SOFIA UNIVERSITY ”ST. KLIMENT OHRIDSKI”

FACULTY OF MATHEMATICS AND INFORMATICS

VARIATIONAL ANALYSIS: METHODS AND APPLICATIONS

Nadia Peycheva Zlateva, PhD

DISSERTATION FOR DOCTOR OF SCIENCES DEGREE IN MATHEMATICS

S O F I A
2 0 1 7

Dedicated to the bright memory of
MD Lidia Todorova Grigorova, my mother
and to MD Peycho Zlatev Zlatev, my father

Contents

Preface **3**

1 Primal lower-nice functions and prox-regular sets **5**

1.1 On primal lower-nice property . . . 7
1.2 Subdifferential characterization of primal lower-nice functions on smooth Banach spaces . . . 14
1.3 Characterizations of prox-regular sets in uniformly convex Banach spaces . . . 19
1.3.1 Prox-regular sets . . . 27
1.3.2 Local Moreau envelopes . . . 33
1.3.3 Characterizations of prox-regular sets . . . 40
1.3.4 Characterizations of uniformly prox-regular sets . . . 48
1.4 Prox-regular sets and epigraphs in uniformly convex Banach spaces: various regularities and other properties . . . 57
1.4.1 Normal and tangential regularity properties of prox-regular sets 66
1.4.2 Epigraphs of J-primal lower regular functions . . . 72
1.4.3 Prox-regularity and N-hyporegularity . . . 78
1.4.4 Comparison of normal cones . . . 82
1.4.5 Preservation of hyporegularity and prox-regularity . . . 83
1.4.6 Conical derivative of the mapping P_C . . . 88
1.4.7 Convergence . . . 92

2 Integrability of subdifferentials of functions **99**

2.1 Integrability of the subdifferential of a convex function through infimal regularization . . . 100
2.2 Integrability of subdifferentials of directionally Lipschitz functions . . 108
2.2.1 Subdifferential properties of directionally Lipschitz functions . 110
2.2.2 Local integrability . . . 114
2.3 Integrability of subdifferentials of certain bivariate functions . . . 118
2.3.1 Concepts of regularity . . . 120

2.3.2 Integrability of subdifferentials of certain locally Lipschitz bivariate functions . . . 122
2.3.3 Concepts of directional Lipschitzness . . . 124
2.3.4 Local integrability of subdifferentials of certain directionally Lipschitz bivariate functions . . . 128
2.4 Partially ball weakly inf-compact saddle functions . . . 136
2.4.1 Saddle functions. Properties . . . 140
2.4.2 Subdifferential of a saddle function. Partially ball weakly inf-compact saddle functions. Definition and properties . . . 145
2.4.3 Integrability of the subdifferential of a proper closed partially ball weakly inf-compact saddle function . . . 153

3 Variational analysis of multivalued maps 157
3.1 Parameterized minimax problem: on Lipschitz-like dependence of the solution with respect to parameter . . . 158
3.1.1 Parameterized minimization problem . . . 160
3.1.2 Parameterized minimax problem . . . 172
3.1.3 Lipschitz continuity of the saddle points map in context of two-player zero sum differential games . . . 178
3.2 Aubin criterion for metric regularity . . . 180
3.2.1 An implicit mapping theorem . . . 185
3.2.2 Proof of Aubin criterion . . . 189
3.2.3 Applications of the Aubin criterion . . . 192
3.3 Long orbit or empty value principle, fixed point and surjectivity theorems . . . 198
3.3.1 Long Orbit or Empty Value (LOEV) principle . . . 199
3.3.2 Caristi-Kirk fixed point theorem . . . 200
3.3.3 Surjectivity theorems . . . 201

Bibliography 209

Preface

Variational Analysis covers a broad field of mathematical theory developed in connection with the study of problems of optimization, equilibrium, control, and stability of linear and nonlinear systems, as stated in the eponymous book of Rockafellar and Wets [152].

For a long time, “variational” problems have been identified mostly with the “calculus of variations” concerning minimization of integral functionals, where a major point is to explore variations in order to characterize solutions and describe them in terms of “variational principles”.

In this connection, notions of perturbation, approximation, generalized differentiability were extensively investigated.

From decades are the attempts to free the term “variational” from the limitations of its past and to use it for much larger area of modern mathematics. The contemporary approach is to consider “variations” not only in a classical sense: as a movement away from a given point along rays or curves, and the geometry of tangent and normal cones associated with that, but also as forms of perturbation and approximation that are describable by set convergence, set-valued mappings and so on. Subgradients and subdifferentials of functions, convex and nonconvex, are crucial in analyzing such “variations”.

At the present time Variational Analysis is considered as a branch of Analysis providing not only powerful tools for the problems that have motivated it so far but also as a mathematical discipline with new applications.

In the dissertation we present several original results in the field of Variational Analysis obtained in the last 15 years and published in 11 journal articles – cited as [18, 19, 63, 94, 95, 96, 140, 162, 163, 164, 172] in the bibliography.

Content is organized into three chapters. Each chapter is divided in sections. Each section bears the title of the eponymous article and for the convenience of the reader it begins with the necessary notations and preliminaries even they were already used before.

In Chapter 1 we study primal lower-nice functions and prox-regular sets. Primal lower-nice functions possess good behaviour as convex functions related to their regularization, integrability, etc. and they are intensively studied last years. We prove in any Banach space a characterization of primal lower-nice functions by hypomonotonicity of certain truncations of its subdifferential (Section 1.1), and that for such functions in smooth Banach spaces proximal subdifferential and Clarke subdifferential coincide (Section 1.2). Primal lower-nice functions belong to the larger class of prox-regular functions. The sets whose indicator functions belong to that class are prox-regular sets. We study their properties and find several characterizations of such sets in uniformly convex Banach spaces in Section 1.3 and Section 1.4.

Chapter 2 contains results concerning integrability of subdifferentials of functions. The purpose of studying integration of subdifferentials is to answer the question whether or not the condition that the subdifferential of one function contains the subdifferential of other function implies that these two functions differ by a constant. We give affirmative answer to that question for several classes of functions defined on Banach space – a new proof of integrability of a lower semicontinuous convex function in Banach space (well-known result of Rockafellar) using regularizations in the spirit of the pioneering proof of Moreau in Hilbert space (Section 2.1); regular function continuous on its domain and strictly directionally Lipschitz at point of its domain is integrable near that point (Section 2.2); bivariate (resp. separately) lower semicontinuous function, continuous on its domain which is (resp. separately) upper-upper regular and (resp. separately) strictly directionally Lipschitz at point of its domain is integrable near that point (Section 2.3); proper closed partially ball weakly inf-compact saddle function (Section 2.4).

In Chapter 3 we study multivalued maps, as well as, their dependence on parameter. Such maps are considered in optimization and are intensively studied recently. In Section 3.1 we present a sufficient condition, ensuring that the map which to any value of the parameter assigns the set of solutions (possibly multi-valued, and unbounded) of a parameterized minimax problem on a product Banach space possesses Aubin property. In Section 3.2 we establish a derivative criterion for metric regularity of set-valued mappings that is based on works of J.-P. Aubin and co-authors. A related implicit mapping theorem is also obtained. A new proof of the radius theorem for metric regularity based on Aubin criterion is given as well. In Section 3.3 a Long Orbit or Empty Value (LOEV) principle is proved and applied to provide unified approach to several fixed point and surjectivity results.

I would like to express my sincere gratitude to all my co-authors with whom I worked together over the years and who played a crucial role in my formation as a mathematician – my husband Assoc. Prof. Milen Ivanov (Sofia University), my supervisor Prof. Pando Georgiev (Sofia University and currently University of Florida), Prof. Lionel Thibault and Dr. Frédéric Bernard (Université Montpellier II), Prof. Marc Quincampoix (Université de Bretagne Occidentale), Prof. Asen Dontchev (University of Michigan), and Assoc. Prof. Boyan Zlatanov (Plovdiv University). I am also grateful to my colleagues from the Department of Probability, Operational Research and Statistics at the Faculty of Mathematics and Informatics of Sofia University, where I have been working for 17 years and to my colleagues from the Department of Operations Research at the Institute of Mathematics and Informatics for their collegial collaboration, and especially to Venelin Chernogorov, who had been my support for all these years.

