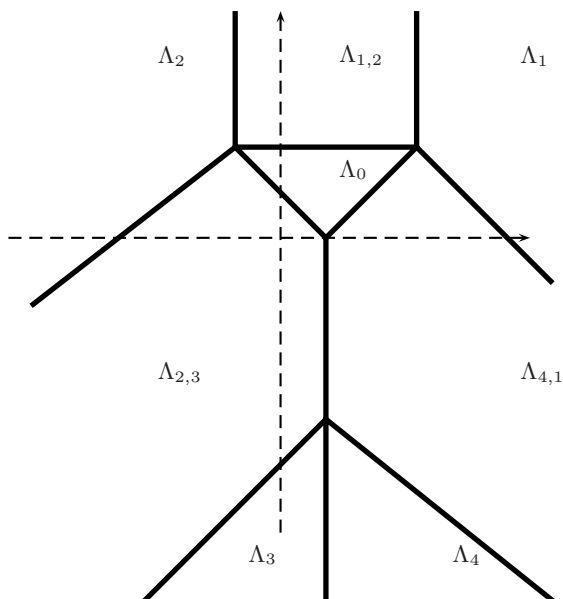
Proof of Lemma 6.2. We first compute $f_\lambda(x) = \inf\{f(z) + \frac{1}{2\lambda}|x - z|^2, z \in \mathbb{R}^2\}$. In fact the infimum in the preceding formula is attained at a point $z \in \mathbb{R}^2$ which satisfies the first order necessary condition

$$(6.5) \qquad 0 \in \partial f(z) + \frac{1}{\lambda}(z - x).$$

One could consider the stronger necessary condition $0 \in \partial_P f(z) + \frac{1}{\lambda}(z - x)$, but the proof with Clarke's generalized gradient is more basic and does not require the computation of $\partial_P f$. We distinguish the nine cases (i) $z \in \text{int } E_1$, (ii) $z \in \text{int } E_2$, (iii) $z \in \text{int } E_3$, (iv) $z \in \text{int } E_4$, (v) $z \in (E_1 \cap E_2) \setminus \{0\}$, (vi) $z \in (E_2 \cap E_3) \setminus \{0\}$, (vii) $z \in (E_3 \cap E_4) \setminus \{0\}$, (viii) $z \in (E_4 \cap E_1) \setminus \{0\}$, (ix) $z = 0$. Wet let $x = (x_1, x_2)$ and $z = (z_1, z_2)$.

Case (i), $z \in \text{int } E_1$. Then $\nabla f(z) = (\frac{3}{2}, 1)$ and $z_1 > 0$, $z_1 + z_2 > 0$. The first order necessary condition can be rewritten as

$$
\begin{cases}
0 & = & \frac{3}{2} + \frac{1}{\lambda}(z_1 - x_1), \\
0 & = & 1 + \frac{1}{\lambda}(z_2 - x_2).
\end{cases}
$$

This implies that $x_1 - \frac{3\lambda}{2} > 0$, $x_1 + x_2 - \frac{5\lambda}{2} > 0$ and that the candidate for $f_\lambda(x)$ is

$$f(z) + \frac{\lambda}{2}|x - z|^2 = \frac{3}{2}x_1 + x_2 - \frac{13\lambda}{8}.$$

Case (ii), $z \in \text{int } E_2$. Then $\nabla f(z) = (-\frac{1}{2}, 1)$ and $z_1 < 0$, $z_1 - z_2 < 0$, and the first order necessary condition implies that $x_1 + \frac{\lambda}{2} < 0$, $x_1 - x_2 + \frac{3\lambda}{2} < 0$ and that the candidate for $f_\lambda(x)$ is

$$f(z) + \frac{1}{2\lambda}|x - z|^2 = -\frac{1}{2}x_1 + x_2 - \frac{5\lambda}{8}.$$

Case (iii), $z \in \text{int } E_3$. Then $\nabla f(z) = (\frac{3}{2}, -1)$ and $z_1 < 0$, $-z_1 + z_2 < 0$, and the first order necessary condition implies that $x_1 - \frac{3\lambda}{2} < 0$, $-x_1 + x_2 + \frac{5\lambda}{2} < 0$ and that the candidate for $f_\lambda(x)$ is

$$f(z) + \frac{1}{2\lambda}|x - z|^2 = \frac{3}{2}x_1 - x_2 - \frac{13\lambda}{8}.$$

Case (iv), $z \in \text{int } E_4$. Then $\nabla f(z) = (-\frac{1}{2}, -1)$ and $z_1 > 0$, $z_1 + z_2 < 0$, and the first order necessary condition implies that $x_1 + \frac{\lambda}{2} > 0$, $x_1 + x_2 + \frac{3\lambda}{2} < 0$ and that the candidate for $f_\lambda(x)$ is

$$f(z) + \frac{1}{2\lambda}|x - z|^2 = -\frac{1}{2}x_1 - x_2 - \frac{5\lambda}{8}.$$

Case (v), $z \in (E_1 \cap E_2) \setminus \{0\}$. Then $\partial f(z) = [(-\frac{1}{2}, 1), (\frac{3}{2}, 1)]$ and $z_1 = 0$, $z_2 > 0$, and the first order necessary condition implies that $x_1 \in [-\frac{\lambda}{2}, \frac{3\lambda}{2}]$, $x_2 > \lambda$ and that the candidate for $f_\lambda(x)$ is

$$f(z) + \frac{1}{2\lambda}|x - z|^2 = \frac{1}{2\lambda}x_1^2 + x_2 - \frac{\lambda}{2}.$$

Case (vi), $z \in (E_2 \cap E_3) \setminus \{0\}$. Then $\partial f(z) = [(-\frac{1}{2}, 1), (\frac{3}{2}, -1)]$ and $z_1 < 0$, $z_1 - z_2 = 0$, and the first order necessary condition implies that $x_1 + x_2 - \frac{\lambda}{2} < 0$, $-x_1 + x_2 + \frac{5\lambda}{2} \in [0, 4\lambda]$, and that the candidate for $f_\lambda(x)$ is

$$f(z) + \frac{1}{2\lambda}|x - z|^2 = \frac{1}{4\lambda}x_1^2 + \frac{1}{4\lambda}x_2^2 - \frac{1}{2\lambda}x_1x_2 + \frac{1}{4}x_1 + \frac{1}{4}x_2 - \frac{\lambda}{16}.$$

Case (vii), $z \in (E_3 \cap E_4) \setminus \{0\}$. Then $\partial f(z) = [(-\frac{1}{2}, -1), (\frac{3}{2}, -1)]$ and $z_1 = 0$, $z_2 < 0$, and the first order necessary condition implies that $x_1 \in [-\frac{\lambda}{2}, \frac{3\lambda}{2}]$, $x_2 < -\lambda$, and that the candidate for $f_\lambda(x)$ is

$$f(z) + \frac{1}{2\lambda}|x - z|^2 = \frac{1}{2\lambda}x_1^2 - x_2 - \frac{\lambda}{2}.$$

Case (viii), $z \in (E_4 \cap E_1) \setminus \{0\}$. Then $\partial f(z) = [(-\frac{1}{2}, -1), (\frac{3}{2}, 1)]$ and $z_1 > 0$, $z_1 + z_2 = 0$, and the first order necessary condition implies that $x_1 - x_2 - \frac{\lambda}{2} > 0$, $-x_1 - x_2 + \frac{5\lambda}{2} \in [0, 4\lambda]$, and that the candidate for $f_\lambda(x)$ is

$$f(z) + \frac{1}{2\lambda}|x - z|^2 = \frac{1}{4\lambda}x_1^2 + \frac{1}{4\lambda}x_2^2 + \frac{1}{2\lambda}x_1x_2 + \frac{1}{4}x_1 - \frac{1}{4}x_2 - \frac{\lambda}{16}.$$

Case (ix), $z = 0$. Then $\partial f(z) = \mathrm{co}\{(\frac{3}{2}, 1), (-\frac{1}{2}, 1), (\frac{3}{2}, -1), (-\frac{1}{2}, -1)\}$, and the first order necessary condition implies that $x \in \mathrm{co}\{(\frac{3\lambda}{2}, \lambda), (-\frac{\lambda}{2}, \lambda), (\frac{3\lambda}{2}, -\lambda), (-\frac{\lambda}{2}, -\lambda)\}$ and that the candidate for $f_\lambda(x)$ is

$$f(z) + \frac{1}{2\lambda}|x - z|^2 = \frac{1}{2\lambda}x_1^2 + \frac{1}{2\lambda}x_2^2.$$

Then, comparing the candidates in the sets where more necessary conditions apply, we obtain the values of f_λ, hence also of its subdifferential. $\qquad\square$

6.2.2. *Calculus of $f_{\lambda,\mu}$, the inf sup convoluted function of f.* For $0 < \mu < \lambda$, we define the sets and the sets ΛM_i on which we give an analytic formula of the inf sup convoluted function $f_{\lambda,\mu}$ which is given in the following lemma:

