Extending Rademacher Theorem to Set-Valued Maps.

Aris Daniilidis & Marc Quincampoix

Abstract. Rademacher theorem asserts that Lipschitz continuous functions between Euclidean spaces are differentiable almost everywhere. In this work we extend this result to set-valued maps using an adequate notion of set-valued differentiability relating to convex processes. Our approach uses Rademacher theorem but also recovers it as a special case.

Keywords: Set-valued map, graphical derivative, convex process, Lipschitz continuity, Rademacher theorem.

AMS Classification: *Primary*: 26E25, 28A15 ; *Secondary*: 49J53, 49J52

1 Introduction

Rademacher theorem asserts that a (locally) Lipschitz continuous function $f : \mathbb{R}^d \to \mathbb{R}^\ell$ is differentiable almost everywhere (that is, the derivative $df(x)$ of f exists at every point $x \in \mathbb{R}^d \setminus \mathcal{N}$ where $\mathcal{N}$ has Lebesgue measure zero). In this paper we study extensions of this result to Lipschitz continuous set-valued maps F from $\mathbb{R}^d$ to $\mathbb{R}^\ell$ with nonempty convex compact values. We recall (see [4], [14], [20] *e.g.*) that a set-valued map F is called locally Lipschitz at $\bar{x}$ if for some $k \geq 0$ and all x, x' in a neighborhood of $\bar{x}$ it holds:

$$F(x) \subset F(x') + k\,|x - x'|\,\mathrm{B},$$

where B is the closed unit ball of $\mathbb{R}^\ell$. A special type of set-valued maps are the so-called *convex processes* (see [18], [19]). A set-valued map $L : \mathbb{R}^d \rightrightarrows \mathbb{R}^\ell$ is called a convex process, if its graph is a closed convex cone in $\mathbb{R}^d \times \mathbb{R}^\ell$, or equivalently, if $0 \in L(0)$ and for every $x, y \in \mathbb{R}^d$ and $\lambda > 0$ it holds:

$$L(x + y) \subset L(x) + L(y) \qquad \text{and} \qquad L(\lambda x) = \lambda\, L(x). \tag{1.1}$$

Convex processes are relevant in many applications in control and optimization (see [1], [3], [13] *e.g.* or the classical monographs [4], [14], [20]). They are considered as set-valued analogues of linear continuous operators; we refer the reader to [4, Chapter 2] for a discussion on this fact. We hereby use this notion to define differentiability of a set-valued map: indeed, we say that F is differentiable at $(x, y) \in \mathbb{R}^d \times \mathbb{R}^\ell$ with $y \in F(x)$ if the graphical derivative of F is a *convex process* (see forthcoming Definition 3.3).

Let us mention for completeness that an alternative notion of differentiability of set-valued maps, called H-differentiability, has been introduced in [15] and studied in [7], [10]. The notion of H-differentiability is based on the (less restrictive) positively homogeneous set-valued operators (which are not necessarily convex processes) and its main drawback is that it leads to conic approximations of F that need not be unique (see discussion in [7, Section 2.3]).

The manuscript is organized as follows: in the next section we fix our notations and give the main definitions and preliminary results. Section 3 is dedicated to differentiability of set-valued maps and its relation with the case of differentiability of functions. In Section 4 we obtain an extension of Rademacher theorem for the class of set-valued maps that are generated by a finite number of pointwise affinely independent Lipschitz functions (Proposition 4.1). This result recovers the classical Rademacher theorem as a special case. In Section 5 we obtain a general differentiability result (Theorem 5.5) for the class of isotropically Lipschitz functions (Definition 5.2) with convex compact values with nonempty interior and smooth boundary.

2 Preliminaries

Throughout this work, we denote by $|x|$ the Euclidean norm of an element $x \in \mathbb{R}^d$, and by $\mathrm{B}_X := \{x \in \mathbb{R}^d : |x| \leq 1\}$ the closed unit ball, centered at the origin of the normed space X. The index will often be omitted if there is no ambiguity about the space. We set $B_\delta(x) := x + \delta \mathrm{B}_X$ for the closed ball centered at x with radius $\delta > 0$. We also denote by $\langle p, x \rangle$ the scalar product of the elements $x, p \in \mathbb{R}^d$ and we identify $\mathbb{R}^d$ with its dual space (the space of all linear functions from $\mathbb{R}^d$ to $\mathbb{R}$) by means of the relation $p(x) := \langle p, x \rangle$, for all $x \in \mathbb{R}^d$. Given a sequence $(t_n)_{n \geq 1}$ of real numbers, we shall use the notation $t_n \searrow 0^+$ to indicate that $t_n > 0$ and $\lim_{n \to \infty} t_n = 0$.

Let us further denote by $\mathcal{L}_d$ the Lebesgue measure of $\mathbb{R}^d$. A set $A \subset \mathbb{R}^d$ is said to have a full measure in $\mathbb{R}^d$ if its complement is contained in a null set, that is, $\mathbb{R}^d \diagdown A \subset \mathcal{N}$ with $\mathcal{L}_d(\mathcal{N}) = 0$. If a property holds for all points of a full-measure set A, then we say that the property holds *almost everywhere* (in short, *a.e.*) or *for almost all* $x \in \mathbb{R}^d$ (in short, $\forall_{a.e.} x \in \mathbb{R}^d$) and omit the explicit reference to the set A.

A function $f : \mathbb{R}^d \longrightarrow \mathbb{R}^\ell$ is called locally Lipschitz continuous at $\bar{x} \in \mathbb{R}^d$ if there exists $k \geq 0$ such that for all x, x' in some neighborhood of $\bar{x}$ it holds:

$$f(x) - f(x') \leq k\,|x - x'|. \tag{2.1}$$

We say that f is locally Lipschitz, if it is locally Lipschitz at every point of its domain. In particular, if (2.1) holds for all $x, x' \in \mathbb{R}^d$, then we say that f is k-Lipschitz continuous. We use the notation $df(x)$ to denote the (Fréchet) derivative of f at x (whenever it exists). Notice that $df(x)$ is a linear function from $\mathbb{R}^d$ to $\mathbb{R}^\ell$ (linear approximation of f aroung x) and its graph is an affine subspace of $\mathbb{R}^d \times \mathbb{R}^\ell$ passing through $(x, f(x))$.
In this classical setting, let us recall the following well-known result (see [9], [8] *e.g.*).

(*Rademacher theorem*) Every locally Lipschitz function $f : \mathbb{R}^d \to \mathbb{R}^\ell$ is *a.e.* differentiable.

Set-valued maps are tightly related to nonsmooth phenomena and have been widely used in problems in calculus of variations, control and optimization (see [10], [11], [12], [16], [17] *e.g.*). For a set-valued map F from $\mathbb{R}^d$ to $\mathbb{R}^\ell$, we will use the notation $F : \mathbb{R}^d \rightrightarrows \mathbb{R}^\ell$. The graph of F is defined as follows:

$$\mathrm{gph}(F) := \{(x, y) \in \mathbb{R}^d \times \mathbb{R}^\ell : \ y \in F(x)\}.$$

A set-valued map F is called locally Lipschitz at $\bar{x}$ (or k-Lipschitz around $\bar{x}$) if for some $k \geq 0$ and all x, x' in a neighborhood of $\bar{x}$ it holds:

$$F(x) \subset F(x') + k\,|x - x'|\,\mathrm{B}. \tag{2.2}$$

Similarly to the case of (single-valued) functions, a set-valued map F is called locally Lipschitz, if it is locally Lipschitz at every $\bar{x} \in \mathbb{R}^d$. Moreover, if (2.2) holds for all $x, x' \in \mathbb{R}^d$, then we say that F is k-Lipschitz. It is easily seen that every locally Lipschitz map has closed graph.