Chapter 1

Primal lower-nice functions and prox-regular sets

In this chapter we present some results concerning properties of an important class of lower semicontinuous extended real-valued functions called “primal lower-nice” and introduced in 1991 by Poliquin in [136] and prox-regular sets which indicator functions are “prox-regular” functions – a class of functions firstly introduced in 1996 by Poliquin and Rockafellar in [137].

The class of prox-regular functions firstly introduced in finite dimensional settings enlarges significantly the class of primal lower-nice functions previously introduced in the same finite dimensional context. Convex functions and lower-C^2 functions belong to these classes, as well as qualified convexly C^2-composite functions also called strongly amenable functions (see Poliquin [136] and Poliquin and Rockafellar [137]).

The class of primal lower nice functions considerably generalizes the scope of functions that possess as good behaviour as convex functions concerning their regularization, their integrability, their second-order properties, etc., see Poliquin [136], Thibault and Zagrodny [160], Levi, Poliquin and Thibault [111], Bernard, Thibault and Zagrodny [16], Marcellin and Thibault [115] and the references therein.

In Poliquin and Rockafellar [137] (resp. Poliquin [136]) one of the key results is an important subdifferential characterization in the line of the well-known result stating that a lower semicontinuous function is convex if and only if its subdifferential is monotone. The latter result characterizing convex functions first obtained by Poliquin in $\mathbb{R}^n$ was later proved in any Banach space by Corea, Jofre and Thibault in [52]. For primal lower-nice functions characterization via hypomonotonicity of certain truncations of their subdifferentials was given by Poliquin in [136] in $\mathbb{R}^n$ and extended to the Hilbert space context by Levy, Poliquin and Thibault in [111]. In Section 1.1 we prove the characterization of primal lower-nice functions by hypomonotonicity of certain truncations of their subdifferentials in any Banach space (see Theorem 1.1.6).

The result is published by Ivanov and Zlateva in [94].

In [136] Poliquin showed that for a primal lower-nice function in $\mathbb{R}^n$ proximal subdifferential and Clarke subdifferential coincide. The result was proved in Hilbert space by Levy, Poliquin and Thibault in [111]. In Section 1.2 we establish this property for primal lower-nice functions in smooth Banach spaces (see Theorem 1.2.2). The result is published by Ivanov and Zlateva in [95].

The extension of second-order considerations was in fact the main motivation of Poliquin and Rockafellar to defining the class of prox-regular functions. They introduced as a special case the concept of prox-regularity of sets. The study of this concept was developed in Hilbert space by Poliquin, Rockafellar and Thibault in [138], where they showed its rich geometric implications.

Prox-regular sets appear under different names in the literature, depending on the point of view chosen by the authors who considered them often independently. The first one was Federer in [73], where he introduced these sets in $\mathbb{R}^n$ as the "sets with positive reach", in order to extend the Steiner polynomial formula to a much larger class of subsets of $\mathbb{R}^n$ than those of convex sets or compact C^2 manifolds. Later, motivated by different purposes other authors focused their analysis on distinct properties of sets and considered the classes of p-convex sets (Canino in [37]), sets with 2-order tangential property (Shapiro in [153]), proximally smooth sets (Clarke, Stern and Wolenski in [46]), prox-regular sets (Poliquin, Rockafellar and Thibault in [138]), and so on. All these concepts are actually known to be the same and to be equivalent in $\mathbb{R}^n$ to the notion of positively reached sets. The class of prox-regular sets is much larger than that of convex sets, but it shares with the latter many good properties with regard to the applications in optimization, control theory, etc. and also has rich geometric implications; see, in addition to the works quoted above, Clarke, Ledyaev, Stern and Wolenski [45], Thibault [158], Marcellin and Thibault [115], Edmond and L. Thibault [68], Maury and Venel [116]. Such sets are also involved in differential inclusions in mechanics (see, e.g., Colombo and Goncharov [48], Edmond and Thibault [68], Thibault [158]), in resource allocation mechanisms in economics (see, e.g., Thibault [158]), in crowd motion problems (Maury and Venel [116]), in the theoretical study of viability for differential inclusions subject to constraints (see, e.g., Thibault [158]), etc. Concerning related concepts for functions, we refer to Bernard and Thibault [15, 14, 17], Bernard, Thibault and Zagrodny [16], Degiovanni, Marino and Toques [56], Marcellin and Thibault [115], Poliquin [136], Poliquin and Rockafellar [137], Rockafellar and Wets [152], Thibault and Zagrodny [160] and the references therein, and concerning the other similar concept of subsmoothness for sets we refer to Aussel, Daniilidis and Thibault [11].

In Section 1.3 and Section 1.4 we establish a lot of properties of prox-regular sets in uniformly convex Banach spaces. These results are published by Bernard, Thibault and Zlateva in [18] and [19], respectively.

1.1 On primal lower-nice property

Poliquin in [136] shows that primal lower-nice functions in finite dimensional spaces are completely characterized by their Clarke-Rockafellar subdifferential. This is the first large class of non-convex lower semicontinuous functions with this property. The properties of these functions in infinite dimensional Hilbert spaces are investigated in detail by Levi, Poliquin and Thibault in [111], Thibault and Zagrodny in [160], Combary, Elhilali, Levi, Poliquin and Thibault in [47]. There are two natural definitions of the primal lower-nice property: one relying on lower estimate of the first order Taylor approximation (see Definition 1.1.4) and other including only subdifferentials (see Definition 1.1.5). The second one states in fact hypomonotonicity of certain truncations of the subdifferential. Their equivalence could be considered also as subdifferential characterization of primal lower-nice property and is established by Poliquin in [136] in finite dimensional spaces and by Levy, Poliquin and Thibault in [111] in Hilbert spaces.

We fill the gap by proving, in more general setting of spaces and subdifferentials, that the two definitions are equivalent. In this way we show that they characterize the same class of functions and answer the question posed by Combary, Elhilali, Levi, Poliquin and Thibault in [47].

The results from this section are published by Ivanov and Zlateva in [94].

We begin by fixing some notations. If it is not stated other, $(X, \|\cdot\|)$ will be a real Banach space, i.e. completed normed space. X^* will stand for its topological *dual space*, i.e., the set of continuous linear functionals on X. If $x^* \in X^*$, we will write $\langle x, x^* \rangle$ for the value of x^* at $x \in X$. Recall that the *weak topology* $w(X, X^*)$ is the smallest topology on X with respect to which all the functions $\langle \cdot, x^* \rangle$ $(x^* \in X^*)$ are continuous and that the *weak-star topology* $w(X^*, X)$ is the smallest topology on X^* with respect to which all the functions $\langle x, \cdot \rangle$ $(x \in X)$ are continuous. We denote by $B[x, r]$ (resp. by $B(x, r)$) the closed (resp. open) ball with centre $x \in X$ and radius r. The closed (resp. open) unit ball in X will be denoted by B (resp. by B°).

A *bornology* β on X is a family of bounded subsets of X together with the properties: $\{x\} \in \beta$ for arbitrary $x \in X$; $A \in \beta,\ D \subset A \Rightarrow D \in \beta$. It is clear that the Gâteaux bornology G consisting of all singletons is contained in any bornology and the Fréchet bornology F of all bounded sets contains any other bornology. The Banach space X is said to be *β-smooth* with respect to certain bornology β if there exists a Lipschitz continuous *bump* (i.e. with non empty bounded support) function $b \in C^1_\beta(X)$, where

$$C^1_\beta(X) = \{\ f : X \to \mathbb{R}\ :\ f \text{ is Gâteaux differentiable and the derivative is a continuous mapping from } X \text{ to the dual space } X^*, \text{ equipped with the topology of uniform convergence on the members of the bornology } \beta\}.$$

It is easy to see that any space possessing an equivalent norm, which is β differentiable on its unit sphere, is β-smooth. The inverse is not true as shown by Haydon in [83].

We consider lower semicontinuous functions from X to $\mathbb{R} \cup \{+\infty\}$. Function $f : X \to \mathbb{R} \cup \{+\infty\}$ is said to be *lower semicontinuous at point* x_0 if $f(x_0) \leq \liminf_{x \to x_0} f(x)$ and is said to be *lower semicontinuous* if it is so at any $x_0 \in X$. Function $f : X \to \mathbb{R} \cup \{+\infty\}$ is *proper* if it is not equal to $+\infty$ everywhere, that is $\operatorname{dom} f \neq \emptyset$, where $\operatorname{dom} f = \{x \in X : f(x) \in \mathbb{R}\}$.