$$
\begin{aligned}
\Lambda M_1 \quad &= \{(x_1, x_2) \in \mathbb{R}^2, \ x_1 - \tfrac{3(\lambda - \mu)}{2} \geq 0, \ x_1 + x_2 - \tfrac{(5\lambda - \mu)}{2} \geq 0\}, \\
\Lambda M_2 \quad &= \{(x_1, x_2) \in \mathbb{R}^2, \ x_1 + \tfrac{(\lambda - \mu)}{2} \leq 0, \ x_1 - x_2 + \tfrac{3(\lambda - \mu)}{2} \leq 0\}, \\
\Lambda M_3 \quad &= \{(x_1, x_2) \in \mathbb{R}^2, \ x_1 - \tfrac{(\lambda - 3\mu)}{2} \leq 0, \ -x_1 + x_2 + \tfrac{5(\lambda - \mu)}{2} \leq 0\}, \\
\Lambda M_4 \quad &= \{(x_1, x_2) \in \mathbb{R}^2, \ x_1 - \tfrac{(\lambda + \mu)}{2} \geq 0, \ x_1 + x_2 + \tfrac{3(\lambda - \mu)}{2} \leq 0\}, \\
\Lambda M_{1,2} \quad &= \{(x_1, x_2) \in \mathbb{R}^2, \ x_1 \in [-\tfrac{(\lambda - \mu)}{2}, \tfrac{3(\lambda - \mu)}{2}], \ x_2 - (\lambda - \mu) \geq 0\}, \\
\Lambda M_{2,3} \quad &= \{(x_1, x_2) \in \mathbb{R}^2, \ x_1 + x_2 - \tfrac{(\lambda - \mu)}{2} \leq 0, \ x_1 - x_2 + \tfrac{3(\lambda - \mu)}{2} \geq 0, \\
&\qquad\qquad -x_1 + x_2 + \tfrac{5(\lambda - \mu)}{2} \geq 0, x_1 - \tfrac{\mu}{2\lambda - \mu}x_2 - \tfrac{(\lambda - \mu)}{2} \leq 0\}, \\
\Lambda M_{4,1} \quad &= \{(x_1, x_2) \in \mathbb{R}^2, \ x_1 - x_2 - \tfrac{(\lambda - \mu)}{2} \geq 0, \ x_1 + x_2 - \tfrac{5(\lambda - \mu)}{2} \leq 0, \\
&\qquad\qquad x_1 + x_2 + \tfrac{3(\lambda - \mu)}{2} \geq 0, x_1 + \tfrac{\mu}{2\lambda - \mu}x_2 - \tfrac{(\lambda - \mu)}{2} \geq 0\}, \\
\Lambda M_0 \quad &= \{(x_1, x_2) \in \mathbb{R}^2, \ x_1 + x_2 - \tfrac{(\lambda - \mu)}{2} \geq 0, \ x_1 - x_2 - \tfrac{(\lambda - \mu)}{2} \leq 0, \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad x_2 \leq (\lambda - \mu)\}, \\
\Lambda M_{2,3 \cap 4,1} &= \{(x_1, x_2) \in \mathbb{R}^2, \ x_2 \in [-2\lambda + \mu, 0], \ x_1 + \tfrac{\mu}{2\lambda - \mu}x_2 - \tfrac{(\lambda - \mu)}{2} \leq 0, \\
&\qquad\qquad -x_1 + \tfrac{\mu}{2\lambda - \mu}x_2 + \tfrac{(\lambda - \mu)}{2} \leq 0\}, \\
\Lambda M_{3 \cap 4} \quad &= \{(x_1, x_2) \in \mathbb{R}^2, \ x_2 \leq -2\lambda + \mu, \ x_1 \in [\tfrac{\lambda - 3\mu}{2}, \tfrac{\lambda + \mu}{2}]\}.
\end{aligned}
$$

Lemma 6.3. *Let $0 < \mu < \lambda$ and let $x \in \mathbb{R}^2$. Then*

$$
\begin{aligned}
f_{\lambda,\mu}(x) &= \quad \tfrac{3}{2}x_1 + \quad\quad x_2 - \tfrac{13(\lambda - \mu)}{8} && \text{if } x \in \Lambda M_1, \\
f_{\lambda,\mu}(x) &= \quad -\tfrac{1}{2}x_1 + \quad\quad x_2 - \tfrac{5(\lambda - \mu)}{8} && \text{if } x \in \Lambda M_2, \\
f_{\lambda,\mu}(x) &= \quad \tfrac{3}{2}x_1 - \quad\quad x_2 - \tfrac{13(\lambda - \mu)}{8} && \text{if } x \in \Lambda M_3, \\
f_{\lambda,\mu}(x) &= \quad -\tfrac{1}{2}x_1 - \quad\quad x_2 - \tfrac{5(\lambda - \mu)}{8} && \text{if } x \in \Lambda M_4, \\
f_{\lambda,\mu}(x) &= \quad \tfrac{1}{2(\lambda - \mu)}x_1^2 + \quad\quad x_2 - \tfrac{(\lambda - \mu)}{2} && \text{if } x \in \Lambda M_{1,2}, \\
f_{\lambda,\mu}(x) &= \tfrac{1}{4(\lambda - \mu)}x_1^2 + \tfrac{1}{4(\lambda - \mu)}x_2^2 - \tfrac{1}{2(\lambda - \mu)}x_1 x_2 + \tfrac{1}{4}x_1 + \tfrac{1}{4}x_2 - \tfrac{(\lambda - \mu)}{16} && \text{if } x \in \Lambda M_{2,3}, \\
f_{\lambda,\mu}(x) &= \tfrac{1}{4(\lambda - \mu)}x_1^2 + \tfrac{1}{4(\lambda - \mu)}x_2^2 + \tfrac{1}{2(\lambda - \mu)}x_1 x_2 + \tfrac{1}{4}x_1 - \tfrac{1}{4}x_2 - \tfrac{(\lambda - \mu)}{16} && \text{if } x \in \Lambda M_{4,1}, \\
f_{\lambda,\mu}(x) &= \tfrac{1}{2(\lambda - \mu)}x_1^2 + \tfrac{1}{2(\lambda - \mu)}x_2^2 && \text{if } x \in \Lambda M_0, \\
f_{\lambda,\mu}(x) &= \quad -\tfrac{1}{2\mu}x_1^2 - \tfrac{1}{2(2\lambda - \mu)}x_2^2 - \tfrac{\lambda}{2\mu}x_1 + \tfrac{\lambda(\lambda - \mu)}{8\mu} && \text{if } x \in \Lambda M_{2,3 \cap 4,1}, \\
f_{\lambda,\mu}(x) &= \quad -\tfrac{x_1^2}{2\mu} + \quad \tfrac{\lambda}{2\mu}x_2^2 - x_2 - \tfrac{\lambda^2 - 4\mu^2 + 7\lambda}{8\mu} && \text{if } x \in \Lambda M_{3 \cap 4}.
\end{aligned}
$$
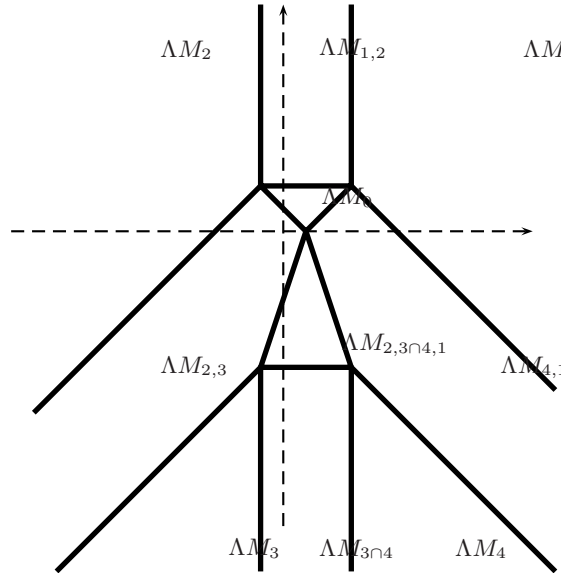
Once the infsup convoluted function $f_{\lambda,\mu}$ is computed, one immediately deduces the values of its gradient, which in particular implies Theorem 2.1, case (b).

Corollary 6.1. *Let $0 < \mu < \lambda$. Then*

$$\nabla f_{\lambda,\mu}(\mathbb{R}^2) = \mathrm{co}\left\{\left(\frac{1}{2},0\right), \left(\frac{3}{2},1\right), \left(-\frac{1}{2},1\right)\right\}$$

$$\cup \, \mathrm{co}\left\{\left(\frac{1}{2},0\right), \left(-\frac{1}{2},-1\right), \left(\frac{3}{2},-1\right)\right\},$$

which does not contain 0.

Proof of Lemma 6.3. We now compute $f_{\lambda,\mu}(x) = -(-f_\lambda)_\mu(x) = -\inf\{-f_\lambda(z) + \frac{1}{2\lambda}|x-z|^2, z \in \mathbb{R}^2\}$. In fact the infimum in the preceeding formula is attained at a point $z \in \mathbb{R}^2$ which satisfies the first order necessary condition

(6.6) $$0 \in \partial(-f_\lambda)(z) + \frac{1}{\mu}(z - x).$$

Noting that f_λ is of class $\mathcal{C}^1$ on $\mathbb{R}^2 \setminus \{\frac{\lambda}{2}\} \times \mathbb{R}_-$, we distinguish the (not necessarily disjoint) cases (i) $z \in \Lambda_1$, (ii) $z \in \Lambda_2$, (iii) $z \in \Lambda_3 \setminus \{\frac{\lambda}{2}\} \times \mathbb{R}_-$, (iv) $z \in \Lambda_4 \setminus \{\frac{\lambda}{2}\} \times \mathbb{R}_-$, (v) $z \in \Lambda_{1,2}$, (vi) $z \in \Lambda_{2,3} \setminus \{\frac{\lambda}{2}\} \times \mathbb{R}_-$, (vii) $z \in \Lambda_{4,1} \setminus \{\frac{\lambda}{2}\} \times \mathbb{R}_-$, (viii) $z \in \Lambda_0$, (ix) $z \in \Lambda_{2,3} \cap \Lambda_{4,1}$, (x) $z \in \Lambda_3 \cap \Lambda_4$.