Before we proceed further, let us recall that for a closed set $K \subset \mathbb{R}^m$ and $\bar{x} \in K$ the tangent (or contingent) cone is defined as follows (see [4], [14], [20] *e.g.*)

$$T_K(\bar{x}) := \{u \in \mathbb{R}^m : \exists (h_n) \searrow 0^+, \exists (u_n)_n \to u \text{ such that } \bar{x} + h_n u_n \in K, \forall n \geq 0\}. \tag{2.3}$$

It follows easily from the definition that $T_K(\bar{x})$ is a nonempty closed cone of $\mathbb{R}^m$.

A nonempty set K is called *convex* if the segments $[x, y] := \{tx + (1-t)y : t \in [0,1]\}$ are contained in K, for all $x, y \in K$. The set K is called *strictly convex* if, in addition, the open segments $(x, y) := \{tx + (1-t)y : t \in (0,1)\}$ are contained in the interior $\text{int}K$ of K, for every $x, y \in K$ with $x \neq y$. Therefore, assuming that a set K with at least two elements is strictly convex, implicitly yields that $\text{int}K \neq \emptyset$.

A function $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ is called convex, if its epigraph $\text{epi}f := \{(x, \beta) : \beta \geq f(x)\}$ is a convex subset of $\mathbb{R}^{d+1}$, or equivalently, if for all $x, y \in \text{dom}f := \{u \in \mathbb{R}^d : f(u) \in \mathbb{R}\}$ and $t \in [0,1]$ we have $f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$. The function f is called strictly convex, if whenever $x \neq y$ and $t \in (0,1)$ it holds $f(tx + (1-t)y) < tf(x) + (1-t)f(y)$ (that is, the graph of f does not contain segments). Given $x \in \text{dom}f$, we define the convex subdifferential $\partial f(x)$ as follows:

$$\partial f(x) := \{p \in \mathbb{R}^d : f(y) \geq f(x) + \langle p, y - x\rangle, \text{ for all } y \in \mathbb{R}^d\}.$$

It is well-known that if the convex function f is differentiable at x, then $\partial f(x) = \{df(x)\}$ (using the identification of $\mathbb{R}^d$ with its dual space).

We finally recall that vectors $\{V_1, \ldots, V_N\} \subset \mathbb{R}^d$ are called affinely independent, if for every $\{\mu_i\}_{i=1}^N \subset \mathbb{R}$ it holds:

$$\sum_{i=1}^N \mu_i = 0 \quad \text{and} \quad \sum_{i=1}^N \mu_i V_i = 0 \qquad \Longrightarrow \qquad \mu_1 = \ldots = \mu_N = 0.$$

Notice that this is equivalent to the fact that the family $\{V_i - V_N\}_{i=1}^{N-1}$ is linearly independent and yields that $N \leq d+1$.

3 Graphical derivative and differentiability of set-valued maps

It is natural to apply the notion of tangent cone to the graph of a set-valued map in order to obtain a linearization of a set-valued map around a point of its graph. This is achieved by setting $K = \text{gph}(F)$ (assuming closed) and taking the tangent cone at a point $(x, y) \in \text{gph}(F)$. We formalize this in the following definition (see [4]).

Definition 3.1 (graphical derivative). Let $F : \mathbb{R}^d \rightrightarrows \mathbb{R}^\ell$ be a set-valued map with closed graph. The graphical derivative of F at $(\bar{x}, \bar{y}) \in \text{gph}(F)$ is a set-valued map $DF(\bar{x}, \bar{y}) : \mathbb{R}^d \rightrightarrows \mathbb{R}^\ell$ whose graph is the tangent cone at $(\bar{x}, \bar{y})$ of the graph of F, that is:

$$\text{gph}(DF(\bar{x}, \bar{y})) := T_{\text{gph}(F)}(\bar{x}, \bar{y}).$$

In view of the above definition, taking into account (2.3) we deduce easily that for every $u \in \mathbb{R}^d$:

$$DF(\bar{x}, \bar{y})(u) := \{v \in \mathbb{R}^\ell : \exists (h_n) \searrow 0^+, \exists (u_n, v_n)_n \to (u, v) : \bar{y} + h_n v_n \in F(\bar{x} + h_n u_n), \forall n\} \tag{3.1}$$

In case of a locally Lipschitz (set-valued) map, the graphical derivative has two interesting properties:

Lemma 3.2 (graphical derivative of a Lipschitz map)**.** *Assume that $F:\mathbb{R}^d \rightrightarrows \mathbb{R}^\ell$ is k-Lipschitz at $\bar{x}$ and pick any $\bar{y}\in F(\bar{x})$.*

(i)*. The graphical derivative $DF(\bar{x},\bar{y})$ admits a simplified formula:*

$$DF(\bar{x},\bar{y})(u) := \{v\in\mathbb{R}^\ell:\ \exists(h_n)\searrow 0^+,\ \exists(w_n)_n\to v:\ \bar{y}+h_nw_n\in F(\bar{x}+h_nu),\ \forall n\}. \tag{3.2}$$

(ii)*. For any $u\in\mathbb{R}^d$, we have*

$$DF(\bar{x},\bar{y})(u)\,\cap\, k\,|u|\,\mathrm{B}\neq\emptyset.$$

Proof. Let us prove (i). A mere comparison with (3.1) shows that the right-hand side of (3.2) is contained in $DF(\bar{x},\bar{y})(u)$. To prove the converse inclusion take $(u,v)\in \mathrm{gph}\,(DF(\bar{x},\bar{y}))$. Then, according to (3.1), there exist sequences $\exists(h_n)\searrow 0^+$, $(u_n)_n\to u$ and $(v_n)_n\to v$ such that $\bar{y}+h_nv_n\in F(\bar{x}+h_nu_n)$ for all $n\geq 1$. Since F is k-Lipschitz, we deduce from (2.2) that $F(\bar{x}+h_nu_n)\subset F(\bar{x}+h_nu)+kh_n|u-u_n|\mathrm{B}$. In particular, there exists a sequence $(b_n)_n\subset\mathrm{B}$ such that for all $n\geq 1$ we have:

$$\bar{y}+h_n(v_n+k|u_n-u|b_n)\in F(\bar{x}+h_nu).$$

Set $w_n:=v_n+k|u_n-u|b_n$ so that $\bar{y}+h_nw_n\in F(\bar{x}+h_nu)$ for all $n\geq 1$. Noticing that $(w_n)_n$ converges to v as $n\to\infty$ yields (3.2).

Let us now prove (ii). Fix $u\in\mathbb{R}^d$ and $(h_n)\searrow 0^+$. By Lipschitz continuity of F we deduce:

$$\bar{y}\in F(\bar{x})\subset F(\bar{x}+h_nu)+\,k\,h_n\,|u|\,\mathrm{B}.$$

Therefore, there exists a sequence $(b_n)_n\subset\mathrm{B}$ such that $\bar{y}+\,k\,h_n\,|u|\,b_n\in F(\bar{x}+h_nu)$ for all $n\geq 1$. Since B is compact, passing eventually to a subsequence, we may assume that $b_n\to b\in\mathrm{B}$ and $w_n:=k\,|u|\,b_n\to k\,|u|\,b:=v$. Since $\bar{y}+h_nw_n\in F(\bar{x}+h_nu)$ we deduce directly from (3.2) that $v\in k\,|u|\,\mathrm{B}\cap DF(\bar{x},\bar{y})(u)$.

The proof is complete. □

Now we are ready to define an appropriate notion of differentiability of a set-valued map.

Definition 3.3 (Differentiability of a set-valued map)**.** A set-valued map $F:\mathbb{R}^d\rightrightarrows\mathbb{R}^\ell$ with closed graph is said to be differentiable at $(\bar{x},\bar{y})\in\mathrm{gph}F$ if its graphical derivative $DF(\bar{x},\bar{y})$ is a convex process (see (1.1)).

In other words, F is differentiable at $(\bar{x},\bar{y})\in\mathrm{gph}F$ if the closed cone $T_{\mathrm{gph}(F)}(\bar{x},\bar{y})$ is also convex, which means that $DF(\bar{x},\bar{y})$ is a set-valued map whose graph is a closed convex cone.