Further we consider β-smooth subdifferential, a notion that goes back to Crandal and Lions [55].

If β is a bornology on X and $f : X \to \mathbb{R} \cup \{+\infty\}$ is a proper and lower semicontinuous function then *β-smooth subdifferential* of f at x is the set

$$D_\beta f(x) = \{u'(x) : \ u \in C^1_\beta(X) \text{ and } f - u \text{ has a local minimum at } x\}$$

if $x \in \operatorname{dom} f$ and $D_\beta f(x) = \emptyset$ if $f(x) = +\infty$.

Following Thibault and Zagrodny [160], Correa, Jofre and Thibault [52], Ioffe [89], Borwein and Ioffe [26] we will consider an abstract subdifferential operator:

Definition 1.1.1. *Abstract subdifferential operator* ∂ is an operator that associates with each function $f : X \to \mathbb{R} \cup \{+\infty\}$ and with each point $x \in X$ a subset $\partial f(x)$ of X^*, that is called an *abstract subdifferential of f at x*, and for which the following properties hold:

Property 1. $\partial f(x) = \emptyset$ if $x \notin \operatorname{dom} f$;

Property 2. $\partial f(x) = \partial g(x)$ whenever f and g coincide on a neighbourhood of x;

Property 3. $\partial f(x)$ is equal to the subdifferential in the sense of convex analysis whenever f is convex;

Property 4. If g is convex and continuous, f is lower semicontinuous and $f + g$ has local minimum point at x_0 then for arbitrary $\varepsilon > 0$ there exist $x, \ y \in X$ and $p \in \partial f(x)$, $q \in \partial g(y)$ such that:

$$\|x - x_0\| < \varepsilon, \ \|y - x_0\| < \varepsilon, \ |f(x) - f(x_0)| < \varepsilon \text{ and}$$

$$\|p + q\| < \varepsilon.$$

We will prove that β-smooth subdifferential on β-smooth Banach space is abstract subdifferential in the above sense.

Let us recall the famous

Ekeland variational principle (e.g. Phelps [133, p.45]). Let f be a proper lower semicontinuous function from a Banach space X into $\mathbb{R} \cup \{+\infty\}$. Let f be bounded below and $\varepsilon > 0$, $\bar{y} \in \operatorname{dom} f$ be such that $f(\bar{y}) \leq \inf f + \varepsilon$. Then for each $\lambda > 0$ there is $\bar{x} \in \operatorname{dom} f$ such that:

(i) $\lambda\|\bar{x}-\bar{y}\| \le f(\bar{y}) - f(\bar{x})$,

(ii) $\|\bar{x}-\bar{y}\| \le \varepsilon/\lambda$, and

(iii) $\lambda\|x-\bar{x}\| + f(x) > f(\bar{x})$ for all $x \neq \bar{x}$.

We will use the following assertion that is straightforward implication of the definition of abstract subdifferential and Ekeland variational principle.

Proposition 1.1.2. Let X be a Banach space, ∂ be an abstract subdifferential, $f : X \to \mathbb{R}\cup\{+\infty\}$ be a lower semicontinuous and $g : X \to \mathbb{R}$ be a convex continuous function. Let μ, ε and δ be given positive numbers. If x_0 is such that

$$(f+g)(x_0) < \inf(f+g)(x) + \varepsilon$$

then there exist $\bar{x}, \bar{y} \in X$ and $p \in \partial f(\bar{x})$, $q \in \partial g(\bar{y})$ such that

(i) $$\|x_0-\bar{x}\| < \frac{\varepsilon}{\mu} + \delta, \quad \|x_0-\bar{y}\| < \frac{\varepsilon}{\mu} + \delta,$$

(ii) $$\|p+q\| < \mu + \delta,$$

(iii) $$|f(x_0) - f(\bar{x})| < \varepsilon + \delta(2+\mu) + |g(\bar{x}) - g(x_0)|.$$

Proof. Applying Ekeland variational principle we find a point $x_1 \in X$ such that $\|x_0 - x_1\| < \varepsilon/\mu$ and such that the function $(f+g)(x) + \mu\|x - x_1\|$ attains its strong minimum at x_1. Moreover, $|(f+g)(x_1) - (f+g)(x_0)| < \varepsilon$.

According to Property 4 in Definition 1.1.1 for the sum of $f(x)$ and convex continuous $g(x) + \mu\|x - x_1\|$ there exist $\bar{x}$, $\bar{y} \in B(x_1, \delta)$ and $p \in \partial f(\bar{x})$, $\bar{q} \in \partial(g(\cdot) + \mu\|\cdot - x_1\|)(\bar{y})$ such that $\|p + \bar{q}\| < \delta$, $|f(\bar{x}) - f(x_1)| < \delta$ and $|g(\bar{x}) + \mu\|\bar{x} - x_1\| - g(x_1)| < \delta$. We may represent $\bar{q} = q + \mu\xi$ with some $q \in \partial g(\bar{y})$ and some $\xi \in \partial\|\cdot\|(\bar{y} - x_1)$. Then $\|p + q\| = \|p + \bar{q} - \mu\xi\| < \delta + \mu$, which is (ii).

Simple applications of the triangle inequality give (i) and (iii). □

From this result it follows that the domain of an abstract subdifferential is densely defined in the domain of a lower semicontinuous function.

Now we are ready to prove that β-smooth subdifferentials are abstract subdifferentials on β-smooth Banach spaces. The statement and the proof will be given for Gâteaux smoothness only since its expose is the same for all types of smoothness of which Gâteaux smoothness is the weakest. Of course, Property 4 is the only one that has to be verified.

In fact we prove a more general result, namely:

Theorem 1.1.3. Let X be a Gâteaux smooth Banach space, $f : X \to \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous and $g : X \to \mathbb{R}$ be a locally uniformly continuous function. If $f + g$ has a local minimum at x_0 then for arbitrary $\varepsilon > 0$ there exist $x,\ y \in X$ and $p \in D_G f(x)$, $q \in D_G g(y)$ such that:

(i) $$\|x - x_0\| < \varepsilon, \quad \|y - x_0\| < \varepsilon,$$

(ii) $$|f(x) - f(x_0)| < \varepsilon \text{ and}$$

(iii) $$\|p + q\| < \varepsilon.$$

A similar result is called enhanced fuzzy sum rule by Borwein, Mordukhovich and Shao in [29].

Proof. We follow El Haddad and Deville [80]. By the trivial invariance of Gâteaux subdifferential D_G with respect to translation and addition of a constant, there is no loss of generality in assuming that $f(0) = g(0) = 0$ and that 0 is a local minimum of $f + g$.

Fix $r > 0$ such that on rB the function f is bounded below, g is uniformly continuous, and 0 is a minimum of $f + g$ on rB.

Let $l(x)$ be a *Leduc function* on X, that is $l(x)$ is continuously Gâteaux smooth away from the origin, Lipschitz continuous, and $b\|x\| \geq l(x) \geq \|x\|$ for some constant $b > 0$, see Deville, Godefroy and Zizler [57].

The function $f + l^2 + g$ has a strong local minimum at 0.

Consider the functions

$$w_n(x,y) = \begin{cases} f(x) + l^2(x) + g(y) + nl^2(x-y) & , \quad x, y \in rB \\ \infty & , \quad \text{otherwise.} \end{cases}$$

For each $n \in \mathbb{N}$ the function w_n is lower semicontinuous and bounded below on $X \times X$, so according to the smooth variational principle, see Deville, Godefroy and Zizler [57], there exists a function $\varphi_n : X \times X \to \mathbb{R}$ that is Lipschitz continuous, Gâteaux smooth and such that $\|\varphi_n\|_\infty < n^{-1}$, $\|\varphi_n'\|_\infty < n^{-1}$, and $w_n + \varphi_n$ attains its strong minimum at (x_n, y_n).

We claim that $\|x_n - y_n\| \underset{n\to\infty}{\longrightarrow} 0$ and x_n is a minimizing sequence for $f + l^2 + g$.