Case (i), $z \in \Lambda_1$. Then $\nabla f_\lambda(z) = (\frac{3}{2},1)$ and $z_1 - \frac{3\lambda}{2} \geq 0, z_1 + z_2 - \frac{5\lambda}{2} \geq 0$. The first order necessary condition implies that $x_1 + \frac{(\lambda-\mu)}{2} \leq 0, x_1 - x_2 + \frac{3(\lambda-\mu)}{2} \leq 0$ and that the candidate for $(-f_\lambda)_\mu(x)$ is

$$-f_\lambda(z) + \frac{\mu}{2}|x-z|^2 = -\frac{3}{2}x_1 - x_2 + \frac{13(\lambda-\mu)}{8}.$$

Case (ii), $z \in \Lambda_2$. Then $\nabla f_\lambda(z) = (-\frac{1}{2}, 1)$ and $z_1 + \frac{\lambda}{2} \leq 0, z_1 - z_2 + \frac{3\lambda}{2} \leq 0$. The first order necessary condition implies that $x_1 - \frac{3(\lambda-\mu)}{2} > 0$, $x_1 + x_2 - \frac{5(\lambda-\mu)}{2} > 0$ and that the candidate for $(-f_\lambda)_\mu(x)$ is

$$-f_\lambda(z) + \frac{\mu}{2}|x - z|^2 = \frac{1}{2}x_1 - x_2 + \frac{5(\lambda - \mu)}{8}.$$

Case (iii), $z \in \Lambda_3 \setminus \{\frac{\lambda}{2}\} \times \mathbb{R}_-$. Then $\nabla f_\lambda(z) = (\frac{3}{2}, -1)$ and $z_1 - \frac{\lambda}{2} < 0, -z_1 + z_2 + \frac{5\lambda}{2} \leq 0$. The first order necessary condition implies that $x_1 - \frac{\lambda}{2} + \frac{3\mu}{2} < 0, -x_1 + x_2 + \frac{5(\lambda-\mu)}{2} \leq 0$ and that the candidate for $(-f_\lambda)_\mu(x)$ is

$$-f_\lambda(z) + \frac{\mu}{2}|x - z|^2 = -\frac{3}{2}x_1 + x_2 + \frac{13\lambda}{8}.$$

Case (iv), $z \in \Lambda_4 \setminus \{\frac{\lambda}{2}\} \times \mathbb{R}_-$. Then $\nabla f_\lambda(z) = (-\frac{1}{2}, -1)$ and $z_1 - \frac{\lambda}{2} > 0, z_1 + z_2 + \frac{3\lambda}{2} \leq 0$. The first order necessary condition implies that $x_1 - \frac{\lambda}{2} + \frac{\mu}{2} > 0, x_1 + x_2 + \frac{3(\lambda-\mu)}{2} \leq 0$ and that the candidate for $(-f_\lambda)_\mu(x)$ is

$$-f_\lambda(z) + \frac{\mu}{2}|x - z|^2 = \frac{1}{2}x_1 + x_2 + \frac{5(\lambda - \mu)}{8}.$$

Case (v), $z \in \Lambda_{1,2}$. Then $\nabla f_\lambda(z) = (\frac{z_1}{\lambda}, 1)$ and $z_1 \in [-\frac{\lambda}{2}, \frac{3\lambda}{2}], z_2 \geq \lambda$. The first order necessary condition implies that $x_1 \in [-\frac{(\lambda-\mu)}{2}, \frac{3(\lambda-\mu)}{2}], x_2 \geq (\lambda - \mu)$ and that the candidate for $(-f_\lambda)_\mu(x)$ is

$$-f_\lambda(z) + \frac{\mu}{2}|x - z|^2 = -\frac{1}{2(\lambda - \mu)}x_1^2 - x_2 + \frac{(\lambda - \mu)}{2}.$$

Case (vi), $z \in \Lambda_{2,3} \setminus \{\frac{\lambda}{2}\} \times \mathbb{R}_-$. Then $\nabla f_\lambda(z) = (\frac{z_1}{2\lambda} - \frac{z_2}{2\lambda} + \frac{1}{4}, -\frac{z_1}{2\lambda} + \frac{z_2}{2\lambda} + \frac{1}{4})$ and $z_1 + z_2 - \frac{\lambda}{2} \leq 0, z_1 - z_2 + \frac{3\lambda}{2} \geq 0, -z_1 + z_2 + \frac{5\lambda}{2} \geq 0, z_1 - \frac{\lambda}{2} < 0$. The first order necessary condition implies that $x_1 + x_2 - \frac{(\lambda-\mu)}{2} \leq 0, x_1 - x_2 + \frac{3(\lambda-\mu)}{2} \geq 0, -x_1 + x_2 + \frac{5(\lambda-\mu)}{2} \geq 0, \frac{2\lambda-\mu}{\lambda-\mu}x_1 - \frac{\mu}{\lambda-\mu}x_2 + \frac{\mu}{2} - \lambda < 0$ and that the candidate for $(-f_\lambda)_\mu(x)$ is

$$-f_\lambda(z) + \frac{\mu}{2}|x - z|^2 = -\frac{1}{4(\lambda - \mu)}x_1^2 - \frac{1}{4(\lambda - \mu)}x_2^2$$
$$+ \frac{1}{2(\lambda - \mu)}x_1 x_2 - \frac{1}{4}x_1 - \frac{1}{4}x_2 + \frac{(\lambda - \mu)}{16}.$$

Case (vii), $z \in \Lambda_{4,1} \setminus \{\frac{\lambda}{2}\} \times \mathbb{R}_-$. Then $\nabla f_\lambda(z) = (\frac{z_1}{2\lambda} + \frac{z_2}{2\lambda} + \frac{1}{4}, \frac{z_1}{2\lambda} + \frac{z_2}{2\lambda} - \frac{1}{4})$ and $z_1 - z_2 - \frac{\lambda}{2} \geq 0, z_1 + z_2 - \frac{5\lambda}{2} \leq 0, z_1 + z_2 + \frac{3\lambda}{2} \geq 0, z_1 - \frac{\lambda}{2} > 0$. The first order necessary condition implies that $x_1 - x_2 - \frac{(\lambda-\mu)}{2} \geq 0, x_1 + x_2 - \frac{5(\lambda-\mu)}{2} \leq 0, x_1 + x_2 + \frac{3(\lambda-\mu)}{2} \geq 0, \frac{2\lambda-\mu}{\lambda-\mu}x_1 + \frac{\mu}{\lambda-\mu}x_2 + \frac{\mu}{2} - \lambda > 0$ and that the candidate for $(-f_\lambda)_\mu(x)$ is

$$-f_\lambda(z) + \frac{\mu}{2}|x - z|^2 = -\frac{1}{4(\lambda - \mu)}x_1^2 - \frac{1}{4(\lambda - \mu)}x_2^2$$
$$- \frac{1}{2(\lambda - \mu)}x_1 x_2 - \frac{1}{4}x_1 + \frac{1}{4}x_2 + \frac{(\lambda - \mu)}{16}.$$

Case (viii), $z \in \Lambda_0$. Then $\nabla f_\lambda(z) = (\frac{z_1}{\lambda}, \frac{z_2}{2\lambda})$ and $z_1 + z_2 - \frac{\lambda}{2} \geq 0, z_1 - z_2 - \frac{\lambda}{2} \leq 0, z_2 \leq \lambda$. The first order necessary condition implies that $x_1 + x_2 - \frac{(\lambda-\mu)}{2} \geq$

$0, x_1 - x_2 - \frac{(\lambda - \mu)}{2} \leq 0, x_2 \leq (\lambda - \mu)$ and that the candidate for $(-f_\lambda)_\mu(x)$ is

$$-f_\lambda(z) + \frac{\mu}{2}|x - z|^2 = -\frac{1}{2(\lambda - \mu)}x_1^2 - \frac{1}{2(\lambda - \mu)}x_2^2.$$

Case (ix), $z \in \Lambda_{2,3} \cap \Lambda_{4,1}$. Then $\partial f_\lambda(z) = [(\frac{1}{2} + \frac{z_1}{2\lambda}, \frac{z_2}{2\lambda}), (\frac{1}{2} - \frac{z_1}{2\lambda}, \frac{z_2}{2\lambda})]$ and $z_1 = \frac{\lambda}{2}, z_2 \in [-2\lambda, 0]$. The first order necessary condition implies that $x_2 \in [-2\lambda + \mu, 0], x_1 + \frac{\mu}{2\lambda - \mu}x_2 - \frac{(\lambda - \mu)}{2} \leq 0, -x_1 + \frac{\mu}{2\lambda - \mu}x_2 + \frac{(\lambda - \mu)}{2} \leq 0$ and that the candidate for $(-f_\lambda)_\mu(x)$ is

$$x_1 - f_\lambda(z) + \frac{\mu}{2}|x - z|^2 = \frac{1}{2\mu}x_1^2 - \frac{1}{2(2\lambda - \mu)}x_2^2 - \frac{\lambda}{2\mu}x_1 + \frac{\lambda(\lambda - \mu)}{8\mu}.$$

Case (x), $z \in \Lambda_3 \cap \Lambda_4$. Then $\partial f_\lambda(z) = [(-\frac{1}{2}, -1), (\frac{3}{2}, -1)]$ and $z_1 = \frac{\lambda}{2}, z_2 \leq -2\lambda$. The first order necessary condition implies that $x_2 \leq -2\lambda + \mu, x_1 \in [\frac{\lambda - 3\mu}{2}, \frac{\lambda + \mu}{2}]$ and that the candidate for $(-f_\lambda)_\mu(x)$ is

$$-f_\lambda(z) + \frac{\mu}{2}|x - z|^2 = \frac{x_1^2}{2\mu} - \frac{\lambda}{2\mu}x_2^2 + x_2 + \frac{\lambda^2 - 4\mu^2 + 7\lambda}{8\mu}.$$

Then, comparing the candidates in the sets where more necessary conditions apply, we obtain the values of $(-f_\lambda)_\mu$, hence of $f_{\lambda,\mu}(x) = -(-f_\lambda)_\mu(x)$, hence also of its gradient. $\qquad\square$

6.3. **Proof of part (c).** Let (f_n) be a sequence of locally Lipschitz functions from $\mathbb{R}^N$ to $\mathbb{R}$ which converges uniformly to f on compact sets. From Proposition 3.2, the set $\liminf_{n \to +\infty} \text{graph}(\partial f_n)$ contains $\text{graph}\,\partial_P f$ and $\text{graph}\,\partial^L f$. The calculus of the proximal subgradient and of the limiting uppergradient is elementary. For example, the computations given below in Section 8.3 (considering the points which have only one projection on the set $\text{epi}\,f$) are sufficient to obtain the set $\partial_P f(0)$. Anyway, we have

$$\partial_P f(0) = \text{co}\left\{\left(\frac{1}{2}, 0\right), \left(\frac{3}{2}, 1\right), \left(-\frac{1}{2}, 1\right)\right\} \cup \left[\left(-\frac{1}{2}, 1\right), \left(\frac{3}{2}, -1\right)\right]$$
$$\cup \left[\left(-\frac{1}{2}, -1\right), \left(\frac{3}{2}, 1\right)\right],$$

$$\partial^L f(0) = \left[\left(-\frac{1}{2}, -1\right), \left(\frac{3}{2}, -1\right)\right] \cup \left\{\left(-\frac{1}{2}, 1\right), \left(\frac{3}{2}, 1\right)\right\}.$$

So the proof of part (c) will be finished once we prove

$$\{0\} \times \text{int co}\left\{\left(\frac{1}{2}, 0\right), \left(-\frac{1}{2}, -1\right), \left(\frac{3}{2}, -1\right)\right\} \subset \liminf_{n \to +\infty} \text{graph}(\partial f_n).$$

It is a consequence of the following lemma.