Definition 3.3 follows the spirit of [5], but it is purely geometrical and guarantees that the derivative is unique (whenever it exists). Moreover, it generalizes the classical differentiability of Lipschitz functions, as forthcoming Proposition 3.5 will underline. Let us first state the following lemma.

Lemma 3.4. *Assume that for some $f:\mathbb{R}^d\longrightarrow\mathbb{R}^\ell$ and $\bar{x}\in\mathbb{R}^d$ we have:*

$$k_{\bar{x}}:=\limsup_{x\to\bar{x}}\frac{|f(x)-f(\bar{x})|}{|x-\bar{x}|}<+\infty \tag{3.3}$$

Then setting $F(x)=\{f(x)\}$, for all $x\in\mathbb{R}^d$ (ie. f is identified to a set-valued map F), we have:

(i). $DF(\bar{x},f(\bar{x}))(u)\neq\emptyset$, for every $u\in\mathbb{R}^d$; and

(ii) for all $(u,v)\in\mathrm{gph}(DF(\bar{x},f(\bar{x})))$ it holds: $|v|\ \leq k_{\bar{x}}\,|u|$.

Proof. Let $u \in \mathbb{R}^d$ and $(h_n) \searrow 0^+$. We deduce from (3.3) that for n sufficiently large:

$$w_n := \frac{f(\bar{x}+h_n u) - f(\bar{x})}{h_n} \in (1+k_{\bar{x}})\mathrm{B}.$$

Since the sequence $(w_n)_n$ is bounded, passing to a subsequence we may assume that it converges to some $v \in \mathbb{R}^\ell$. Since $f(\bar{x}) + h_n w_n = f(\bar{x}+h_n u) \in F(\bar{x}+h_n u)$ we conclude from (3.2) that $v \in DF(\bar{x}, f(\bar{x}))(u)$ and (i) holds.

To prove (ii), fix $(u,v) \in \mathrm{gph}(DF(\bar{x}, f(\bar{x})))$ and consider the associated sequences $(h_n) \searrow 0^+$ and $(u_n, v_n)_n \to (u,v)$ such that $f(\bar{x}) + h_n v_n = f(\bar{x}+h_n u_n)$. In view of (3.3), passing possibly to subsequences, we deduce that there exists a sequence of positive numbers $\varepsilon_n \to 0^+$ such that

$$|v_n| = \frac{|f(\bar{x}+h_n u_n) - f(\bar{x})|}{h_n} \leq (k_{\bar{x}} + \varepsilon_n)|u_n|, \ \forall n \geq 0.$$

Passing to the limit we obtain $|v| \leq k_{\bar{x}}\,|u|$ as desired. □

Notice that every Lipschitz function at $\bar{x}$ satisfies (3.3). Notice that in Lemma 3.4, f is considered both as a (single-valued) function and as a set-valued map with singleton values, and consequently we can define the usual (Fréchet) derivative and the graphical derivative. The following proposition clarifies the relation between these two objects.

Proposition 3.5 (compatibility)**.** *Let $f : \mathbb{R}^d \longrightarrow \mathbb{R}^\ell$ satisfy* (3.3) *at $\bar{x} \in \mathbb{R}^n$. Then f is differentiable at $\bar{x}$ if and only if the map $F(x) = \{f(x)\}$ is differentiable at $(\bar{x}, f(\bar{x}))$. In this case, it holds*

$$\mathrm{gph}(DF(\bar{x}, f(\bar{x}))) = \mathrm{gph}(df(\bar{x})).$$

Proof. Let us first assume that f is differentiable at $\bar{x}$, that is,

$$\lim_{|z|\to 0} \frac{|f(\bar{x}+z) - f(\bar{x}) - df(\bar{x})(z)|}{|z|} = 0. \tag{3.4}$$

Then it follows easily from (3.2) that for any $u \in \mathbb{R}^d$ it holds $df(\bar{x})(u) \in DF(\bar{x}, \bar{y})(u)$, therefore $\mathrm{gph}(df(\bar{x})) \subset \mathrm{gph}(DF(\bar{x}, f(\bar{x})))$. Let now $v \in DF(\bar{x}, f(\bar{x}))(u)$. It follows from (3.1) that there exist sequences $(h_n) \searrow 0^+$ and $(u_n, v_n)_n \to (u,v)$ such that $f(\bar{x}) + h_n v_n \in F(\bar{x}+h_n u_n) = \{f(\bar{x}+h_n u_n)\}$, for all $n \geq 1$. It follows readily that

$$v_n = \frac{f(\bar{x}+h_n u_n) - f(\bar{x})}{h_n}.$$

In view of (3.4), taking the limit as $n \to \infty$ we obtain $v = df(\bar{x})(u)$, yielding the equality

$$\mathrm{gph}(DF(\bar{x}, f(\bar{x}))) = \mathrm{gph}\,(df(\bar{x}))\,.$$

The above ensures the differentiability of F at $(\bar{x}, f(\bar{x}))$, since $\mathrm{gph}\,(df(\bar{x}))$ is a linear subspace of $\mathbb{R}^d \times \mathbb{R}^\ell$ and consequently a closed convex cone.

Let us now assume that F is differentiable at $(\bar{x}, f(\bar{x}))$, that is, $DF(\bar{x}, f(\bar{x}))$ is a convex process and consequently $\mathrm{gph}\,(DF(\bar{x}, f(\bar{x})))$ is a closed convex cone. We are going to prove successively that $\mathrm{gph}\,(DF(\bar{x}, f(\bar{x})))$ is the graph of some linear function L from $\mathbb{R}^d$ to $\mathbb{R}^\ell$ and then that L is the Fréchet derivative of f at $\bar{x}$.

To this end, fix $(u, v) \in \mathrm{gph}(DF(\bar{x}, f(\bar{x})))$. We claim that

$$DF(\bar{x}, f(\bar{x}))(-u) = \{-v\}. \tag{3.5}$$

In view of Lemma 3.4(ii), this is clear when $u = 0$, therefore, we may assume $u \neq 0$. Let $w \in DF(\bar{x}, f(\bar{x}))(-u)$ and notice that (u, v) and $(-u, w)$ are two points of the convex set $\mathrm{gph}\,(DF(\bar{x}, f(\bar{x})))$. It follows that

$$\frac{1}{2}[(u, v) + (-u, w)] = (0, \frac{1}{2}(v + w) \in DF(\bar{x}, f(\bar{x})),$$

which again in view of Lemma 3.4(ii) yields $v + w = 0$ and (3.5) follows. Moreover, since this is true for any $u \in \mathbb{R}^d$, we conclude that the set $DF(\bar{x}, f(\bar{x}))(u)$ reduces to the singleton $\{v\}$. We claim that for every u the directional derivative

$$\lim_{t \to 0} \frac{f(\bar{x} + tu) - f(\bar{x})}{t} \tag{3.6}$$

exists and is equal to v. Indeed let $t_n \searrow 0^+$ and observe that for n sufficiently large the sequence

$$v_n := \frac{f(\bar{x} + t_n u) - f(\bar{x})}{t_n}$$

is contained in the ball $B(0,\ 1 + k_{\bar{x}})$ therefore it has accumulation points, which all belong to $DF(\bar{x}, f(\bar{x}))(u) = \{v\}$. This proves that $v_n \to v$ and that the directional limit in (3.6) exists when $t \to 0^+$ and is equal to $DF(\bar{x}, f(\bar{x}))(u)$. Changing u into $-u$ we obtain that the corresponding directional limit also exists and it is equal to $-Df(\bar{x}, f(\bar{x}))(-u)$, which in turn, in view of (3.5), is equal to $DF(\bar{x}, f(\bar{x}))(u)$. It remains to show that the linear function

$$\mathbb{R}^d \ni u \mapsto L(u) := v \in \mathbb{R}^d \qquad \text{where } DF(\bar{x}, f(\bar{x}))(u) = \{v\},$$

is the derivative of f at $\bar{x}$ namely:

$$\lim_{|z| \to 0^+} \frac{|f(\bar{x} + z) - f(\bar{x}) - L(z)|}{|z|} = 0.$$

Let us assume, towards a contradiction, that this is not the case. Then there exist $\alpha > 0$ and a sequence $(z_n) \to 0$, with $z_n \neq 0$ for all $n \geq 1$ such that

$$\frac{|f(\bar{x} + z_n) - f(\bar{x}) - L(z_n)|}{|z_n|} \geq \alpha,\ \forall n \geq 0. \tag{3.7}$$