First, observe that

$$(w_n + \varphi_n)(0,0) > (w_n + \varphi_n)(x_n, y_n)$$

and using $w_n(0,0) = 0$ we obtain

$$\varphi_n(0,0) > f(x_n) + l^2(x_n) + g(y_n) + nl^2(x_n - y_n) + \varphi_n(x_n, y_n). \tag{1.1}$$

Let K be a lower bound of f and g on rB. From both $\|\varphi_n\|_\infty < n^{-1}$ and $l^2(x_n - y_n) \geq \|x_n - y_n\|^2$ we have that $2n^{-1} > 2K + n\|x_n - y_n\|^2$, hence $\|x_n - y_n\| \leq \sqrt{2(1-K)n^{-1}} \underset{n\to\infty}{\longrightarrow} 0$.

By the uniform continuity of g on rB° it follows that $|g(x_n) - g(y_n)| \underset{n\to\infty}{\longrightarrow} 0$. From (1.1) we have $2n^{-1} > f(x_n) + l^2(x_n) + g(y_n) > (f + l^2 + g)(x_n) - |g(x_n) - g(y_n)|$, and then

$$2n^{-1} + |g(x_n) - g(y_n)| > (f + l^2 + g)(x_n) \geq 0, \tag{1.2}$$

hence $(f + l^2 + g)(x_n) \underset{n\to\infty}{\longrightarrow} 0$ and x_n is a minimizing sequence for $f + l^2 + g$. As 0 is a strong local minimum point for $f + l^2 + g$ this implies that $x_n \underset{n\to\infty}{\longrightarrow} 0$ and thus, $y_n \underset{n\to\infty}{\longrightarrow} 0$ too. Hence, for n large enough the points (x_n, y_n) are interior points for $rB \times rB$. Moreover, (1.2) implies that $f(x_n) \underset{n\to\infty}{\longrightarrow} 0$.

It is clear that for large n the function $f(\cdot) + l^2(\cdot) + \varphi_n(\cdot, y_n) + nl^2(\cdot - y_n)$ has a local minimum at x_n and, from the definition of D_G^- it follows that $p_n = -2l(x_n)l'(x_n) - (\varphi_n)'_x(x_n, y_n) - 2nl(x_n - y_n)l'(x_n - y_n) \in D_G^- f(x_n)$. Similarly, the function $g(\cdot) + \varphi_n(x_n, \cdot) + nl^2(x_n - \cdot)$ has a local minimum at y_n and $q_n = -(\varphi_n)'_y(x_n, y_n) + 2nl(x_n - y_n)l'(x_n - y_n) \in D_G^- g(y_n)$. Hence, $p_n + q_n = -2l(x_n)l'(x_n) - (\varphi_n)'_x(x_n, y_n) - (\varphi_n)'_y(x_n, y_n)$ and $\|p_n + q_n\| \leq 2n^{-1} + 2l(x_n)\|l'(x_n)\|$, so we may take sufficiently large n to complete the proof. □

We shall state the two known definitions of primal lower-nice property before showing that they are equivalent. They are given in arbitrary Banach space X and for arbitrary fixed abstract subdifferential ∂.

Definition 1.1.4. Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous function. Function f is said to be ∂ *primal lower-nice* (∂*-pln* in short) at $\bar{x} \in \operatorname{dom} f$ if there exist $\lambda > 0$, $c > 0$, $T > 0$ such that

$$f(y) \geq f(x) + \langle p, y - x \rangle - \frac{t}{2}\|y - x\|^2$$

whenever $t \geq T$, $x \in \bar{x} + \lambda B^\circ$, $y \in x + \lambda B^\circ$, $p \in \partial f(x)$, and $\|p\| \leq ct$.

Definition 1.1.5. Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous function. Function f is said to be ∂ *primal lower-nice* (∂*-pln* in short) at $\bar{x} \in \operatorname{dom} f$ if there exist $\lambda > 0$, $c > 0$, $T > 0$ such that

$$\langle p - q, x - y \rangle \geq -t\|x - y\|^2$$

whenever $t \geq T$, x, $y \in \bar{x} + \lambda B^\circ$, $p \in \partial f(x)$, $q \in \partial f(y)$ and $\max\{\|p\|, \|q\|\} \leq ct$.

Theorem 1.1.6. Any function f that is ∂-pln at $\bar{x} \in \operatorname{dom} f$ according to Definition 1.1.4 is also ∂-pln at $\bar{x}$ according to Definition 1.1.5 and vice versa.

Proof. The fact that function which is ∂-pln according to Definition 1.1.4 is ∂-pln according to Definition 1.1.5 is well known (see for example Combary, Elhilali, Levi, Poliquin and Thibault [47]). To establish the opposite direction we will use some ideas from Thibault and Zagrodny [160] and Levi, Poliquin and Thibault [111].

Let now f be ∂-pln at $\overline{x} \in \operatorname{dom} f$ according to Definition 1.1.5 and the constants λ, c, T be as they are therein. We will find positive constants λ', c', T' such that f satisfies Definition 1.1.4 with these "prime" constants.

The proof is divided in two parts.

Part 1. Localization. Take $0 < \lambda' < \max\{\lambda/4,\ c/16\}$ and such that f is bounded below on $B(\overline{x}, 4\lambda')$ and let $f(x) \geq K$ for $x \in B(\overline{x}, 4\lambda')$.

Fix c' and T' in a way that

$$0 < c' < \lambda'/8 \quad \text{and} \quad T' > \max\{2T, (\lambda')^{-2}(1 + f(\overline{x}) - K)\}. \tag{1.3}$$

We claim that for arbitrary $x_0 \in B(\overline{x}, \lambda')$ and $p \in \partial f(x_0)$ such that $\|p\| \leq c't$, $t \geq T'$ and arbitrary $y_0 \in B(\overline{x}, 4\lambda')$ such that

$$f(y_0) + \langle p, x_0 - y_0 \rangle + \frac{t}{2}\|y_0 - x_0\|^2 < \inf_{y \in B(\overline{x}, 4\lambda')} \left\{ f(y) + \langle p, x_0 - y \rangle + \frac{t}{2}\|y - x_0\|^2 \right\} + 1$$

it follows that $y_0 \in B(\overline{x}, 3\lambda')$.

Assume the contrary. That is, there are $x' \in B(\overline{x}, \lambda')$ and $p' \in \partial f(x')$ such that $\|p'\| \leq c't$, for some $t \geq T'$, and $y' \in B(\overline{x}, 4\lambda') \setminus B(\overline{x}, 3\lambda')$ such that

$$f(y') + \langle p', x' - y' \rangle + \frac{t}{2}\|y' - x'\|^2 < \inf_{y \in B(\overline{x}, 4\lambda')} \left\{ f(y) + \langle p', x' - y \rangle + \frac{t}{2}\|y - x'\|^2 \right\} + 1.$$

In particular

$$f(y') + \langle p', x' - y' \rangle + \frac{t}{2}\|y' - x'\|^2 < f(\overline{x}) + \langle p', x' - \overline{x} \rangle + \frac{t}{2}\|\overline{x} - x'\|^2 + 1.$$

Thus,

$$1 + f(\overline{x}) - K > \langle p', \overline{x} - y' \rangle + \frac{t}{2}(\|y' - x'\|^2 - \|\overline{x} - x'\|^2).$$

Observe that $\|y' - x'\| - \|\overline{x} - x'\| \geq \|y' - \overline{x}\| - 2\|\overline{x} - x'\| \geq 3\lambda' - 2\lambda' = \lambda'$ to estimate $\|y' - x'\|^2 - \|\overline{x} - x'\|^2 = (\|y' - x'\| - \|\overline{x} - x'\|)(\|y' - x'\| + \|\overline{x} - x'\|) \geq \lambda'.\|y' - \overline{x}\| \geq \lambda'.3\lambda' = 3(\lambda')^2$.

Also, $\langle p', \overline{x} - y' \rangle \geq -\|p'\|.\|\overline{x} - y'\| \geq -4\lambda'\|p'\| \geq -4\lambda' c't$, because $\|p'\| \leq c't$.

Hence,

$$1 + f(\overline{x}) - K > -4\lambda' c't + \frac{t}{2}3(\lambda')^2 = t\lambda'\left(\frac{3}{2}\lambda' - 4c'\right) > t(\lambda')^2$$

using for the last inequality that $c' < \lambda'/8$ (see (1.3)). Then we have that

$$1 + f(\overline{x}) - K > t(\lambda')^2 > T'(\lambda')^2$$

which is a contradiction to the choice of T' (see (1.3)).