Lemma 6.4. *For every* $v \in \text{int co}\left\{\left(\frac{1}{2}, 0\right), \left(-\frac{1}{2}, -1\right), \left(\frac{3}{2}, -1\right)\right\}$ *and every* $\alpha > 0$, *there exists* $\beta > 0$ *such that, for every locally Lipschitz function* $g : \mathbb{R}^N \to \mathbb{R}$ *such that* $\|f - g\|_\infty \leq \beta$,[9] *there exists* $x \in [-\alpha, \alpha]^2$ *and* $\zeta \in \partial f(x)$ *such that*

$$\|\zeta - v\| \leq \alpha.$$

---

[9]It is sufficient to assume $\|f|_{[-\alpha,\alpha]^2} - g|_{[-\alpha,\alpha]^2}\|_\infty \leq \beta$.

Proof of Lemma 6.4. First note that one can approach a locally Lipschitz function g by C^∞ functions having their gradients close to the generalized gradient of g by using a classical convolution argument. Hence, without loss of generality, we only prove Lemma 6.4 for C^∞ functions.

Consider an element $v = (v_1, v_2) \in \operatorname{int} \operatorname{co} \left\{ \left(\frac{1}{2}, 0 \right), \left(-\frac{1}{2}, -1 \right), \left(\frac{3}{2}, -1 \right) \right\}$, i.e., $v_2 < v_1 - \frac{1}{2}$, $v_2 < -v_1 + \frac{1}{2}$ and $-1 < v_2$. Take a real number $\alpha > 0$. We define the elements A and B by

$$A := \frac{\alpha}{2} \left(\frac{1 + v_2}{\frac{1}{2} + v_1}, -1 \right),$$

$$B := \frac{\alpha}{2} \left(1, \frac{\frac{3}{2} - v_1}{v_2 - 1} \right).$$

A straightforward computation gives the level sets of the function $fv := f - \langle v, . \rangle$. In particular, we note that the value of fv on the set $[A, 0] \cup [0, B]$ is zero, and that the set $\{x \in \mathbb{R}^2 | f_v(x) \geq 0\} \setminus \{0\}$ has two connected components, a component $C_{1,2}$ (containing B) in the set $E_1 \cap E_2$ and a component $C_{3,4}$ (containing A) in the set $E_3 \cap E_4$. An elementary calculus shows that, given the assumption on the element v, we have $d(A, C_{1,2}) > 0$ and $d(B, C_{3,4}) > 0$.

We now choose a real number $\beta > 0$ small enough, namely $\beta \leq \frac{\alpha^2}{6}$, $\beta \leq \frac{\alpha}{3} d(A, C_{1,2})$, $\beta \leq \frac{\alpha}{3} d(B, C_{3,4})$. Take a C^∞ function $g : \mathbb{R}^N \to \mathbb{R}$ such that $\|f - g\|_\infty \leq \beta$ and assume that Lemma 6.4 does not hold, i.e.,

$$\|\nabla g(x) - v\| > \alpha \text{ for every } x \in [-\alpha, \alpha]^2.$$

Define $gv := g - \langle v, . \rangle$ and note that $\|\nabla gv(x)\| > \alpha$ for every $x \in [-\alpha, \alpha]^2$. Let $(t, x) \mapsto \varphi(t, x)$ be the flow of the following differential equation:

$$\dot{x}(t) = \frac{\nabla gv}{\|\nabla gv\|^2}(x(t)), \quad x(0) = x_0, \text{ with } x_0 \in \left[-\frac{\alpha}{2}, \frac{\alpha}{2} \right]^2.$$

Noting that $\|\dot{x}(t)\| < \frac{1}{\alpha}$, and in view of the Peano existence theorem, or a theorem on the extension of solutions, any solution of the above differential equation is defined on the interval $[-\frac{\alpha^2}{2}, \frac{\alpha^2}{2}]$, which contains the interval $[0, 3\beta]$ (since $\beta \leq \frac{\alpha^2}{6}$). Since

$$fv([A, 0] \cup [0, B]) = \{0\},$$

then

$$gv([A, 0] \cup [0, B]) \subset [-\beta, \beta].$$

Note that

$$\frac{\partial}{\partial t} gv(\varphi(t, x)) = 1.$$

Hence

$$gv \left(\varphi \left(3\beta, [A, 0] \cup [0, B] \right) \right) \subset [2\beta, 4\beta],$$

which implies

$$fv \left(\varphi \left(3\beta, [A, 0] \cup [0, B] \right) \right) \subset [\beta, 3\beta].$$
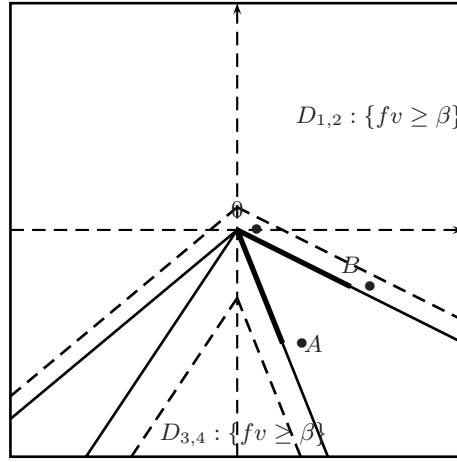
Note that the set

$$\{x | fv(x) \geq \beta\}$$

has two connected components, a component $D_{1,2}$ in the set $E_1 \cap E_2$ and a component $D_{3,4}$ in the set $E_3 \cap E_4$. Since ∇gv is locally Lipschitz, the flow φ is continuous,

hence the set $\varphi\left(3\beta, [A,0] \cup [0,B]\right)$ is connected. Hence it is included in one of the connected components of $\{x | fv(x) \geq \beta\}$. On the other hand, for every x,

$$\|\varphi(3\beta, x) - x\| \leq \frac{3\beta}{\alpha}.$$

We remark that

$$d(A, D_{1,2}) > d(A, C_{1,2}) \geq \frac{3\beta}{\alpha},$$

$$d(B, D_{3,4}) > d(B, C_{3,4}) \geq \frac{3\beta}{\alpha}.$$

Consequently, $\varphi(3\beta, A) \notin D_{1,2}$ and $\varphi(3\beta, B) \notin D_{3,4}$, a contradiction. $\square$

7. Proofs of the special cases: Dimension 1 for functions and 2 for sets, and the related results

In this section, we first prove Proposition 3.1 on the stability of normal cones, which implies Proposition 3.2 on the stability of subgradients. We then prove the results of Section 2.4.

7.1. **Proof of Proposition 3.1.** First note that (3.2), respectively (3.4), is deduced from (3.1), respectively (3.3), by observing that $\mathrm{graph}\,(N_M^L) = \mathrm{cl}\,\mathrm{graph}\,(N_M^P)$. We first prove (3.3) and then deduce (3.1).

7.1.1. *Proof of* (3.3). Let $x \in M$ and $\zeta \in N_M^P(x)$. By definition of the proximal normal cone, there exists $\lambda > 0$ such that $x = \mathrm{proj}_M(y + \lambda\zeta)$. Since $M = \limsup_{n\to+\infty} M_n = \liminf_{n\to+\infty} M_n$, then[10]

$$\lim_{n\to+\infty} d_{Haus}\left(M \cap \overline{B}(x,1) \cup S(x,1), M_n \cap \overline{B}(x,1) \cup S(x,1)\right) = 0.$$

---

[10]Recall that, if (M_n) is a sequence of nonempty closed subsets of $\mathbb{R}^N$, then $M = \liminf_{n\to+\infty} M_n = \limsup_{n\to+\infty} M_n$ if and only if, for every nonempty bounded subset $\mathcal{B}$, $\lim_{n\to+\infty} d_{Haus}(M \cap \mathcal{B} \cup \mathrm{bd}\,\mathcal{B}, M_n \cap \mathcal{B} \cup \mathrm{bd}\,\mathcal{B}) = 0$.

For every $n \in \mathbb{N}$, take

$$x_n \in \mathrm{proj}_{M_n \cap \overline{B}(x,1) \cup S(x,1)} \left(x + \frac{\zeta}{2} \right)$$

and set $2\zeta_n = x + \frac{\zeta}{2} - x_n$. From Lemma 4.5, $\lim_{n \to +\infty} |x_n - x| = 0$, hence $\lim_{n \to +\infty} |\zeta_n - \zeta| = 0$. But, by construction, $\zeta_n \in N^P_{M_n \cap \overline{B}(x,1) \cup S(x,1)}(x_n)$. For n large enough $x_n \in B(x,1)$, which implies $x_n \in M_n$ and

$$N^P_{M_n \cap \overline{B}(x,1) \cup S(x,1)}(x_n) = N^P_{M_n}(x_n),$$

which completes the proof of (3.3). $\qquad\square$

7.1.2. *Proof of* (3.1). It can be found in [12, Lemma 4.1], but we give here a shorter argument. Let $x \in M$ and $\zeta \in N^P_M(x)$. Since $x \in \limsup_{n \to +\infty} M_n$, there is a subsequence $(M_{\varphi(n)})$ such that $x \in \liminf_{n \to +\infty} M_{\varphi(n)}$. Up to a subsequence, in view of the generalized Bolzano-Weierstrass theorem (see [22, §25, VIII]), we may assume that $\liminf_{n \to +\infty} M_{\varphi(n)} = \limsup_{n \to +\infty} M_{\varphi(n)}$. Since $\liminf_{n \to +\infty} M_{\varphi(n)} \subset M$ and $x \in \liminf_{n \to +\infty} M_{\varphi(n)}$, then $N^P_M(x) \subset N^P_{\liminf_{n \to +\infty} M_{\varphi(n)}}(x)$. Applying (3.3) to the sequence $(M_{\varphi(n)})$, we obtain that $(x, \zeta) \in \limsup_{n \to \infty} \mathrm{graph}(N^P_{M_{\varphi(n)}}) \subset \limsup_{n \to \infty} \mathrm{graph}(N^P_{M_n})$. $\qquad\square$