The sequence $u_n := z_n/|z_n|$ is bounded and converges (up to a subsequence) to some $u \in \mathbb{R}^d$. Moreover, using (3.3) for n sufficiently large we obtain

$$\begin{aligned}
\left| \frac{f(\bar{x} + z_n) - f(\bar{x}) - L(z_n)}{z_n} \right| &= \left| \frac{f(\bar{x} + |z_n| u_n) - f(\bar{x})}{|z_n|} - L(u_n) \right| \leq \left| \frac{f(\bar{x} + |z_n| u) - f(\bar{x})}{|z_n|} - L(u) \right| \\
&+ \frac{|f(\bar{x} + |z_n| u_n) - f(\bar{x} + |z_n| u)|}{|z_n|} + |L(u) - L(u_n)| \\
&\leq \left| \frac{f(\bar{x} + |z_n| u) - f(\bar{x})}{|z_n|} - L(u) \right| + (1 + k_{\bar{x}})\, |u_n - u| + \|L\|\, |u - u_n|
\end{aligned}$$

where $\|L\|$ denotes the operator norm of the linear map L. Passing to the limsup to the above inequalities and taking into account that

$$u_n \to u \qquad \text{and} \qquad \lim_n \frac{f(\bar{x} + |z_n|u) - f(\bar{x})}{|z_n|} = L(u)$$

we obtain

$$\limsup_n \frac{|f(\bar{x} + z_n) - f(\bar{x}) - L(z_n)|}{|z_n|} \leq 0$$

which contradicts (3.7). Therefore, $L(u) = df(\bar{x})(u)$ and $DF(\bar{x}, f(\bar{x}))(u) = \{df(\bar{x})(u)\}$.
The proof is complete. □

4 Set-valued maps generated by a finite family

In this section we consider a particular type of set-valued maps $F : \mathbb{R}^d \rightrightarrows \mathbb{R}^\ell$, namely those generated by a finite family of locally Lipschitz functions, that is,

$$F(x) = \text{conv}\{f_1(x), \ldots, f_N(x)\}$$

where $f_i : \mathbb{R}^d \to \mathbb{R}^\ell$ is a (locally) Lipschitz function, for all $i \in \{1, \ldots, N\}$. It is easy to see that $F : \mathbb{R}^d \rightrightarrows \mathbb{R}^\ell$ is Lipschitz continuous (as set-valued function) with nonempty convex compact values. We shall further assume that for every $x \in \mathbb{R}^d$ the family $\{f_1(x), \ldots, f_N(x)\}$ is affinely independent. This assumption yields in particular that $N \leq \ell + 1$. Moreover, setting

$$\Delta := \left\{ (\lambda_1, \ldots, \lambda_N) \in [0,1]^N : \sum_{j=1}^N \lambda_j = 1 \right\},$$

we have a unique representation of every $y \in F(x)$, that is, there exist unique $(\lambda_1, \ldots, \lambda_N) \in \Delta$ such that $y = \sum_{j=1}^N \lambda_j f_j(x)$.

We are now ready to state the following generalization of Rademacher theorem for this type of set-valued maps.

Proposition 4.1 (generalized Rademacher theorem)**.** *Let $f_i : \mathbb{R}^d \to \mathbb{R}^\ell$, $i \in \{1, \ldots, N\}$ be a family of Lipschitz continuous functions for which $\{f_1(x), \ldots, f_N(x)\}$ are affinely independent for every $x \in \mathbb{R}^d$. Then the set-valued map $F : \mathbb{R}^d \rightrightarrows \mathbb{R}^\ell$ given by*

$$F(x) = \text{conv}\,\{f_1(x), \ldots, f_N(x)\}, \quad x \in \mathbb{R}^d,$$

is differentiable at $(\bar{x}, \bar{y})$ for a.e. *$\bar{x} \in \mathbb{R}^d$ and all $\bar{y} \in F(x)$. In particular, for every $u \in \mathbb{R}^d$*

$$DF(\bar{x}, \bar{y})(u) \;=\; \sum_{j=1}^N \bar{\lambda}_j \, df_j(\bar{x})(u) \;+\; T_{F(\bar{x})}(\bar{y}),$$

where $(\bar{\lambda}_1, \ldots, \bar{\lambda}_N) \in \Delta$ are such that

$$\bar{y} = \sum_{j=1}^N \bar{\lambda}_j \, f_j(\bar{x}). \tag{4.1}$$

Proof. Let us denote by $\mathcal{D} \subset \mathbb{R}^d$ the set of points of common differentiability of the functions f_i, $i \in \{1, \ldots, N\}$ (by Rademacher theorem this set is of full measure) and let us fix $\bar{x} \in \mathcal{D}$, $\bar{y} \in F(\bar{x})$ and $u \in \mathbb{R}^d$. Let further $w \in DF(\bar{x}, \bar{y})(u)$. Then there exist $(w_n)_n \to w$ and $(t_n) \searrow 0^+$ such that

$$\bar{y} + t_n w_n \in F(\bar{x} + t_n u).$$

Then for every $n \geq 1$ there exists $(\lambda_1^n, \ldots, \lambda_N^n) \in \Delta$ such that

$$\bar{y} + t_n w_n = \sum_{j=1}^{N} \lambda_j^n \, f_j(\bar{x} + t_n u). \tag{4.2}$$

Thanks to the compactness of Δ, passing to a subsequence, we may assume that

$$(\lambda_1^n, \ldots, \lambda_N^n) \to (\lambda_1^*, \ldots, \lambda_N^*), \quad \text{as } n \to \infty.$$

For every $j \in \{1, \ldots, N\}$, the first-order Taylor series of f_j at $\bar{x}$ yields:

$$f_j(\bar{x} + t_n u) = f_j(\bar{x}) \, + \, t_n \, df_j(\bar{x})(u) \, + o_j(t_n u), \qquad \text{where } \lim_{|v| \to 0} \frac{o_j(v)}{|v|} = 0. \tag{4.3}$$

Replacing (4.1), (4.3) into (4.2) we obtain:

$$t_n w_n \, = \sum_{j=1}^{N} \left(\lambda_j^n - \bar{\lambda}_j\right) f_j(\bar{x}) \, + \, t_n \sum_{j=1}^{N} \lambda_j^n \, df_j(\bar{x})(u) \, + \, t_n \, ||u|| \sum_{n j=1}^{N} \lambda_j^n \left(\frac{o_j(t_n u)}{||t_n u||} \right), \tag{4.4}$$

and passing to the limit as $n \to \infty$ we deduce:

$$\sum_{j=1}^{N} \left(\lambda_j^* - \bar{\lambda}_j\right) f_j(\bar{x}) = 0, \qquad \text{yielding} \quad \lambda_j^* = \bar{\lambda}_j, \text{ for all } j \in \{1, \ldots, N\}.$$

Returning to (4.2), we deduce easily that:

$$w_n \, = \sum_{j=1}^{N} \lambda_j^n \left(\frac{f_j(\bar{x} + t_n u) - f_j(\bar{x})}{t_n} \right) \, + \, \frac{1}{t_n} \left(\sum_{j=1}^{N} \lambda_j^n \, f_j(\bar{x}) - \bar{y} \right). \tag{4.5}$$