Part 2. Variation. Let λ', c', T' be fixed as in Part 1. We claim that f is ∂-pln at $\overline{x}$ with λ', c', T' according to Definition 1.1.4.

Assume the contrary. Then, there are $x_0 \in B(\overline{x}, \lambda')$ and $p \in \partial f(x_0)$ such that $\|p\| \le c't$, $t \ge T'$ and

$$\inf_{y \in B(x_0,\lambda')} \left\{ f(y) + \langle p, x_0 - y\rangle + \frac{t}{2}\|y - x_0\|^2 \right\} < f(x_0).$$

Now, it is clear that

$$\inf_{y \in B(\overline{x},4\lambda')} \left\{ f(y) + \langle p, x_0 - y\rangle + \frac{t}{2}\|y - x_0\|^2 \right\} < f(x_0). \tag{1.4}$$

Define the function

$$h(y) = \begin{cases} f(y) + \langle p, x_0 - y\rangle + \frac{t}{2}\|y - x_0\|^2 &, \ y \in \overline{x} + 4\lambda' B \\ +\infty &, \ \text{otherwise.} \end{cases}$$

Let y_n be such that $h(y_n) < \inf h(y) + n^{-1}$. Because of (1.4) there is no subsequence of $\{y_n\}$ norm convergent to x_0. Hence, there exists $r > 0$ such that $\|y_n - x_0\| \ge r$.

According to Part 1 the points y_n are in $B(\overline{x}, 3\lambda')$.

Through applying Proposition 1.1.2 for $\varepsilon = (2n)^{-1}$, $\mu = (4n)^{-1/2}$, $\delta = (2n)^{-1}$, and sufficiently large n, we obtain points $x_n \in B(y_n, n^{-1/2})$, $z_n \in B(y_n, n^{-1/2})$ and $p_n \in \partial f(x_n)$, $q_n \in -p + 2^{-1}t\partial(\|\cdot - x_0\|^2)(z_n)$ such that $\|p_n + q_n\| < n^{-1/2}$.

We can represent $q_n = -p + t\xi_n$ with some $\xi_n \in 2^{-1}\partial(\|\cdot - x_0\|^2)(z_n)$ which implies that $\langle p + q_n, x_0 - z_n\rangle = -t\|x_0 - z_n\|^2$.

We need to estimate $\|p_n\|$. First, we have that $\left|\|p_n\| - \|q_n\|\right| \underset{n\to\infty}{\longrightarrow} 0$. Second, we consider $\|q_n\| = \|p - t\xi_n\| \le \|p\| + t\|x_0 - z_n\| \le c't + t(\|x_0 - \overline{x}\| + \|\overline{x} - y_n\| + \|y_n - z_n\|) \le c't + t(\lambda' + 3\lambda' + n^{-1/2}) \le (ct)/3$ for sufficiently large n.

Hence, for n large enough, we have that $x_n \in B(\overline{x}, \lambda)$, $p_n \in \partial f(x_n)$ is such that $\|p_n\| \le ct/2$, also $x_0 \in B(\overline{x}, \lambda)$, $p \in \partial f(x_0)$ is such that $\|p\| \le c't < ct/2$, and $t/2 \ge T'/2 > T$. According to Definition 1.1.5 we have that

$$\langle p - p_n, x_0 - x_n\rangle \ge -\frac{t}{2}\|x_0 - x_n\|^2. \tag{1.5}$$

We estimate the left-hand side by the following sequence of inequalities:

$$\begin{aligned}\langle p - p_n, x_0 - x_n\rangle &= \langle p + q_n, x_0 - x_n\rangle - \langle p_n + q_n, x_0 - x_n\rangle \\ &\leq \langle p + q_n, x_0 - z_n\rangle + \langle p + q_n, z_n - x_n\rangle + \|p_n + q_n\|.\|x_0 - x_n\| \\ &\leq -t\|x_0 - z_n\|^2 + \|p + q_n\|.\|z_n - x_n\| + \|p_n + q_n\|.\|x_0 - x_n\| \\ &\leq -t\|x_0 - x_n\|^2 + 2t\|x_n - x_0\|.\|z_n - x_n\| + \\ &\quad + \|p + q_n\|.\|z_n - x_n\| + \|p_n + q_n\|.\|x_0 - x_n\| \\ &\quad \text{and using that } \|x_0 - x_n\| \text{ and } \|p + q_n\| \text{ are bounded} \\ &\quad \text{and } \|x_n - z_n\| \underset{n\to\infty}{\longrightarrow} 0,\ \|p_n + q_n\| \underset{n\to\infty}{\longrightarrow} 0 \text{ we have that} \\ &\leq -t\|x_0 - x_n\|^2 + \alpha_n, \text{ where } \alpha_n \underset{n\to\infty}{\longrightarrow} 0.\end{aligned}$$

Combining with (1.5) we obtain

$$\frac{t}{2}\|x_0 - x_n\|^2 \leq \alpha_n.$$

Passing, if necessary, to subsequence we have that $\|x_0 - x_n\| \underset{n\to\infty}{\longrightarrow} a \geq r > 0$, while the right-hand side tends to zero, so $\frac{t}{2}a^2 \leq 0$ and $t \leq 0$, which is a contradiction. □

1.2 Subdifferential characterization of primal lower-nice functions on smooth Banach spaces

Poliquin was proved in [136] that Clarke-Rockafellar subdifferential and proximal subdifferential of a primal lower-nice function on finite-dimensional space coincide. This means in particular that if the definition of primal lower-nice property (see Definition 1.1.4 and Definition 1.1.5) is taken with respect to Clarke-Rockafellar subdifferential, this will produce the same class of functions. Also, these functions are completely characterized by their Clarke-Rockafellar subdifferential, see Poliquin [136]. This was the first large class of non-convex lower semicontinuous functions with this property.

The coincidence of proximal subdifferential and Clarke-Rockafellar subdifferentials of a primal lower-nice function defined on Hilbert space was proved by Levy, Poliquin and Thibault [111]. Their proof uses the representation of Clarke-Rockafellar subdifferential in Hilbert space as a sequential limit of proximal subdifferentials due to Loewen [114]. Since a similar representation is available in general smooth Banach spaces (see Borwein and Ioffe [26], and Ivanov [91]) it is possible to extend the result of Levy, Poliquin and Thibault to such spaces.

Here we show that Clarke-Rockafellar subdifferential and proximal subdifferential of a primal lower-nice function defined on a β-smooth Banach space coincide (see

Theorem 1.2.2). The result obtained demonstrates that the class of primal lower-nice functions does not depend on what reasonable subdifferential is used in defining the class. This would mean that it is possible to characterize primal lower-nice property in terms not involving subdifferentials.

The results from this section are published by Ivanov and Zlateva in [95].

Let $(X, \|.\|)$ be a Banach space, X^* be its dual space. Recall that the *Clarke-Rockafellar subdifferential* (or simply the *Clarke subdifferential*) of a proper lower semicontinuous function $f : X \to \mathbb{R} \cup \{+\infty\}$ at point $x \in \operatorname{dom} f$ is the set

$$\partial_C f(x) = \{p \in X^* : f^{\uparrow}(x; v) \geq \langle p, v\rangle, \ \forall v \in X\},$$

where

$$f^{\uparrow}(x; v) = \lim_{\varepsilon \downarrow 0} \limsup_{\substack{y \to_f x \\ t \downarrow 0}} \inf_{w \in v + \varepsilon B} \frac{f(y + tw) - f(y)}{t}$$

is the Clarke generalized derivative and $y \to_f x$ means that $(y, f(y))$ tend to $(x, f(x))$ in $X \times \mathbb{R}$. If $f(x) = +\infty$ then $\partial_C f(x) = \emptyset$.

Also, recall that $p \in X^*$ is said to be a *proximal subgradient* of the lower semicontinuous function $f : X \to \mathbb{R} \cup \{+\infty\}$ at $x \in \operatorname{dom} f$, written $p \in \partial_p f(x)$, if for some $t > 0$ the inequality

$$f(y) \geq f(x) + \langle p, y - x\rangle - \frac{t}{2}\|y - x\|^2$$

holds for all y in a neighbourhood of x. The set of all such p is called *proximal subdifferential* of f at x.