7.2. **Proof of Proposition 3.2.**

7.2.1. *Proof of* (3.5). Let $x \in \mathbb{R}^N$ and $\zeta \in \partial_P f(x)$. We immediately obtain by definition of the proximal subdifferential that $(\zeta, -1) \in N^P_{\mathrm{epi}(f)}(x)$. On the other hand, since the sequence $(f_n)_{n \in \mathbb{N}}$ epi-converges to f, then $\mathrm{epi}\, f = \limsup_{n \to +\infty} \mathrm{epi}\, f_n = \liminf_{n \to +\infty} \mathrm{epi}\, f_n$. Hence by Proposition 3.1 (3.1), since these sets are closed, there exists a sequence $(x_n)_{n \in \mathbb{N}}$ converging to x and a sequence $(\zeta_n, \lambda_n)_{n \in \mathbb{N}}$ of points in $\mathbb{R}^N \times \mathbb{R}$ which converges to $(\zeta, -1)$ and such that

$$\forall n \in \mathbb{N}, (\zeta_n, \lambda_n) \in N^P_{\mathrm{epi}(f_n)}(x_n).$$

Therefore we deduce that for n large enough $-\frac{1}{\lambda_n} \zeta_n \in \partial_P f_n(x_n)$, and that the sequence $(-\frac{1}{\lambda_n} \zeta_n)_{n \in \mathbb{N}}$ converge to ζ. $\qquad\square$

7.2.2. *Proof of* (3.6). It follows immediately by construction of the limiting subdifferential and by the closedness of the lim inf of sets. $\qquad\square$

7.2.3. *Proof of* (3.7). This is a consequence of Proposition 4.5, Chapter 3, page 138, of the book of Clarke, Ledyaev, Stern and Wolenski [9], which follows from Subbotin's theorem [36, Theorem 4.2] which asserts that every D-subdifferential can be approximated by proximal subgradients. $\qquad\square$

7.3. **Proof of Lemma 2.1.** Let $x \in \mathbb{R}$ and $\zeta \in \partial f(x)$. We want to prove that $\zeta \in \partial_P f(x) \cup \partial^L f(x)$ (the second equality is left to the reader). Adding an affine function if necessary, we can assume that $\zeta = 0$. We argue by contradiction, hence we suppose that $0 \notin \partial_P f(x)$ and that $0 \notin \partial^L f(x)$. Since the multivalued mapping $y \mapsto \partial^L f(y)$ is upper semicontinuous, this means that there exists two positive constants α, β such that for any y with $|y - x| < \varepsilon$,

$$\partial^L f(y) \cap (-\alpha, \beta) = \emptyset.$$

Let us remark that if a function has a local extrema at a point $z \in \mathbb{R}$, then $0 \in \partial_P f(z) \cup \partial^P f(z)$. Therefore, we deduce that our function f has neither local maximum nor local minimum and hence that f is either increasing or decreasing. But if a function is increasing, then its proximal superdifferential has values in $[0, \infty)$, and if it is decreasing, then its proximal superdifferential has values in $(-\infty, 0]$. We deduce in both cases that the limiting supergradient has a constant sign. Since $0 \in \partial f(x) = \mathrm{co}\,\partial^L f(x_0)$ and since $\partial^L f(x) \cap (-\alpha, \beta) = \emptyset$, we get a contradiction. $\square$

7.4. **Proof of Lemma 2.2.** Let x be in M. If x is in the interior of M, then the conclusion of the lemma is obvious. If x is on the boundary of M, then since M is epi-Lipschitz there exists a neighbourhood $\mathcal{V}$ of x such that $M \cap \mathcal{V}$ can be seen as the epigraph of some Lipschitz function. In fact, without loss of generality, we can assume that $x = 0$ and that there exists $f : H \to \mathbb{R}$ Lipschitz such that

$$M \cap \mathcal{V} = \mathrm{epi}\,(f) \cap \mathcal{V},$$

where H denotes the hyperplane such that $H \overset{\oplus}{\perp} e_N = \mathbb{R}^N$ (where $e_N = (0, \cdots, 0, 1)$ is the last vector of the canonical basis of $\mathbb{R}^N$). Hence by the fundamental properties of the proximal sub- and superdifferentials, we have that for any $y \in H \cap \mathcal{V}$,

$$\zeta \in \partial_P f(y) \iff (\zeta, -1) \in N_M^P(y, f(y)),$$
$$\zeta \in \partial^P f(y) \iff (\zeta, -1) \in N_{\hat{M}}^P(y, f(y)).$$

The same equivalences hold for the limiting sup- and superdifferentials, and moreover

$$\zeta \in \partial f(y) \iff (\zeta, -1) \in N_M^C(y, f(y)).$$

We let the reader deduce the conclusions of the lemma from Lemma 2.1. $\square$

7.5. **Proof of Proposition 2.3.** It is an immediate consequence of Proposition 3.2 and Lemma 2.1. $\square$

7.6. **Proof of Proposition 2.4.** It is an immediate consequence of Proposition 3.1 and Lemma 2.2. $\square$

8. Proof of the consequences

8.1. **Proof of Corollary 3.2.** Let $\Omega = \{x \in U | \delta(x) > 0\}$. Since Theorem 2.2 corresponds to the case $\Omega = U$, the proof of Corollary 3.2 consists in properly gluing the approximation obtained on Ω and the function f on $U \setminus \Omega$. This kind of gluing was largely developed in [15], where equivalence was shown between formulations with $\Omega = U$ and $\Omega \neq U$. So the techniques used below can already be found in [15] and the associated references. Let

$$\varepsilon(x) = \min\left\{\delta(x), \frac{d(\mathbb{R}^N \setminus U, x)}{2}, d(\mathbb{R}^N \setminus \Omega, x), d(\mathbb{R}^N \setminus \Omega, x)^2\right\}.$$

Noting that $\varepsilon(x) > 0 \Leftrightarrow \delta(x) > 0 \Leftrightarrow x \in \Omega$, let $f_\varepsilon : \Omega \to \mathbb{R}$ be given by Theorem 2.2 (associated to the function ε). Let us define the function $f_\delta : U \to \mathbb{R}$ by

(8.1)
$$\begin{aligned} f_\delta(x) &:= f_\varepsilon(x) \text{ if } x \in \Omega, \\ f_\delta(x) &:= f(x) \text{ if } x \in U \setminus \Omega. \end{aligned}$$

Proof of (3.8). Let W be a compact convex subset of U. Let L be the Lipschitz constant of f on the compact set $\bigcup_{x \in W} \overline{B}(x, \varepsilon(x))$. Consider x and x' in W. If $(x, x') \in U \setminus \Omega \times U \setminus \Omega$, clearly $|f_\delta(x) - f_\delta(x')| = |f(x) - f(x')| \leq L|x - x'|$. If $\max\{d(\mathbb{R}^N \setminus \Omega, x), d(\mathbb{R}^N \setminus \Omega, x')\} \leq |x - x'|$, since $|f_\delta(x) - f(x)| \leq \delta(x) \leq d(\mathbb{R}^N \setminus \Omega, x)$ and $|f_\delta(x') - f(x')| \leq \delta(x') \leq d(\mathbb{R}^N \setminus \Omega, x')$, we obtain

$$|f_\delta(x) - f_\delta(x')| \leq d(\mathbb{R}^N \setminus \Omega, x) + d(\mathbb{R}^N \setminus \Omega, x') + |f(x) - f(x')| \leq (L + 2)|x - x'|.$$

Without any loss of generality, let us finally consider the case $|x - x'| < d(\mathbb{R}^N \setminus \Omega, x)$. Then $[x, x'] \subset B(x, d(\mathbb{R}^N \setminus \Omega, x)) \subset \Omega$, and, for some $z \in [x, x']$,

$$|f_\delta(x) - f_\delta(x')| \leq |\nabla f_{\varepsilon(x)}(z)| \, |x - x'|.$$

But $[x, x'] \subset W$ and from (2.7), $\nabla f_\varepsilon(z) \in \partial f(z + \varepsilon(z)B) + \varepsilon(z)B$. Hence $|\nabla f_{\varepsilon(x)}(z)| \leq L + \varepsilon(z)$ and $|f_\delta(x) - f_\delta(x')| \leq (L + \max \varepsilon(W))|x - x'|$, which finishes the proof of (3.8). $\qquad \square$

Proof of (3.9). It is immediate from (2.6) and from (8.1). $\qquad \square$

We now need the following result.

Lemma 8.1. *Let $x \in U \setminus \Omega$. Then f is differentiable at x if and only if f_δ is differentiable at x, and then*

$$\nabla f_\delta(x) = \nabla f(x).$$

Proof of Lemma 8.1. From (8.1), the proof is immediate if $x \in U \setminus \mathrm{cl}\,\Omega$. Now assume that $x \in \mathrm{bd}\,\Omega$. We first assume that f is differentiable at x. Then, for every $x' \in U$,

$$f(x') = f(x) + \langle \nabla f(x), x' - x \rangle + o(|x' - x|).^{11}$$

We proved $|f_\varepsilon - f| \leq \varepsilon$, we recall $\varepsilon(x') \leq d(\mathbb{R}^N \setminus \Omega, x')^2$, and moreover $d(\mathbb{R}^N \setminus \Omega, x') \leq |x' - x|$ since $x \in \mathrm{bd}\,\Omega$. Hence

$$|f_\varepsilon(x') - f(x')| \leq |x' - x|^2 = o(|x' - x|).$$

Since $\varepsilon(x) = 0$, we have $f_\varepsilon(x) - f(x) = 0$, and we obtain

$$f_\varepsilon(x') = f_\varepsilon(x) + \langle \nabla f(x), x' - x \rangle + o(|x' - x|).$$

Thus f_ε is differentiable at x and $\nabla f_\varepsilon(x) = \nabla f(x)$. The converse is proved in the same way. $\qquad \square$