Noticing that

$$\frac{1}{t_n} \left(\sum_{j=1}^{N} \lambda_j^n \, f_j(\bar{x}) - \bar{y} \right) \in \mathbb{R}_+ \left(F(\bar{x}) - \bar{y}\right) \subset T_{F(\bar{x})}(\bar{y})$$

and taking the limit in (4.5) as $n \to \infty$ we deduce that

$$w = \lim_{n \to \infty} \, w_n \, \in \, \sum_{j=1}^{N} \bar{\lambda}_j df_j(\bar{x})(u) + T_{F(\bar{x})}(\bar{y}).$$

For the reverse inclusion, it suffices to prove that for every $u \in \mathbb{R}^d$ we have:

$$\sum_{j=1}^{N} \bar{\lambda}_j df_j(\bar{x})(u) + \mathbb{R}_+ \left(F(\bar{x}) - \bar{y}\right) \subset DF(\bar{x}, \bar{y})(u).$$

To this end, fix $\mu \geq 0$, $(\lambda_1, \dots, \lambda_N) \in \Delta$ and define

$$v = \mu \left(\sum_{k=1}^{N} \lambda_k \, f_k(\bar{x}) - \bar{y} \right) \qquad \text{(arbitrary element of } \mathbb{R}_+ \, (F(\bar{x}) - \bar{y}) \,).$$

Let further $(t_n) \searrow 0^+$ and set

$$w_n := \frac{1}{t_n} \left((1 - \mu t_n) \sum_{j=1}^{N} \bar{\lambda}_j \, f_j(\bar{x} + t_n u) + \mu t_n \sum_{k=1}^{N} \lambda_k \, f_k(\bar{x} + t_n u) - \bar{y} \right). \tag{4.6}$$

Since $F(\bar{x} + t_n u)$ is convex, the above definition yields

$$\bar{y} + t_n w_n \in F(\bar{x} + t_n u), \quad \text{for all } n \text{ sufficiently large.} \tag{4.7}$$

Replacing $\bar{y}$ in (4.6) by its representation given in (4.1) we deduce:

$$w_n := (1 - \mu t_n) \sum_{j=1}^{N} \bar{\lambda}_j \left(\frac{f_j(\bar{x} + t_n u) - f_j(\bar{x})}{t_n} \right) + \mu \left(\sum_{k=1}^{N} \lambda_k \, f_k(\bar{x} + t_n u) - \bar{y} \right).$$

Taking the limit as $n \to \infty$ we obtain:

$$w := \lim_{n \to \infty} w_n = \sum_{j=1}^{N} \bar{\lambda}_j \, df_j(\bar{x})(u) + v,$$

and consequently, in view of (4.7)

$$w = \lim_{n \to \infty} w_n \in DF(\bar{x}, \bar{y})(u)$$

as desired. The proof is complete. □

Remark 4.2. Taking $N = 1$ in Proposition 4.1 we recover the classical Rademacher theorem.

5 A Rademacher Theorem for isotropically Lipschitz maps

In this section we deal with general set-valued maps for which we do not dispose a concrete description. Let us first notice that if a Lipschitz set-valued map $F : \mathbb{R}^d \rightrightarrows \mathbb{R}^\ell$ has closed convex values with nonempty interior, then differentiability of F at (x, y) for every $x \in \mathbb{R}^d$ and for *a.e.* $y \in F(x)$ (with respect to the $\mathcal{L}_\ell$-measure of $F(x)$) is straightforward. Indeed, for every $y \in \text{int}F(x)$ we have $\text{gph}\,(DF(x, y)) = \mathbb{R}^d \times \mathbb{R}^\ell$ and the boundary $\partial F(x)$ is a null subset of the convex set $F(x)$. In this case, the interesting part is clearly the behavior of F around boundary points $y \in \partial F(x)$ which is what we are going to investigate in the forthcoming Theorem 5.5.

Let $F : \mathbb{R}^d \rightrightarrows \mathbb{R}^\ell$ be a Lipschitz set-valued map with nonempty compact values. For every $p \in \mathbb{R}^\ell$ we set:

$$\sigma_{F(x)}(p) := \max_{y \in F(x)} \langle p, y \rangle \quad \text{and} \quad Y(x, p) := \{ y \in F(x) : \langle p, y \rangle = \sigma_{F(x)}(p) \}.$$

Notice that $Y(x,p)$ is a nonempty compact subset of $\mathbb{R}^\ell$ and that the function $p \mapsto \sigma_{F(x)}(p)$ is convex and Lipschitz continuous. It is also easy to verify that

$$(x,p) \mapsto \sigma_{F(x)}(p) \quad \text{is (locally) Lipschitz continuous,}$$

while the set-valued function $(x,p) \rightrightarrows Y(x,p)$ is continuous, but in general not Lipschitz continuous, as shows the following example.

Example 5.1. Consider the real valued functions

$$\phi_0(u) \;:=\; \begin{cases} u^2 - u, & \text{if} \quad u \le 0 \\ u^2, & \text{if} \quad u > 0, \end{cases}$$

and for $\tau > 0$:

$$\phi_\tau(u) \;:=\; \begin{cases} u^2 - u + \tau, & \text{if} \quad u \le 0 \\ -u\sqrt{\tau} + \tau, & \text{if} \quad u \in [0, \sqrt{\tau}], \\ u^2 - \tau, & \text{if} \quad u > \sqrt{\tau}, \end{cases}$$

We define $F : [-1,1] \rightrightarrows \mathbb{R}^2$ by

$$F(x) = \operatorname{epi} \phi_{|x|} \cap (\mathbb{R} \times [-2,2]).$$

Then F is Lipschitz continuous with compact convex values (with a slight modification we can obtain that $F(x)$ is strictly convex, but for $\bar{p} = (0,-1)$, the (single-valued) map $x \to Y(x,\bar{p}) := \{(\sqrt{|x|}, 0)\}$ is not Lipschitz continuous.

The above example motivates the following definition.

Definition 5.2 (isotropically Lipschitz map)**.** Let $F : \mathbb{R}^d \rightrightarrows \mathbb{R}^\ell$ be a set-valued map with nonempty convex compact values. We say that F is (locally) isotropically Lipschitz continuous, if (for every $\bar{x} \in \mathbb{R}^d$) there exists $k > 0$ such that for all $p \in \mathbb{R}^\ell$ and all x, x' (in a neighborhood of $\bar{x}$) it holds:

$$Y(x,p) \subset Y(x',p) + k\,|x - x'|\,\mathrm{B}. \tag{5.1}$$

The following result holds.

Proposition 5.3 (isotropically Lipschitz vs Lipschitz)**.** *Every (locally) isotropically Lipschitz map $F : \mathbb{R}^d \rightrightarrows \mathbb{R}^\ell$ is (locally) Lipschitz.*

Proof. Let F be k-isotropically Lipschitz around $\bar{x} \in \mathbb{R}^d$ and let us assume, towards a contradiction, that F is not k-Lipschitz there. Then there exist sequences $(x_n)_n$, $(x'_n)_n$ converging to $\bar{x}$ and $z_n \in F(x_n)$ such that $z_n \notin F(x'_n) + k\,|x_n - x'_n|\,\mathrm{B}$. Denoting by z'_n the projection of z_n onto the convex compact set $F(x'_n)$ and setting

$$p_n = \frac{z_n - z'_n}{|z_n - z'_n|}$$

we deduce that $z'_n \in Y(x'_n, p_n)$ and

$$\langle p_n, z_n - z'_n \rangle = |z_n - z'_n| > k\,|x_n - x'_n|.$$

Then for every $y_n \in Y(x_n, p_n) \subset F(x_n)$ one has:

$$|y_n - z_n'| \geq \langle p_n, y_n - z_n' \rangle \geq \langle p_n, z_n - z_n' \rangle = |z_n - z_n'| > k\,|x_n - x_n'|.$$

The above clearly yields

$$z_n' \notin Y(x_n, p_n) + k\,|x_n - x_n'|\,\mathrm{B}$$

contradicting (5.1). □

Remark 5.4. (i). If F is a (single-valued) function, then the notions of isotropic Lipschitz continuity and Lipschitz continuity coincide.
(ii). If F is isotropically Lipschitz and has strictly convex values, then for every $p \in S^{\ell-1}$ the set-valued map

$$x \rightrightarrows [Y(x, -p), Y(x, p)]$$

is Lipschitz. (Notice that strict convexity of $F(x)$ guarantees that both $Y(x, p)$ and $Y(x, -p)$ are singletons, so that $[Y(x, -p), Y(x, p)]$ is a closed segment in $\mathbb{R}^\ell$.)