Borwein and Ioffe in [26] mentioned that it is quite useful to split β-smooth subdifferential in the following manner. For any $k > 0$ one defines

$$D^k_\beta f(x_0) = \{ p \in X^* : \text{there is a } \beta \text{ smooth function } g : X \to \mathbb{R} \text{ with a Lipschitz constant } k \text{ such that the function } f - g \text{ has a local minimum at } x_0 \text{ and } g'(x_0) = p\}.$$

It is clear that for $x_0 \in \operatorname{dom} f$

$$D_\beta f(x_0) = \bigcup_{k>0} D^k_\beta f(x_0).$$

If $x_0 \notin \operatorname{dom} f$ then $D^k_\beta f(x_0) = D_\beta f(x_0) = \emptyset$.

Representation formulae of Clarke-Rockafellar subdifferential of a lower semicontinuous function in terms of different smaller subdifferentals are obtained, for example, in Rockafellar [151], Loewen [114], Ioffe [89], Borwein and Ioffe [26], Ivanov [91], etc. We use the following

Representation formulae (Ivanov [91]). Let X be a β-smooth Banach space and $f : X \to \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous function. Then

$$\partial_C f(x) = \overline{\mathrm{co}}^* \left(\widetilde{\partial}_G f(x) + \widetilde{\partial}_G^\infty f(x) \right),$$

where

$$\widetilde{\partial}_G f(x) = \bigcup_{k=1}^{\infty} \left\{ w^* - \lim_{n\to\infty} p_n;\ p_n \in D_\beta^k f(x_n), x_n \to_f x \right\},$$

and

$$\widetilde{\partial}_G^\infty f(x) = \bigcup_{k=1}^{\infty} \left\{ w^* - \lim_{n\to\infty} \lambda_n^{-1} p_n;\ p_n \in D_\beta^{\lambda_n k} f(x_n),\ x_n \to_f x,\ \lambda_n \to \infty \right\}.$$

Further we will work in a β-smooth Banach space and will consider β-smooth subdifferential. In this context is natural to consider D_β-pln functions.

A lower semicontinuous function $f : X \to \mathbb{R} \cup \{+\infty\}$ defined on a β-smooth Banach space X is said to be D_β *primal lower-nice* (D_β*-pln* in short) at $x_0 \in \mathrm{dom}\, f$ if there exist $\lambda > 0$, $c > 0$, $T > 0$ such that

$$f(y) \geq f(x) + \langle p, y - x \rangle - \frac{t}{2}\|y - x\|^2$$

whenever $t \geq T$, $x \in B(x_0, \lambda)$, $y \in B(x, \lambda)$, $p \in D_\beta f(x)$, and $\|p\| \leq ct$.

In fact this is Definition 1.1.4 taken for the subdifferential $\partial = D_\beta$.

Obviously, in any Banach space X,

$$\partial_p f(x) \subseteq \partial_C f(x), \tag{1.6}$$

and in β-smooth Banach space X,

$$\partial_p f(x) \subseteq D_\beta f(x) \subseteq \partial_C f(x).$$

By Definition 1.1.4 it is clear that if f is ∂-pln at x_0 then

$$\partial f(x) \cap ctB^* \subset \partial_p f(x) \text{ for all } x \in B(x_0, \lambda).$$

Lemma 1.2.1. Let X be a Banach space. Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be a ∂-pln at $x_0 \in \mathrm{dom}\, f$ function with respect an abstract subdifferential ∂ including ∂_p. Then $\partial_p f(x_0)$ is a convex and w^* closed subset of X^*.

Proof. The convexity of $\partial_p f(x_0)$ is obvious from its definition. The set $\partial_p f(x_0)$ being a convex subset of X^* has the same closures in w^* and bw^* topologies (see Holmes [86, Corollary 2, p. 154]). Let us recall that a set in X^* is bw^* closed when it contains the w^* limits of all bounded and w^* converging nets of its elements. Hence, it is enough to show that $\partial_p f(x_0)$ is a bw^* closed set.

To this end, let $\{p_\alpha\}_{\alpha\in A}$ be a norm bounded net such that $p_\alpha \in \partial_p f(x_0)$, $\alpha \in A$ and $p_\alpha \to p$ in w^* topology.

Let λ, c and T be as in the definition of ∂ primal lower-nice property of f at x_0.

Since the net $\{p_\alpha\}_{\alpha\in A}$ is norm bounded, we can take $t > T$ such that $\|p_\alpha\| < ct$ for all $\alpha \in A$. Since f is ∂-pln at x_0 and $\partial f(x) \supset \partial_p f(x)$ for all x, we have that

$$f(x) \geq f(x_0) + \langle p_\alpha, x - x_0\rangle - \frac{t}{2}\|x - x_0\|^2$$

for any x such that $\|x - x_0\| \leq \lambda$ and for all $\alpha \in A$. Passing to limit in w^* topology we obtain that $p \in \partial_p f(x_0)$. □

In particular, since in β-smooth Banach space D_β is a subdifferential including ∂_p, then for a D_β-pln at $x_0 \in \operatorname{dom} f$ function $f : X \to \mathbb{R} \cup \{+\infty\}$ the proximal subdifferential $\partial_p f(x_0)$ is a convex and w^* closed subset of X^*.

When a function f is ∂-pln at all points in $\operatorname{dom} f$, then f is said to be ∂-pln.

Now we are ready to prove

Theorem 1.2.2. Let X be β-smooth Banach space and $f : X \to \mathbb{R} \cup \{+\infty\}$ be a D_β-pln function. Then $\partial_p f \equiv D_\beta f \equiv \partial_C f$.

Proof. Let $x_0 \in \operatorname{dom} f$ and λ, c and T be as in the definition of D_β primal lower-nice property of f at x_0.

First, we show that $\widetilde{\partial}_G f(x_0) \subseteq \partial_p f(x_0)$. Let $p \in \widetilde{\partial}_G f(x_0)$.

Let $k \geq 1$ and p be a w^* limit of a sequence $\{p_n\}$, $p_n \in D^k_\beta f(x_n)$, $x_n \to_f x_0$. The sequence $\{p_n\}$ is norm bounded by k, so we can take $t > T$ such that $\|p_n\| \leq k < ct$. Eventually, $\|x_n - x_0\| \leq \lambda$. Since f is D_β-pln at x_0 for all large enough n and all x such that $\|x - x_0\| \leq \lambda$ we have that

$$f(x) \geq f(x_n) + \langle p_n, x - x_n\rangle - \frac{t}{2}\|x - x_n\|^2.$$

Passing to limit as $n \to \infty$ we obtain that $p \in \partial_p f(x_0)$.

Further, we show that $\widetilde{\partial}_G f(x_0) + \widetilde{\partial}^\infty_G f(x_0) \subseteq \partial_p f(x_0)$.

Let $p \in \widetilde{\partial}_G f(x_0)$ and $p^\infty \in \widetilde{\partial}^\infty_G f(x_0)$.

Let $k \geq 1$ and p^∞ be a w^* limit of a sequence $\{\lambda_n^{-1} p_n\}$, $p_n \in D^{\lambda_n k}_\beta f(x_n)$, $x_n \to_f x_0$, $\lambda_n \to \infty$. The sequence $\{\lambda_n^{-1} p_n\}$ is norm bounded by k, so $\|p_n\| \leq k\lambda_n$. Since $\lambda_n \to \infty$ we have that $k\lambda_n \geq Tc$ for sufficiently large n. Eventually $\|x_n - x_0\| \leq \lambda$. Since f is D_β-pln at x_0 the inequality

$$f(x) \geq f(x_n) + \langle p_n, x - x_n\rangle - \frac{k\lambda_n}{2c}\|x - x_n\|^2$$

holds for any x such that $\|x - x_0\| \le \lambda$. Dividing by λ_n and taking limit as $n \to \infty$ we obtain that

$$0 \ge \langle p^\infty, x - x_0\rangle - \frac{k}{2c}\|x - x_0\|^2$$

for any $x \in \operatorname{dom} f \cap B(x_0, \lambda)$.

We have from the first part of the proof the inequality

$$f(x) \ge f(x_0) + \langle p, x - x_0\rangle - \frac{t}{2}\|x - x_0\|^2$$

for any $x \in \operatorname{dom} f \cap B(x_0, \lambda)$.