Proof of (3.10). If $x \in \Omega$, the inclusion is immediate from (2.8). It is also clear if $x \in U \setminus \mathrm{cl}\,\Omega$. Now consider $x \in \mathrm{bd}\,\Omega$. In view of the definition of ∂f_δ and the convexity of ∂f, it is sufficent to prove the following inclusion:

$$\left\{ \lim_{n \to +\infty} \nabla f_\delta(x_n), (x_n) \subset \mathrm{Dom}\,(\nabla f_\delta), x_n \to x \right\} \subset \partial f(x).$$

Indeed, let $v \in \mathbb{R}^N$ and let (x_n) be a sequence in $\mathrm{Dom}\,(\nabla f_\delta)$ converging to x such that $v = \lim_{n \to +\infty} \nabla f_\delta(x_n)$. Without any loss of generality, we may assume that one of the two following cases holds: (a) $(x_n) \subset U \setminus \Omega$ or (b) $(x_n) \subset \Omega$. In case (a), then, from Lemma 8.1, f is differentiable at x_n and $\nabla f_\delta(x_n) = \nabla f(x_n)$. Hence $v = \lim_{n \to +\infty} \nabla f(x_n) \in \partial f(x)$. Assume now that case (b) holds. Then $\nabla f_\delta(x_n) = \nabla f_\varepsilon(x_n)$. But $\nabla f_\varepsilon(x_n) \in \partial f(x_n + \varepsilon(x_n)B) + \varepsilon(x_n)B$. Since $\lim_{n \to +\infty} \varepsilon(x_n) = 0$, we obtain $v = \lim_{n \to +\infty} \nabla f_\varepsilon(x_n) \in \partial f(x)$. $\qquad \square$

---

[11]I.e., there is a function $\varepsilon : \mathbb{R} \to \mathbb{R}$ such that $\lim_{t \to 0} \varepsilon(t) = 0$, and $o(|h|) = |h|\varepsilon(|h|)$ for all $h \in \mathbb{R}^N$.

Proof of (3.11). In view of (2.8), the proof is done in the same way as the previous one, when one notices additionally that $B(x, \varepsilon(x)) \subset \Omega$ for every $x \in \Omega$. $\square$

8.2. **Proof of Corollary 3.3.** Let f_M be given by Corollary 3.2, for $f = d_M$ and $\delta = \frac{d_M}{2}$. For $\lambda > 0$, we define the set M_λ as follows:

$$M_\lambda = \{x \in \mathbb{R}^N | f_M(x) \leq \lambda\}.$$

We claim that, for some $\lambda_0 > 0$, the family $(M_\lambda)_{\lambda_0 > \lambda > 0}$ satisfies:
   (i)′   for every λ, M_λ is a compact and smooth subset of $\mathbb{R}^n$, i.e., is a closed C^∞ submanifold with a boundary of $\mathbb{R}^N$, of full dimension;
   (ii)′   for every $\lambda' < \lambda$, $M_{\lambda'} \subset M_\lambda \subset B(M, 1)$, and $M = \bigcap_{\lambda_0 > \lambda > 0} M_\lambda$;
   (ret)′   for every λ, M is a deformation retract of M_λ;
   (cn)′ $\mathrm{graph}\,(\widetilde{N}_M) = \limsup_{\lambda \to 0, \lambda_0 > \lambda > 0} \mathrm{graph}\,(N^P_{M_\lambda})$.
We refer to [14, Proof of Theorem 2.1] for the definition of λ_0, (i)′, (ii)′, (ret)′ and the inclusion $\limsup_{\lambda_0 > \lambda > 0} \mathrm{graph}\,(N_{M_\lambda}) \subset \mathrm{graph}\,(\widetilde{N}_M)$. We only note that, since $|f - d_M| < \frac{d_M}{2}$, we have the inclusion $f_M^{-1}((0, 1/2]) \subset B(\mathrm{bd}\,M, 1)$ which proves the inclusion $M_\lambda \subset B(M, 1)$ and the compacity of M_λ.

We now prove the inclusion $\mathrm{graph}\,(\widetilde{N}_M) \subset \limsup_{\lambda \to 0, \lambda_0 > \lambda > 0} \mathrm{graph}\,(N^P_{M_\lambda})$. Take $(x, v) \in \mathrm{graph}\,(\widetilde{N}_M)$. From the definition of $\widetilde{N}_M$, there is an element w in the set $\limsup_{x' \to x, d_M(x') > 0} \partial d_M(x')$ and $\mu \geq 0$ such that $v = \mu w$. Then there is a sequence (x_n, w_n) in $\mathbb{R}^N \times \mathbb{R}^N$ such that $x = \lim_{n \to +\infty} x_n$, $d_M(x_n) > 0$, $w_n \in \partial d_M(x_n)$ and $w = \lim_{n \to +\infty} w_n$. From Corollary 3.2

$$\partial d_M(x_n) \subset \partial f_M\Big(\overline{B}(x, \frac{d_M(x_n)}{2})\Big) + \overline{B}(0, \frac{d_M(x_n)}{2})$$
$$= \nabla f_M\Big(\overline{B}(x, \frac{d_M(x_n)}{2})\Big) + \overline{B}(0, \frac{d_M(x_n)}{2}).$$

Hence there is $x'_n \in \overline{B}(x, \frac{d_M(x_n)}{2})$ such that

$$|w_n - \nabla f_M(x'_n)| \leq \frac{d_M(x_n)}{2}.$$

Since $x = \lim_{n \to +\infty} x_n$, then $\lim_{n \to +\infty} d_M(x_n) = 0$, hence $x = \lim_{n \to +\infty} x'_n$ and $\lim_{n \to +\infty} d_M(x'_n) = 0$. Therefore $\lim_{n \to +\infty} f_M(x'_n) = 0$, and for n large enough, $f_M(x'_n) \leq \lambda_0$, hence

$$\nabla f_M(x'_n) \neq 0, \quad x'_n \in \mathrm{bd}\,M_{f_M(x'_n)}, \quad \text{and} \quad \nabla f_M(x'_n) \in N^P_{M_{f_M(x'_n)}}(x'_n).$$

Then

$$(x'_n, \mu \nabla f_M(x'_n)) \in \mathrm{graph}\,(N^P_{M_{f_M(x'_n)}}),$$
$$(x, v) = (x, \mu w) = \lim_{n \to +\infty} (x'_n, \mu \nabla f_M(x'_n)).$$

Since $\lim_{n \to +\infty} f_M(x'_n) = 0$, then $(x, v) \in \limsup_{\lambda \to 0, \lambda_0 > \lambda > 0} \mathrm{graph}\,(N^P_{M_\lambda})$. The proof of Corollary 3.3 is finished in view of the following lemma.

Lemma 8.2. *Let $(G_\lambda)_{\lambda > 0}$ be a family of closed subsets of $\mathbb{R}^N$. Then there is a sequence $(\lambda_n) \subset \mathbb{R}_+ \setminus \{0\}$ such that $\lim_{n \to +\infty} \lambda_n = 0$ and*

$$\limsup_{\lambda \to 0} G_\lambda = \limsup_{n \to +\infty} G_{\lambda_n}.$$

The proof of Lemma 8.2 is achieved by considering a dense countable subset of the closed set $\limsup_{\lambda \to 0} G_\lambda$ and with a diagonal argument. It is left to the reader.

8.3. **Proof of Proposition 3.4.** The proof of Proposition 3.4 is elementary and is achieved by computing Clarke's generalized gradients of $d_{\text{epi}\,f}(x)$, the distance function to the epigraph of f, at $x \notin \text{epi}\,f$, with $x = (x_1, x_2, x_3)$. Indeed, since $f(\lambda x_1, \lambda x_2) = \lambda f(x_1, x_2)$, then $\text{proj}_{\text{epi}\,f}(\lambda x) = \lambda \text{proj}_{\text{epi}\,f}(x)$, $\partial d_{\text{epi}\,f}(\lambda x) = \partial d_{\text{epi}\,f}(x)$ and

$$(8.2) \qquad \limsup_{x \to 0, d_{\text{epi}\,f}(x) > 0} \partial d_{\text{epi}\,f}(x) = \partial d_{\text{epi}\,f}(\mathbb{R}^3 \setminus \text{epi}\,f).$$

The proof is achieved by considering 11 different cases. For the comfort of the reader, the number of each case corresponds to the place of $\text{proj}_{\text{epi}\,f}(x)$, as one can see below. The sets E_1, E_2, E_3 and E_4 are defined in Section 6. We denote

$$n_1 := \nabla f|_{\text{int}\,E_1} = (\tfrac{3}{2}, 1),\ n_2 := \nabla f|_{\text{int}\,E_2} = (-\tfrac{1}{2}, 1),$$

$$n_3 := \nabla f|_{\text{int}\,E_3} = (\tfrac{3}{2}, -1),\ n_4 := \nabla f|_{\text{int}\,E_4} = (-\tfrac{1}{2}, -1),$$

$$e_{12} := (0, 1, 1), e_{23} := (-1, -1, -\tfrac{1}{2}), e_{34} := (0, -1, 1), e_{41} := (1, -1, \tfrac{1}{2}),$$

and note

$$f(E_1) = \mathbb{R}_+ e_{41} + \mathbb{R}_+ e_{12}, f(E_2) = \mathbb{R}_+ e_{12} + \mathbb{R}_+ e_{23},$$

$$f(E_3) = \mathbb{R}_+ e_{23} + \mathbb{R}_+ e_{34}, f(E_4) = \mathbb{R}_+ e_{34} + \mathbb{R}_+ e_{41},$$

and graph $f = f(E_1) \cup f(E_2) \cup f(E_3) \cup f(E_4)$.