We shall show that locally isotropically Lipschitz set-valued maps having *strictly convex* values with nonempty interior satisfy a Rademacher-type result. Before going further, let us make some comments about this last assumption.
(A). Fix $\bar{x} \in \mathbb{R}^d$ and set $K = F(\bar{x}) \subset \mathbb{R}^\ell$. Since the strictly convex compact set K has nonempty interior, its boundary ∂K is locally the graph of a (strictly) convex Lipschitz function $g : \mathbb{R}^{\ell-1} \mapsto \mathbb{R}$, that is, for every $y_0 \in \partial K$ there exists a (strictly) convex Lipschitz function g and $\varepsilon > 0$ such that

$$\partial K \cap B(y_0, \varepsilon) = \mathrm{gph}(g) \cap B(y_0, \varepsilon).$$

It follows that every $y \in \partial K \cap B(y_0, \varepsilon)$ can be represented in local coordinates as $y = (\xi, g(\xi))$ and that for *a.e.* $\xi \in \mathbb{R}^{\ell-1}$ (with respect to the Lebesgue measure of $\mathbb{R}^{\ell-1}$) the gradient $\nabla g(\xi)$ exists and therefore

$$N_K((\xi, g(\xi)) = \mathbb{R}_+(\nabla g(\xi), -1).$$

Since both the projection $\pi(y) = \pi(\xi, g(\xi)) = \xi \in \mathbb{R}^{\ell-1}$ and its inverse $\pi^{-1}(\xi) = (\xi, g(\xi))$ are Lipschitz continuous, we get a bi-Lipschitz bijection between $\partial K \cap B(y_0, \varepsilon)$ and $\mathrm{gph}(g) \cap B(y_0, \varepsilon)$ and conclude that these sets have the same null subsets (in the sense that they are identified through this bijection). We say that a subset $\mathcal{N}$ of $\partial F(\bar{x}))$ is null, if for every $y_0 \in \partial F(\bar{x})$ the set $\mathcal{N} \cap B(y_0, \varepsilon)$ is identified to a null subset of $\mathbb{R}^{\ell-1}$ by some bi-Lipschitz bijection. Similarly, we say that a subset $A \subset F(\bar{x})$ is of full measure in $\partial F(\bar{x})$, if $\partial F(\bar{x}) \setminus A$ is null. (Notice that although there is no canonical way to define a measure on $\partial F(\bar{x})$, null sets and sets of full measure are unambiguously determined in $\partial F(\bar{x})$.) In what follows, we say that a property holds *a.e.* on $\partial F(\bar{x})$ if it holds for all $y \in \partial F(\bar{x}) \setminus \mathcal{N}$ where $\mathcal{N}$ is a null subset of $\partial F(\bar{x})$.
Therefore for *a.e* $\bar{y} \in \partial F(\bar{x})$ we have $N_{\partial F(\bar{x})}(\bar{y}) = \mathbb{R}\bar{p}$ and $\bar{y} = Y(\bar{x}, \bar{p})$, which means that $\bar{y}$ is a point of smoothness of the boundary of $\partial F(\bar{x})$ and

$$\mathcal{M}_{\bar{x}} := \{\bar{y} \in \partial F(\bar{x}) : \ \dim N_{\partial F(\bar{x})}(\bar{y}) = 1\,\} \tag{5.2}$$

has a full measure in $\partial F(\bar{x})$.

(B). Thanks to the assumption of strict convexity of $F(\bar{x})$, the mapping

$$y_{\bar{x}} : p \in S^{\ell-1} \mapsto Y(\bar{x}, p)$$

is single-valued. It is also clearly onto $\partial F(\bar{x})$ (for every $y_0 \in \partial F(\bar{x})$, use Hahn-Banach theorem to separate $y_0 \in \partial F(\bar{x})$ from $\mathrm{int} F(\bar{x})$). Set

$$\mathbf{M}_{\bar{x}} := \{p \in S^{\ell-1} : \ y_{\bar{x}}^{-1}(y_{\bar{x}}(p)) = \{p\}\} = \{p \in S^{\ell-1} : \ y_{\bar{x}}(p) \in \mathcal{M}_{\bar{x}}\}.$$

Then $p \mapsto y_{\bar{x}}$ is a bijection from $\mathbf{M}_{\bar{x}}$ onto $\mathcal{M}_{\bar{x}}$ and $N_{\partial F(\bar{x})}(y_{\bar{x}}(p)) = \mathbb{R}_+ p$ for all $p \in \mathcal{M}_{\bar{x}}$. Notice further that $\mathcal{M}_{\bar{x}}$ is of full-measure in $\partial F(\bar{x})$ and $\mathbf{M}_{\bar{x}}$ is of full-measure in $S^{\ell-1}$. Moreover, for any full-measure subset P of $S^{\ell-1}$, the set $y_{\bar{x}}(S^{\ell-1} \setminus P)$ is null in $\partial F(\bar{x})$ and $\mathcal{M}_{\bar{x}} \setminus y_{\bar{x}}(S^{\ell-1} \setminus P)$ is of full-measure in $\partial F(\bar{x})$.

We are now ready to state a Rademacher type result for isotropically Lipschitz set-valued maps.

Theorem 5.5 (Rademacher result for isotropically Lipschitz maps.)**.** *Let $F : \mathbb{R}^d \rightrightarrows \mathbb{R}^\ell$ be a locally isotropically Lipschitz set-valued map with strictly convex values of nonempty interior. Then for* a.e. $\bar{x} \in \mathbb{R}^d$ *(with respect to the $\mathcal{L}_d$-Lebesgue measure of $\mathbb{R}^d$) and for* a.e. $\bar{y} \in \partial F(\bar{x})$*, the map F is differentiable at $(\bar{x}, \bar{y})$ and moreover,*

$$\mathrm{gph}(DF(\bar{x}, \bar{y})) \quad \textit{is a half-space.}$$

Proof. The mapping $(x, p) \mapsto \sigma_{F(x)}(p)$ is Lipschitz. By Rademacher theorem, there exists a set $\hat{D} \subset \mathbb{R}^d \times \mathbb{R}^\ell$ of full measure such that for any $(x, p) \in \hat{D}$ the derivative $\nabla \sigma_{F(x)}(p)$ exists. By Fubini theorem, and using the homogeneity of σ with respect to p we deduce the existence of a full measure subset $D \subset \mathbb{R}^d$ such that for any $x \in D$ the set

$$\{p \in S^{\ell-1} : \ (x, p) \in \hat{D}\} \text{ has a full measure in } S^{\ell-1}.$$

Fix $\bar{x} \in D$ and set

$$P_{\bar{x}} := \{p \in S^{\ell-1} : \ (x, p) \in \hat{D}\} \quad \text{(which has full-measure in } S^{\ell-1}\text{)}.$$

Recalling the definition of $\mathcal{M}_{\bar{x}}$ from (5.2) we set

$$\widehat{\mathcal{M}_{\bar{x}}} = \mathcal{M}_{\bar{x}} \setminus y_{\bar{x}}(S^{\ell-1} \setminus P_{\bar{x}}) \quad \text{(which has full-measure in } \partial F(\bar{x})\text{)}.$$

Fix $\bar{y} \in \widehat{\mathcal{M}_{\bar{x}}}$ and set $\bar{p} = y_{\bar{x}}^{-1}(\bar{y})$ (i.e. $Y(\bar{x}, \bar{p}) = \{\bar{y}\}$). Set

$$C_{\bar{x},\bar{y}} := \{ \, (u, w) \in \mathbb{R}^d \times \mathbb{R}^\ell : \ \nabla_x \sigma_{F(\bar{x})}(\bar{p})(u) \geq \langle \bar{p}, w \rangle \, \}$$

which is a closed half-space. We are going to prove that

$$T_{\mathrm{gph}(F)}(\bar{x}, \bar{y}) = C_{\bar{x},\bar{y}} \tag{5.3}$$

which immediately yields that the graphical derivative is convex (a half-space in $\mathbb{R}^d \times \mathbb{R}^\ell$) and consequently F is differentiable at $(\bar{x}, \bar{y})$, according to Definition 3.3.