We only need to sum up the last two inequalities to obtain that

$$f(x) \ge f(x_0) + \langle p + p^\infty, x - x_0\rangle - \left(\frac{t}{2} + \frac{k}{2c}\right)\|x - x_0\|^2$$

for any $x \in \operatorname{dom} f \cap B(x_0, \lambda)$ (hence, for any $x \in B(x_0, \lambda)$) which means that $p + p^\infty \in \partial_p f(x_0)$.

Therefore,

$$\widetilde{\partial}_G f(x_0) + \widetilde{\partial}_G^\infty f(x_0) \subseteq \partial_p f(x_0). \tag{1.7}$$

To complete the proof, we remind that $\partial_p f(x_0)$ is a convex and w^* closed set, see Lemma 1.2.1. From the latter, the representation formulae of Ivanov, (1.7) and (1.6) it follows that

$$\begin{aligned}\partial_C f(x_0) &= \overline{\text{co}}^*\left(\widetilde{\partial}_G f(x_0) + \widetilde{\partial}_G^\infty f(x_0)\right) \subseteq \\ &\subseteq \overline{\text{co}}^*\partial_p f(x_0) = \partial_p f(x_0) \subseteq \\ &\subseteq \partial_C f(x_0). \qquad \square\end{aligned}$$

Remark 1.2.3. Not every function f such that $\partial_p f \equiv \partial_C f$ is necessarily primal lower-nice. Consider for example $f : \mathbb{R} \to \mathbb{R}$ defined as

$$f(x) := \int_0^x t^2 \sin\frac{1}{t^2}\,dt.$$

It is easy to check that f is everywhere twice differentiable. So, $\partial_p = \partial_C f = \{f'\}$, but f is not primal lower-nice at 0. Otherwise it must locally satisfy $f'' \ge -C(|f'| + 1)$ for some constant $C > 0$. The latter will imply that f'' is bounded from below at a neighbourhood of 0, but $\liminf_{x\to 0} f''(x) = -\infty$.

1.3 Characterizations of prox-regular sets in uniformly convex Banach spaces

In this section we extend to the setting of uniformly convex Banach spaces several results obtained for prox-regular sets in Hilbert spaces. Prox-regularity of a set C at a point $x \in C$ is a variational condition related to normal vectors and which is common to many types of sets. In the context of uniformly convex Banach spaces, the prox-regularity of a closed set C at x is shown to be still equivalent to the property of the distance function d_C to be continuously differentiable outside of C on some neighbourhood of x. Additional characterizations are provided in terms of metric projection mapping. We also examine the global level of prox-regularity corresponding to the continuous differentiability of the distance function d_C over an open tube of uniform thickness around the set C.

The differentiability of the distance function d_C to a nonempty closed subset C of a Banach space X and the single valuedness of the metric projection mapping P_C are longstanding subjects of study. For a convex set C, the differentiability of d_C^2 and the single valuedness and continuity of P_C on the whole space X are well known in smooth Banach spaces. In the finite dimensional Euclidean case, Motzkin [130] seems to be the first to prove that a nonempty closed set C is convex if and only if its metric projection mapping P_C is single-valued everywhere. In the Hilbert setting, Klee [103] proved that for weakly closed sets C, the convexity of C is also characterized in this way.

The characterization in the Hilbert setting of the convexity of a norm closed set C by the metric projection mapping P_C being single-valued and norm-to-weak continuous is due to Asplund [1]. When the dual space of X is rotund, Vlasov [167] extended in some sense Asplund's result by showing that a norm closed set C of X is convex if and only if the metric projection mapping is single-valued and norm-to-norm continuous. See also Vlasov [166] for the case of approximately compact sets.

Using the original result by Fitzpatrick [75] reducing the differentiability of d_C to its Gâteaux directional derivability in a certain key direction, Borwein, Fitzpatrick and Giles [24] characterized closed convex sets C of a Banach space X with rotund dual as closed sets C for which the distance function d_C is Gâteaux differentiable on $X \setminus C$ with $\|\nabla^G d_C(x)\| = 1$ for all $x \in X \setminus C$. This result of Fitzpatrick will be stated below as a theorem (see Theorem 1.3.20) because of its importance.

In order to extend the Steiner polynomial formula concerning the n-dimensional measure of the r-neighbourhood (with respect to the Euclidean norm) of a closed convex subset or a compact C^2-submanifold of $\mathbb{R}^n$ to a much larger class of sets, Federer [73] introduced the concept of subsets of $\mathbb{R}^n$ with *positive reach*. For a nonempty closed set $C \subset \mathbb{R}^n$, denoting by Unp(C) the set of all points $x \in \mathbb{R}^n$ for which C contains a unique nearest point to x, Federer defined its *reach* (that he denoted by reach(C)) as the largest r (possibly $+\infty$) such that $\{x \in \mathbb{R}^n : 0 < d_C(x) < r\} \subset \mathrm{Unp}(C)$. Then, he declared C to be positively reached whenever reach(C) > 0 and established, among other results, that d_C is continuously differentiable on the set $\{x \in \mathbb{R}^n : 0 < d_C(x) < \mathrm{reach}(C)\}$. Note that Federer also worked, for a fixed point $\bar{x} \in C$ with reach($C, \bar{x}$), i.e., the supremum of all $r > 0$ such that the open ball centered at $\bar{x}$ with radius r is included in Unp(C). Considering, in Hilbert space, the concept of *p-convex set* C of Degiovanni, Marino and Tosques [56], Canino [37] established that on a suitable open neighbourhood of C the metric projection mapping P_C is single-valued and locally Lipschitz continuous. Staying on the global level in Hilbert space, Clarke, Stern and Wolenski [46] introduced and studied the *proximally smooth sets*. Such sets correspond to closed sets C for which the distance function d_C is continuously differentiable on an open tube around C of the type

$$U_C(r) := \{u \in X : 0 < d_C(u) < r\}$$

for some $r > 0$. In view of Federer's result recalled above, in finite dimensions those sets are positively reached and vice versa. Clarke, Stern and Wolenski characterized, in Hilbert space, proximal smooth sets in several interesting ways, in particular in terms of proximal normals and proximal mapping P_C. They also provided (in finite dimensions) a detailed analysis of locally Lipschitz continuous functions for which the epigraph is proximally smooth. Another previous interesting result was obtained by Shapiro [153] on the local level, in the Hilbert setting. He proved, for a closed set C and a point $\bar{x} \in C$, that the metric projection mapping P_C is single-valued on a neighbourhood of $\bar{x}$ whenever the distance to the general Boulingand contingent cone to C satisfies a property referred to as the *Shapiro property* by Poliquin, Rockafellar and Thibault [138].

On the local level, in the study of sets C for which d_C is *locally* differentiable and its consequences for the metric projection mapping P_C, Poliquin, Rockafellar and Thibault [138] recently made advance in the Hilbert setting with a different point of view, by making the link with the local property of C called *prox-regularity*. This property has been introduced as a new important regularity in variational analysis by Poliquin and Rockafellar [137]. They defined this concept for functions and for sets, and studied it in the finite dimensional setting. Rich geometric implications and characterizations of such a concept were obtained by Poliquin, Rockafellar and Thibault. In their work [138], they relied, in Hilbert space, on the prox-regularity of a closed set C at $\bar{x} \in C$ and characterized it in terms of d_C and P_C, e.g. d_C being continuously differentiable outside of C on a neighbourhood of $\bar{x}$ or P_C being single-valued and norm-to-weak continuous on this same neighbourhood. They also gave a subdifferential characterization of such sets with the normal cone to C. Coming back to the *global level*, they showed that proximally smooth sets are exactly *uniformly prox-regular* sets and provided new insights on those sets.

We extend the scope of these results to the more general setting of uniformly convex Banach spaces (e.g. l_p, L^p and W^p_m with $1 < p < \infty$) and find their analogues in the context of such spaces. We will rely on a property that we introduce, analogous to the prox-regularity.