Case 1: $\det(n_1, e_{41}, x) \leq 0$ *and* $\det(n_1, e_{12}, x) \leq 0$, *i.e.,* $2x_1 + 7x_2 + 10x_3 \geq 0$ *and* $4x_1 - 3x_2 + 3x_3 \geq 0$. The fact that $x \notin \text{epi}\,f$ implies $\tfrac{3}{2}x_1 + x_2 - x_3 > 0$. Then $\text{proj}_{\text{epi}\,f}(x) \in f(E_1)$ and

$$\nabla d_{\text{epi}\,f}(x) = \frac{2}{\sqrt{17}} \left(\frac{3}{2}, 1, -1 \right).$$

Case 1 ∩ 2: $\det(n_1, e_{12}, x) \geq 0$, $\det(n_2, e_{12}, x) \geq 0$, *and* $\langle e_{12}, x \rangle \geq 0$, *i.e.,* $4x_1 - 3x_2 + 3x_3 \leq 0$, $-4x_1 - x_2 + x_3 \geq 0$ $x_2 + x_3 \geq 0$. Then $\text{proj}_{\text{epi}\,f}(x) \in \mathbb{R}_+ e_{12}$ and

$$\nabla d_{\text{epi}\,f}(x) = \lambda \left(x_1, \frac{x_2 - x_3}{2}, \frac{x_3 - x_2}{2} \right),$$

with $\lambda = (x_1^2 + \tfrac{1}{2}(x_2 - x_3)^2)^{-\frac{1}{2}}$.

Case 2: $\det(n_2, e_{12}, x) \leq 0$ *and* $\det(n_2, e_{23}, x) \leq 0$, *i.e.,* $-4x_1 - x_2 + x_3 \leq 0$ *and* $2x_1 - x_2 - 2x_3 \geq 0$. The fact that $x \notin \text{epi}\,f$ implies $-\tfrac{1}{2}x_1 + x_2 - x_3 > 0$. Then $\text{proj}_{\text{epi}\,f}(x) \in f(E_2)$ and

$$\nabla d_{\text{epi}\,f}(x) = \frac{2}{3} \left(-\frac{1}{2}, 1, -1 \right).$$

Case 2 ∩ 3: $\det(n_2, e_{23}, x) \geq 0$, $\det(n_3, e_{23}, x) \geq 0$ *and* $\langle e_{23}, x \rangle \geq 0$, *i.e.,* $2x_1 - x_2 - 2x_3 \leq 0$, $2x_1 + x_3 < 0$ *and* $-x_1 - x_2 - \tfrac{1}{2}x_3 \geq 0$. Then $\text{proj}_{\text{epi}\,f}(x) \in \mathbb{R}_+ e_{23}$ and

$$\nabla d_{\text{epi}\,f}(x) = \mu \left(\frac{5}{9}x_1 - \frac{4}{9}x_2 - \frac{2}{9}x_3, -\frac{4}{9}x_1 + \frac{5}{9}x_2 - \frac{2}{9}x_3, -\frac{2}{9}x_1 - \frac{2}{9}x_2 + \frac{8}{9}x_3 \right),$$

with $\mu = (\tfrac{5}{9}x_1^2 + \tfrac{5}{9}x_2^2 + \tfrac{8}{9}x_3^2 - \tfrac{8}{9}x_1 x_2 - \tfrac{4}{9}x_2 x_3 - \tfrac{4}{9}x_1 x_3)^{-\frac{1}{2}}$.

Case 3: $\det(n_3, e_{23}, x) \leq 0$ $(d(f(E_3), x) < d(f(E_4), x)$ *and* $\det(n_4, e_{41}, x) \leq 0)$ *or* $(d(f(E_3), x) < d(\mathbb{R}_+ e_{41}, x)$ *and* $\det(n_4, e_{41}, x) \geq 0)$, *i.e.,* $2x_1 - 7x_2 + 10x_3 \geq 0$,

$(11 + 3\sqrt{17}x_1 + 4x_2 + 4x_3 < 0$ *and* $-2x_1 - x_2 + 2x_3 \geq 0)$ *or* $(4x_1^2 + 49x_2^2 + 100x_3^2 + 244x_1x_2 + 40x_1x_3 - 4x_2x_3 > 0$ *and* $-2x_1 - x_2 + 2x_3 \leq 0)$. The fact that $x \notin \text{epi } f$ implies $\frac{3}{2}x_1 - x_2 - x_3 > 0$. Then $\text{proj}_{\text{epi } f}(x) \in f(E_3)$ and

$$\nabla d_{\text{epi } f}(x) = \frac{2}{\sqrt{17}}\left(\frac{3}{2}, -1, -1\right).$$

Case $3 \cap 4$ a): $\det(n_3, e_{23}, x) \leq 0$, $\det(n_4, e_{41}, x) \leq 0$ *and* $d(f(E_3), x) = d(f(E_4), x)$, i.e., $2x_1 - 7x_2 + 10x_3 \geq 0$, $-2x_1 - x_2 + 2x_3 \geq 0$ *and* $11 + 3\sqrt{17}x_1 + 4x_2 + 4x_3 = 0$. Then $\text{proj}_{\text{epi } f}(x)$ has two elements, one in $f(E_3)$ and the other in $f(E_4)$, and

$$\partial d_{\text{epi } f}(x) = \left[\frac{2}{\sqrt{17}}\left(\frac{3}{2}, -1, -1\right), \frac{2}{3}\left(-\frac{1}{2}, -1, -1\right)\right].$$

Case $3 \cap 4$ b): $\det(n_3, e_{23}, x) \leq 0$, $d(f(E_3), x) = d(\mathbb{R}_+ e_{41}, x)$, *and* $\det(n_4, e_{41}, x) \geq 0$, i.e., $2x_1 - 7x_2 + 10x_3 \geq 0$, $4x_1^2 + 49x_2^2 + 100x_3^2 + 244x_1x_2 + 40x_1x_3 - 4x_2x_3 = 0$ *and* $-2x_1 - x_2 + 2x_3 \leq 0$. Then $\text{proj}_{\text{epi } f}(x)$ has two elements, one in $f(E_3)$ and the other in $\mathbb{R}_+ e_{41}$ and

$$\partial d_{\text{epi } f}(x) = \left[\frac{2}{\sqrt{17}}\left(\frac{3}{2}, -1, -1\right), \right.$$
$$\left. \nu\left(\frac{5}{9}x_1 + \frac{4}{9}x_2 - \frac{2}{9}x_3, \frac{4}{9}x_1 + \frac{5}{9}x_2 + \frac{2}{9}x_3, -\frac{2}{9}x_1 + \frac{2}{9}x_2 + \frac{8}{9}x_3\right)\right],$$

with $\nu = (\frac{5}{9}x_1^2 + \frac{5}{9}x_2^2 + \frac{8}{9}x_3^2 + \frac{8}{9}x_1x_2 + \frac{4}{9}x_2x_3 - \frac{4}{9}x_1x_3)^{-\frac{1}{2}}$.

Case $2 \cap 3 \bigcap 4 \cap 1$: $\det(n_3, e_{23}, x) \geq 0$, $\langle e_{23}, x\rangle \geq 0$, $\langle e_{41}, x\rangle \geq 0$ *and* $d(\mathbb{R}_+ e_{23}, x) = d(\mathbb{R}_+ e_{41}, x)$, i.e., $2x_1 - 7x_2 + 10x_3 < 0$, $-x_1 - x_2 - \frac{1}{2}x_3 \geq 0$, $x_1 - x_2 + \frac{1}{2}x_3 \geq 0$, *and* $-2x_1 - x_2 + 2x_3 = 0$. Then $\text{proj}_{\text{epi } f}(x)$ has two elements, one in $\mathbb{R}_+ e_{23}$ and the other in $\mathbb{R}_+ e_{41}$, and

$$\partial d_{\text{epi } f}(x) = \left[\mu\left(\frac{5}{9}x_1 - \frac{4}{9}x_2 - \frac{2}{9}x_3, -\frac{4}{9}x_1 + \frac{5}{9}x_2 - \frac{2}{9}x_3, -\frac{2}{9}x_1 - \frac{2}{9}x_2 + \frac{8}{9}x_3\right), \right.$$
$$\left. \nu\left(\frac{5}{9}x_1 + \frac{4}{9}x_2 - \frac{2}{9}x_3, \frac{4}{9}x_1 + \frac{5}{9}x_2 + \frac{2}{9}x_3, -\frac{2}{9}x_1 + \frac{2}{9}x_2 + \frac{8}{9}x_3\right)\right],$$

with $\mu = (\frac{5}{9}x_1^2 + \frac{5}{9}x_2^2 + \frac{8}{9}x_3^2 - \frac{8}{9}x_1x_2 - \frac{4}{9}x_2x_3 - \frac{4}{9}x_1x_3)^{-\frac{1}{2}}$ and $\nu = (\frac{5}{9}x_1^2 + \frac{5}{9}x_2^2 + \frac{8}{9}x_3^2 + \frac{8}{9}x_1x_2 + \frac{4}{9}x_2x_3 - \frac{4}{9}x_1x_3)^{-\frac{1}{2}}$.

Case 4: $\det(n_4, e_{41}, x) \leq 0$ *and* $d(f(E_3), x) > d(f(E_4), x)$, i.e., $-2x_1 - x_2 + 2x_3 \geq 0$ *and* $11 + 3\sqrt{17}x_1 + 4x_2 + 4x_3 > 0$. The fact that $x \notin \text{epi } f$ implies $-\frac{1}{2}x_1 - x_2 - x_3 > 0$. Then $\text{proj}_{\text{epi } f}(x) \in f(E_4)$ and

$$\nabla d_{\text{epi } f}(x) = \frac{2}{3}\left(-\frac{1}{2}, -1, -1\right).$$

Case $4 \cap 1$: $\det(n_4, e_{41}, x) \geq 0$, $\det(n_1, e_{41}, x) \geq 0$ *and* $\langle e_{41}, x\rangle \geq 0$, i.e., $-2x_1 - x_2 + 2x_3 \leq 0$, $2x_1 + 7x_2 + 10x_3 \leq 0$ *and* $x_1 - x_2 + \frac{1}{2}x_3 \geq 0$. Then $\text{proj}_{\text{epi } f}(x) \in \mathbb{R}_+ e_{41}$ and

$$\nabla d_{\text{epi } f}(x) = \nu\left(\frac{5}{9}x_1 + \frac{4}{9}x_2 - \frac{2}{9}x_3, \frac{4}{9}x_1 + \frac{5}{9}x_2 + \frac{2}{9}x_3, -\frac{2}{9}x_1 + \frac{2}{9}x_2 + \frac{8}{9}x_3\right)$$

with $\nu = (\frac{5}{9}x_1^2 + \frac{5}{9}x_2^2 + \frac{8}{9}x_3^2 + \frac{8}{9}x_1x_2 + \frac{4}{9}x_2x_3 - \frac{4}{9}x_1x_3)^{-\frac{1}{2}}$.