To this end, take $(u, w) \in T_{\mathrm{gph}(F)}(\bar{x}, \bar{y})$. In view of Proposition 5.3 and (3.2), there exist sequences $t_n \searrow 0^+$ and $w_n \to w$ such that $\bar{y} \in F(\bar{x} + t_n u) - t_n w_n$.

For $\bar p = y_{\bar x}^{-1}(\bar y)$, we have $Y(\bar x,\bar p)=\{\bar y\}$ and

$$\sigma_{F(\bar x)}(\bar p) = \langle \bar p, \bar y\rangle \;\le\; \sigma_{F(\bar x+t_n u)}(\bar p) - t_n\langle \bar p, w_n\rangle, \quad \text{yielding} \quad \frac{\sigma_{F(\bar x+t_n w_n)}(\bar p)-\sigma_{F(\bar x)}(\bar p)}{t_n} \;\ge\; \langle w_n,\bar p\rangle.$$

Therefore, $\nabla_x\sigma_{F(\bar x)}(\bar p)(u)\ge\langle \bar p, w\rangle$ and the inclusion $T_{\mathrm{gph}(F)}(\bar x,\bar y)\subset C_{\bar x,\bar y}$ follows.
For the opposite inclusion, since $T_{\mathrm{gph}(F)}(\bar x,\bar y)$ is closed cone, it suffices to show that

$$\operatorname{int} C_{\bar x,\bar y}\subset T_{\mathrm{gph}(F)}(\bar x,\bar y).$$

To this end, let $(u,w)\in\mathbb{R}^d\times\mathbb{R}^\ell$ be such that

$$\nabla_x\sigma_{F(\bar x)}(\bar p)(u) \;>\; \langle \bar p, w\rangle. \tag{5.4}$$

We may also assume that $w\notin\mathbb{R}\bar p$ (if $w\in\mathbb{R}\bar p$ then we replace w by another vector $\tilde w\notin\mathbb{R}\bar p$ arbitrarily close to it). Let us assume by contradiction that $(u,w)\notin T_{\mathrm{gph}(F)}(\bar x,\bar y)$. Then by a standard argument (see [16] *e.g.*)

$$\exists\delta>0,\ \forall\tau\in(0,\delta):\quad (\bar y+\tau w+\tau\delta\mathrm{B})\cap F(\bar x+\tau u)=\emptyset. \tag{5.5}$$

Let us first prove that $u\neq 0$. Indeed if $u=0$, then $\nabla_x\sigma_{F(\bar x)}(\bar p)(u)=0>\langle\bar p,w\rangle$ and $(\bar y+\tau w+\tau\delta\mathrm{B})\cap F(\bar x)=\emptyset$ for all $\tau\in(0,\delta)$. Therefore $w\notin T_{F(\bar x)}(\bar y)$, which is the half-space $[\bar p\le 0]$ because $\partial F(\bar x)$ is smooth at $\bar y$ and $\operatorname{int}F(\bar x)\neq\emptyset$. This contradicts the assertion $\langle\bar p,w\rangle<0$.
Set

$$\left\{\begin{array}{l} e_1 := w-\langle\bar p,w\rangle\,\bar p\\ e_2 := -\bar p\end{array}\right. \qquad\text{and}\qquad Z:=\operatorname{span}\{e_1,e_2\}. \tag{5.6}$$

We deduce from (5.5) that

$$\Big[\{(\bar y+\tau w+\tau\delta\mathrm{B})-Y(\bar x+\tau u,\bar p)\}\cap Z\Big]\bigcap\Big[\{F(\bar x+\tau u)-Y(\bar x+\tau u,\bar p)\}\cap Z\Big]=\emptyset. \tag{5.7}$$

Let us further set:

$$K_\tau:=\{F(\bar x+\tau u)-Y(\bar x+\tau u)\}\cap Z \qquad\text{and}\qquad K_0:=\{F(\bar x)-Y(\bar x)\}\cap Z,$$

and define, for every $\tau\in[0,\delta)$ the function $\phi_\tau:\mathbb{R}\mapsto\mathbb{R}\cup\{+\infty\}$ as follows:

$$\phi_\tau(t):=\inf\{\beta\ge 0,\ te_1+\beta e_2\in K_\tau\},\quad t\in\mathbb{R}.$$

Note that ϕ_τ is convex with $\phi_\tau(0)=0$ and $\phi_\tau(t)>0$ when $t\neq 0$. Since F is Lipschitz, we have

$$(F(\bar x+\tau u)-Y(\bar x+\tau u,\bar p)) \underset{\tau\to 0}{\longrightarrow} F(\bar x)-Y(\bar x,\bar p) \qquad \text{(in the Painlevé-Kuratowski sense)}$$

and consequently $\mathrm{epi}(\phi_\tau)$ converges to $\mathrm{epi}(\phi_0)$ as $\tau\to 0$ in the Painlevé-Kuratowski sense. Therefore, by Attouch Theorem (see [6, Theorem 2.1] *e.g.*, or [2]) we get

$$\mathrm{gph}\ (\partial\phi_\tau) \underset{\tau\to 0}{\longrightarrow} \mathrm{gph}\ (\partial\phi_0) \quad \text{(in the Painlevé-Kuratowski sense).} \tag{5.8}$$

If $(-\alpha,\alpha)\in\operatorname{dom}\phi_0:=\{t\in\mathbb{R}:\ \phi_0(t)<+\infty\}$, for some $\alpha>0$, then for all $\alpha_1\in(0,\alpha)$ there exists $\tau_0>0$ such that $(-\alpha_1,\alpha_1)\subset\operatorname{dom}\phi_\tau$, for all $\tau\in(0,\tau_0)$ (thanks to the epiconvergence of ϕ_τ to ϕ_0). Let us set

$$a(\tau):=\langle e_1,\bar p-Y(\bar x+\tau u,\bar p)\rangle,\quad \tau\ge 0.$$

Recalling that $\sigma_{F(\bar{x}+\tau u)}(\bar{p}) = \langle \bar{p}, Y(\bar{x}+\tau u, \bar{p})\rangle$ and $\sigma_{F(\bar{x})}(\bar{p}) = \langle \bar{p}, \bar{y}\rangle$ and denoting by Λ the Lipschitz constant of $Y(\cdot, \bar{p})$, we deduce that

$$|a(\tau)| \leq \Lambda\tau|e_1|. \tag{5.9}$$

Translating formula (5.7) into the coordinates $\{e_1, e_2\}$ given in (5.6) we readily obtain that for all $b_1 \in \mathrm{B}$ the following inequality holds:

$$\langle -\bar{p}, \bar{y}\rangle - \tau\langle \bar{p}, w\rangle + \tau\delta + \langle \bar{p}, Y(\bar{x} + \tau u, \bar{p})\rangle \;<\; \phi_\tau\left(a(\tau) + \tau\langle e_1, w\rangle + \tau\delta b_1\right). \tag{5.10}$$