Case 0: $\langle e_{12}, x \rangle \leq 0$, $\langle e_{23}, x \rangle \leq 0$ *and* $\langle e_{41}, x \rangle \leq 0$. Then $\text{proj}_{\text{epi } f}(x) = 0$ and
$$\nabla d_{\text{epi } f}(x) = \frac{x}{|x|}.$$

By a direct computation, and in view of the equality (8.2), one obtains the two sets $\limsup_{x \to 0, d_{\text{epi } f}(x) > 0} \partial d_{\text{epi } f}(x)$ and $\widetilde{N}_{\text{epi } f}(0)$. $\qquad\square$

References

[1] Attouch, Hédy, Convergence de fonctions convexes, des sous-différentiels et semi-groupes associés. (French) *C. R. Acad. Sci. Paris* Sér. A-B 284 (1977), no. 10, A539–A542 MR0473929 (57:13587)

[2] Attouch, Hédy, *Variational convergence for functions and operators.* Applicable Mathematics Series. Pitman (Advanced Publishing Program), Boston, MA, 1984 MR0773850 (86f:49002)

[3] Attouch, Hédy and Beer, Gerald, On the convergence of subdifferentials of convex functions. *Arch. Math.* (Basel) 60 (1993), no. 4, 389–400 MR1206324 (94b:49018)

[4] Aubin, Jean-Pierre and Frankowska, Halina, *Set-valued analysis.* Systems and control: foundations and applications, Birkhäuser, 1990 MR1048347 (91d:49001)

[5] Benoist, Joël, Convergence de la dérivée de la régularisée Lasry-Lions. (French) [Convergence of the derivative of the Lasry-Lions regularization] *C. R. Acad. Sci. Paris Sér. I Math.* 315 (1992), no. 8, 941–944 MR1187139 (94c:46082)

[6] Benoist, Joël, Approximation and regularization of arbitrary sets in finite dimension, *Set Valued Anal.* 2 (1994), no. 1-2, pp. 95-115 MR1285823 (95e:49019)

[7] Bourbaki, Nicolas, Eléments de mathématique, Hermann, Paris

[8] Clarke, Frank H., Optimization and nonsmooth analysis. Canadian Mathematical Society Series of Monographs and Advanced Texts. A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1983. Reprint of the 1983 original. Université de Montréal, Centre de Recherches Mathèmatiques, Montréal, QC, 1989. Second edition. Classics in Applied Mathematics, 5. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1990 MR0709590 (85m:49002), MR1019086 (90g:49011), MR1058436 (91e:49001)

[9] Clarke, Frank. H.; Ledyaev, Yuri S.; Stern, Ronald J.; Wolenski, Peter R., *Nonsmooth analysis and control theory.* Graduate Texts in Mathematics, 178. Springer-Verlag, New York, 1998 MR1488695 (99a:49001)

[10] Clarke, Frank. H.; Stern, Ronald. J.; Wolenski, Peter R., Proximal smoothness and the lower-C^2 property. *J. Convex Anal.* 2 (1995), no. 1-2, 117–144 MR1363364 (96j:49014)

[11] Cornet, Bernard and Czarnecki, Marc-Olivier, Smooth representations of epi-Lipschitz subsets of $\mathbb{R}^n$. *Nonlinear Anal.* 37 (1999), no. 2, Ser. A: Theory Methods, 139–160 MR1689740 (2000a:49031)

[12] Cornet, Bernard and Czarnecki, Marc-Olivier, Smooth normal approximations of epi-Lipschitz subsets of R^n. *SIAM J. Control Optim.* 37 (1999), no. 3, 710–730 MR1675157 (2000b:49031)

[13] Cornet, Bernard and Czarnecki, Marc-Olivier, Existence of (generalized) equilibria: necessary and sufficient conditions. *Comm. Appl. Nonlinear Anal.* 7 (2000), no. 1, 21–53 MR1733400 (2000j:49029)

[14] Cornet, Bernard and Czarnecki, Marc-Olivier, Existence of generalized equilibria. *Nonlinear Anal.* 44 (2001), no. 5, Ser. A: Theory Methods, 555–574 MR1822230 (2001m:90131)

[15] Cornet, Bernard and Czarnecki, Marc-Olivier, Smoothing the distance function to a closed subset of $\mathbb{R}^n$ and applications, manuscript (1998)

[16] Czarnecki, Marc-Olivier, thèse de doctorat, Université Paris 1 Panthéon Sorbonne, 1996

[17] Federer, Herbert, Curvature measures. *Trans. Amer. Math. Soc.* 93 (1959) 418–491 MR0110078 (22:961)

[18] Georgiev, Pando Gr. and Zlateva, Nadia P., Reconstruction of the Clarke subdifferential by the Lasry-Lions regularizations. *J. Math. Anal. Appl.* 248 (2000), no. 2, 415–428 MR1775860 (2001h:49025)

[19] Ioffe, Alexander. D., Approximate subdifferentials and applications. I. The finite-dimensional theory. *Trans. Amer. Math. Soc.* 281 (1984), no. 1, 389–416 MR0719677 (84m:49029)

[20] Jourani, Abderrahim, Limit superior of subdifferentials of uniformly convergent functions. *Positivity* 3 (1999), no. 1, 33–47 MR1675463 (2000i:49022)

[21] Kruger, A. Ya. and Mordukhovich, Boris. S., Generalized normals and derivatives, and necessary optimality conditions in nondifferentiable programming (in Russian), *Depon. VINITI 494-80*, Moscow, 1980

[22] Kuratowski, Casimir, Topologie. Vol. II. (French) 2ème éd. Monografie Matematyczne, Tom XXI. Polskie Towarzystwo Matematyczne, Warszawa, 1952 MR0054232 (14:892a)

[23] Lasry, Jean-Michel. and Lions, Pierre-Lions, A remark on regularization in Hilbert spaces. *Israel J. Math.* 55 (1986), no. 3, 257–266 MR0876394 (88b:41020)

[24] Lebourg, Gérard, Sous-dérivabilité de fonctions semi-continues et convergence de dérivées: quelques résultats en densité. (French) *C. R. Acad. Sci. Paris Sér. A-B* 288 (1979), no. 16, A753–A755 MR0535804 (80d:49004)

[25] Levy, A. B.; Poliquin, René A; Thibault, Lionel, Partial extensions of Attouch's theorem with applications to proto-derivatives of subgradient mappings. *Trans. Amer. Math. Soc.* 347 (1995), no. 4, 1269–1294 MR1290725 (95k:49035)

[26] Mordukhovich, Boris. Sh., Maximum principle in the problem of time optimal response with nonsmooth constraints. *J. Appl. Math. Mech.* 40 (1976), no. 6, 960–969 (1977); translated from *Prikl. Mat. Meh.* 40 (1976), no. 6, 1014–1023 (Russian) MR0487669 (58:7284)

[27] Mordukhovich, Boris. S., *Approximation methods in problems of optimization and control*, Nauka, Moscow, 1988 MR0945143 (89m:49001)

[28] Plaskacz, Sławomir, On the solution sets for differential inclusions. *Boll. Un. Mat. Ital.* A (7) 6 (1992), no. 3, 387–394 MR1196133 (93m:34021)

[29] Poliquin, René A., An extension of Attouch's theorem and its application to second-order epi-differentiation of convexly composite functions. *Trans. Amer. Math. Soc.* 332 (1992), no. 2, 861–874 MR1145732 (93a:49013)

[30] Poliquin, René. A.; Rockafellar, R. Tyrrell., Prox-regular functions in variational analysis. *Trans. Amer. Math. Soc.* 348 (1996), no. 5, 1805–1838 MR1333397 (96h:49039)

[31] Poliquin, René. A.; Rockafellar, R. Tyrrell.; Thibault, Lionel, Local differentiability of distance functions. *Trans. Amer. Math. Soc.* 352 (2000), no. 11, 5231–5249 MR1694378 (2001b:49024)

[32] Rademacher, Hans, Uber partielle und totale Differenzierbarkeit von Funktionen mehrerer Variabeln und über die Transformation der Doppelintegrale, *Math. Ann.* 79 (1919), 340–359 MR1511935

[33] Rockafellar, R. Tyrrell, Clarke's tangent cones and the boundaries of closed sets in R^n. *Nonlinear Anal.* 3 (1979), no. 1, 145–154 (1978) MR0520481 (80d:49032)

[34] Rockafellar, R. Tyrrell, Proximal subgradients, marginal values, and augmented Lagrangians in nonconvex optimization. *Math. Oper. Res.* 6 (1981), no. 3, 424–436 MR0629642 (83m:90088)

[35] Rockafellar, R. Tyrrell and Wets, Roger J.-B., Variational analysis. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 317. Springer-Verlag, Berlin, 1998 MR1491362 (98m:49001)

[36] Subbotin, Andrei I., *Generalized solutions of first order PDEs*, Birkhäuser, Boston, 1995 MR1320507 (96b:49002)

[37] Warga, Jack, Fat homeomorphisms and unbounded derivate containers. *J. Math. Anal. Appl.* 81 (1981), no. 2, 545–560 MR0622836 (83f:58007)

[38] Zolezzi, Tullio, Convergence of generalized gradients. Set convergence in nonlinear analysis and optimization. *Set-Valued Anal.* 2 (1994), no. 1-2, 381–393 MR1285841 (95f:49016)

INSTITUT DE MATHEMATIQUES ET MODELISATION DE MONTPELLIER UMR 5030 CNRS, UNIVERSITÉ MONTPELLIER 2, PLACE EUGÈNE BATAILLON, 34095 MONTPELLIER CEDEX 5, FRANCE
E-mail address: `marco@math.univ-montp2.fr`

INSTITUT GIRARD DESARGUES, UNIVERSITÉ CLAUDE BERNARD LYON I, 43 BD DU 11 NOVEMBRE 1918, 69622 VILLEURBANNE CEDEX, FRANCE
E-mail address: `rifford@igd.univ-lyon1.fr`
Current address: Departement de Mathématiques (Bâtiment 425), Université de Paris-Sud, 91405 Orsay, France
E-mail address: `Ludovic.Rifford@math.u-psud.fr`