Observing that $\langle e_1, w\rangle = |w|^2 - \langle \bar{p}, w\rangle^2 > 0$ and recalling that $\sigma_{F(\bar{x})}(\bar{p}) = \langle \bar{p}, \bar{y}\rangle$, we deduce from (5.10) that

$$\phi_\tau(a(\tau) + \tau\langle e_1, w\rangle + \tau\delta b_1) > \tau\left(\frac{\sigma_{F(\bar{x}+\tau u)}(\bar{p}) - \sigma_{F(\bar{x})}(\bar{p})}{\tau} - \langle \bar{p}, w\rangle + \delta\right).$$

Therefore we deduce from (5.6) that for sufficiently small $\tau > 0$

$$\phi_\tau(X_\tau) \;>\; \frac{\delta}{2}\tau\,, \qquad \text{where} \quad X_\tau := a(\tau) + \tau\langle e_1, w\rangle + \tau\delta b_1 \;\neq\; 0.$$

By the mean value theorem, there exist $Z_\tau \in [0, X_\tau]$ and $\xi_\tau \in \partial\phi_\tau(Z_\tau)$ (subdifferential of ϕ_τ at Z_τ) such that

$$\phi_\tau(X_\tau) - \phi_\tau(0) = \xi_\tau X_\tau > \frac{\delta\tau}{2}.$$

Clearly when $\tau \to 0$ we have $X_\tau \to 0$ hence $Z_\tau \to 0$. Moreover since

$$\frac{1}{\tau}|X_\tau| \;\leq\; \Lambda|e_1| \;+\; |\langle e_1, w\rangle| \;+\; \delta$$

we have

$$|\xi_\tau| \;\geq\; \frac{\delta/2}{\Lambda|e_1| + |\langle e_1, w\rangle| + \delta} \;:=\; \Delta_0 \;>\; 0.$$

Claim. The set $\{\,|\xi_\tau| :\; \tau \in (0, \delta)\}$ is bounded.

Let $(-\alpha_1, \alpha_1) \subset \operatorname{dom}\phi_\tau$ for all $\tau \in [0, \delta)$. Since the convex function ϕ_0 is κ-Lipschitz on $[-\alpha_1, \alpha_1)$ for some constant $\kappa > 0$, we have $\partial\phi_0(t) \subset [-\kappa, \kappa]$ for all $t \in (-\alpha_1, \alpha_1)$. We now deduce from (5.8) that $\partial\phi_\tau(Z_\tau) \subset [-\kappa, \kappa]$, for all $\tau \in (0, \delta)$, and consequently, we may assume (passing to a subsequence) that $(\xi_\tau)_\tau$ converges to some ξ and the sequence $(Z_\tau, \xi_\tau) \in \operatorname{gph}(\partial\phi_\tau)$ converges to $(0, \xi)$ which belongs to $\operatorname{gph}(\partial\phi_0)$ by (5.8). We deduce that $|\xi| = \lim|\xi_\tau| > \delta_0 > 0$, which is a contradiction since ϕ_0 is differentiable at 0 with $\phi_0'(0) = 0$, because $\bar{y} = Y[\bar{x}, \bar{p}]$ is a smooth point of $\partial F(\bar{x})$. The proof is complete. □

Remark 5.6. (i). As already mentioned, differentiability of F at (x, y) is straightforward whenever $y \in \operatorname{int}F(x)$. Theorem 5.5 establishes differentiability also for *a.e.* $y \in \partial F(x)$.

(ii). The proof of Theorem 5.5 uses in a crucial way (namely, in (5.9)) that the set-valued map F is not only Lipschitz, but isotropically Lipschitz. Although this assumption seems indeed essential, we do not have any counterexample so far. This could be a topic for further investigations.

Acknowledgement Major parts of this work were done during mutual research visits at the University Bretagne-Occidentale (October 2020, November 2022) and at the TU Wien (April 2022). In each case the authors express their gratitude to the host institutes for hospitality.
The first author acknowledges support from the Austrian Science Fund (FWF, P-36344-N).

References

[1] F. Alvarez, R. Correa, P. Gajardo, Inner estimation of the eigenvalue set and exponential series solutions to differential inclusions, *J. Convex Anal.* **12** (2005), 1–11.

[2] H. Attouch, Convergence de fonctions convexes, des sous-différentiels et semi-groupes associés. *C. R. Acad. Sci. Paris Sér.* **284** (1977), 539–542.

[3] J.-P. Aubin, H. Frankowska, C. Olech, Controllability of Convex Processes, *SIAM J. Control Optim.* **24** (1986), 1192–1211.

[4] J.-P. Aubin, H. Frankowska, *Set-Valued Analysis*, (Birkhäuser, 1990).

[5] D. Aze, An inversion theorem for set-valued maps, *Bull. Austral. Math. Soc.* **37** (1988), 411–414.

[6] C. Combari, L. Thibault, On the graph convergence of subdifferentials of convex functions. *Proc. Amer. Math. Soc.* **126** (1998), 2231–2240.

[7] A. Daniilidis, J. Pang, Continuity and differentiability of set-valued maps revisited in the light of tame geometry, *J. Lond. Math. Soc.* **83** (2011), 637–658.

[8] L. C. Evans, R. F. Gariepy, *Measure theory and fine properties of functions* (Boca Raton CRC Press, 1992).

[9] H. Federer, *Geometric measure theory.* Die Grundlehren der mathematischen Wissenschaften **153** (Springer, 1969).

[10] M. Gaydu, M. Geoffroy, C. Jean-Alexis, An inverse mapping theorem for H-differentiable set-valued maps, *J. Math. Anal. Appl.* **421** (2015), 298–313.

[11] M. Gaydu, M. Geoffroy, Y. Marcelin, Prederivatives of convex set-valued maps and applications to set optimization problems, *J. Global Optim.* **64** (2016), 141–158.

[12] M. Gaydu, N. Silva, A general iterative procedure to solve generalized equations with differentiable multifunction, *J. Optim. Theory Appl.* **185** (2020), 207–222.

[13] A. Leizarowitz, Eigenvalues of convex processes and convergence properties of differential inclusions, *Set-Valued Anal.* **2** (1994), 505–527.

[14] B. Mordukhovich, *Variational analysis and generalized differentiation* (Vol. I and II), Grundlehren der mathematischen Wissenschaften **330** (Springer, 2006)

[15] J. Pang, Generalized differentiation with positively homogeneous maps: applications in set-valued analysis and metric regularity, *Math. Oper. Res.* **36** (2011), 377–397.

[16] M. Quincampoix, Differential inclusions and target problems, *SIAM J. Control Optim.* **30** (1992), 324–335.

[17] A. Uderzo, On differential properties of multifunctions defined implicitly by set-valued inclusions, *Pure Appl. Funct. Anal.* **6** (2021), 1509–1531.

[18] S. Robinson, Normed convex processes, *Trans. Amer. Math. Soc.* **174** (1972), 127–140.

[19] R. T. Rockafellar, Monotone processes of convex and concave type, *Memoirs of the American Mathematical Society* **77** (AMS, 1967).

[20] R. T. Rockafellar, R. Wets, *Variational analysis*, Grundlehren der Mathematischen Wissenschaften **317**, (Springer, 1998).

Aris DANIILIDIS

Institut für Stochastik und Wirtschaftsmathematik, VADOR E105-04
TU Wien, Wiedner Hauptstraße 8, A-1040 Wien

E-mail: `aris.daniilidis@tuwien.ac.at`
`https://www.arisdaniilidis.at/`

Research supported by the Austrian Science Fund (Grant FWF P-36344N)

Marc QUINCAMPOIX

Univ Brest, CNRS UMR 6205
Laboratoire de Mathématiques de Bretagne Atlantique
6, avenue Victor Le Gorgeu, F-29200 Brest, France

E-mail: `marc.quincampoix@univ-brest.fr`
`http://marc.quincampoix.perso.math.cnrs.fr/`