In Subsection 1.3.1 we consider the definition and the first properties of prox-regular sets on a uniformly convex Banach space X. We also introduce related definitions for functions and for set-valued mappings. Subsection 1.3.2 is devoted to the study of properties of local Moreau envelopes of functions on X. In Subsection 1.3.3 we establish several characterizations of prox-regular sets in X (see Theorem 1.3.25) extending in this way the results of Poliquin, Rockafellar and Thibault [138]. In the final Subsection 1.3.4 we use some techniques developed in the previous subsections to obtain in Theorem 1.3.27 various results similar to the characteristic Theorem 1.3.25 but on the global level of proximally smooth sets. So, with Theorems 1.3.25 and 1.3.27 we extend several results of Federer [73], Canino [37], Clarke, Stern and Wolenski [46], Poliquin, Rockafellar and Thibault [138], and Colombo and Goncharov [48].

The results from this section are published by Bernard, Thibault and Zlateva in [18].

We begin by recalling some of the properties of uniformly convex Banach spaces which can be found in the books of Diestel [58], Brezis [34], Beauzamy [12], Deville, Godefroy and Zizler [57].

For a Banach space X the following are equivalent:

(X1) X has an equivalent *uniformly convex* norm $\|.\|$, i.e., such that its *modulus of convexity*

$$\delta_{\|.\|}(\varepsilon) := \inf\left\{1-\left\|\frac{x+y}{2}\right\| : \|x\| = \|y\| = 1, \|x-y\|\geq\varepsilon\right\}$$

satisfies $\delta_{\|.\|}(\varepsilon)>0$ for all $\varepsilon\in]0,2]$.

(X2) X has an equivalent uniformly convex norm $\|.\|$ with modulus of convexity of *power type* q, i.e., for some $k > 0$ one has $\delta_{\|.\|}(\varepsilon) \geq k\varepsilon^q$, for all $\varepsilon \in]0,2]$. The function $\delta_{\|.\|}$ is increasing in $]0,2]$. From Dvoretzky's theorem necessarily $q \geq 2$.

(X3) X has an equivalent *uniformly smooth* norm $\|.\|$, i.e., such that its *modulus of smoothness*

$$\rho_{\|.\|}(\tau) := \frac{1}{2}\sup\{\|x+y\| + \|x-y\| - 2 : \|x\| = 1, \|y\| \leq \tau\} \quad \text{for } \tau \geq 0$$

satisfies $\lim_{\tau\downarrow 0} \frac{\rho_{\|.\|}(\tau)}{\tau} = 0$.

(X4) X has an equivalent uniformly smooth norm $||.||$ with modulus of smoothness of *power type* s, i.e., such that for some $c > 0$ one has $\rho_{||.||}(\tau) \leq c\tau^s$ for all $\tau \geq 0$. From Dvoretzky's theorem necessarily $1 < s \leq 2$.

(X5) X has an equivalent norm which is both uniformly convex and uniformly smooth and which has moduli of convexity and smoothness of power type q and s, respectively.

It is well-known (see Diestel [58], Lindenstrauss and Tzafriri [112, 113]) that all Hilbert spaces H and the Banach spaces l^p, L^p, and W^p_m $(1 < p < \infty)$ all are (for their usual norms) uniformly convex and uniformly smooth with moduli of convexity and smoothness of power type. More precisely, for $\varepsilon \in]0, 2]$ and $\tau \geq 0$

$$\delta_H(\varepsilon) = 1 - \sqrt{1 - (1/4)\varepsilon^2} \geq \varepsilon^2/4,$$

$$\delta_{l^p}(\varepsilon) = \delta_{L^p}(\varepsilon) = \delta_{W^p_m}(\varepsilon) = \begin{cases} \frac{p-1}{8}\varepsilon^2 + o(\varepsilon^2) > \frac{p-1}{8}\varepsilon^2, & 1 < p < 2, \\ 1-\left[1-\left(\frac{\varepsilon}{2}\right)^p\right]^{1/p} > \frac{1}{p}\left(\frac{\varepsilon}{2}\right)^p, & p \geq 2, \end{cases}$$

$$\rho_H(\tau) = (1 + \tau^2)^{1/2} - 1 < \tau,$$

$$\rho_{l^p}(\tau) = \rho_{L^p}(\tau) = \rho_{W^p_m}(\tau) = \begin{cases} (1 + \tau^p)^{1/p} - 1 < \frac{1}{p}\tau^p, & 1 < p < 2, \\ \frac{p-1}{2}\tau^2 + o(\tau^2) < \frac{p-1}{2}\tau^2, & p \geq 2. \end{cases}$$

In current Section 1.3 we work in an uniformly convex Banach space X which is equipped with an equivalent norm $||.||$ that satisfies (X5).

Such a norm is a *Kadec norm*, i.e., it satisfies the property that whenever $x_n \xrightarrow[n\to\infty]{w} x$, meaning that the sequence x_n converges to x in the weak topology, with $||x_n|| \xrightarrow[n\to\infty]{} ||x||$, then $x_n \xrightarrow[n\to\infty]{||.||} x$, meaning that the sequence x_n converges to x in the norm topology.

Let q and s be the power types of moduli of convexity and smoothness of $||.||$, respectively. Then X^* is also uniformly convex and its dual norm has modulus of convexity of power type $q^* = s(s-1)^{-1}$ and modulus of smoothness of power type $s^* = q(q-1)^{-1}$.

The mapping $J : X \to X^*$ defined by

$$J(x) := \{x^* \in X^* : \langle x^*, x\rangle = ||x^*||.||x||, \quad ||x^*|| = ||x||\}$$

is generally called the *normalized duality mapping*. Let us put together, in the context of uniformly convex Banach space X satisfying (X5), some of its properties that we will use hereafter (see Cioranescu [43]):

(J) The mapping $J : X \to X^*$ is single-valued, bijective and norm-to-norm uniformly continuous on bounded sets, $J(\lambda x) = \lambda J(x)$ for all $\lambda \in \mathbb{R}$, $\|J(x)\| = \|x\|$, and $J(x) = \nabla \frac{1}{2}\|.\|^2(x)$ for all $x \in X$.

An analogous property (J^*) holds for the normalized duality mapping $J^* : X^* \to X$. Moreover, $J^* = J^{-1}$.

It is known (see e.g. Xu and Roach [169]), that for $r > 0$ there exist positive constants K_r, K'_r such that

$$\langle J(x) - J(y), x - y \rangle \geq K_r \|x - y\|^q, \quad \forall x, y \in rB, \tag{1.8}$$

$$\|J(x) - J(y)\| \leq K'_r \|x - y\|^{s-1}, \quad \forall x, y \in rB. \tag{1.9}$$

The space $X \times \mathbb{R}$ will be endowed with the norm $|||\cdot|||$ given by $|||(x, r)||| = \sqrt{\|x\|^2 + r^2}$. So, for the normalized duality mapping $J_{X\times\mathbb{R}} : X \times \mathbb{R} \to X^* \times \mathbb{R}$ associated with the norm $|||\cdot|||$, one has the equality

$$J_{X\times\mathbb{R}}(x, r) = (J(x), r). \tag{1.10}$$

Recall that a normed vector space $(Y, \|\cdot\|)$ is *rotund* or *strictly convex* provided that for any $y, y' \in Y$ with $\|y\| = \|y'\| = 1$ and $y \neq y'$ one has $\left\|\frac{1}{2}(y + y')\right\| < 1$. According to (X1), the uniform convexity holds when this inequality is fulfilled in some uniform way. Recall (see for example Deville, Godefroy and Zizler [57] and Fabian, Habala, Hájek, Montesinos, Pelant and Zizler [71]) that the strict convexity of the norm $\|\cdot\|$ is equivalent to require for any non zero $y, y' \in Y$, $y \neq y'$ the equality

$$\|y + y'\| = \|y'\| + \|y\| \tag{1.11}$$

to entail $y' = \mu y$ for some $\mu > 0$.

We will need the following elementary result concerning nearest points of a closed subset C to a point in Y. It can be found in Hilbert space for example in Clarke, Ledyaev, Stern and Wolenski [45, p.4]. It must also be known in the general strictly convex setting but we did not find it in the literature.

As usual, $d_C(u)$ denotes the distance from u to the set C, i.e. $d_C(u) := \inf_{x\in C} \|u - x\|$ and $P_C(u) := \{x \in C : d_C(u) := \inf_{x\in C} \|u - x\|\}$ denotes the set of all nearest points of C to u.

Lemma 1.3.1. Let $(Y, \|\cdot\|)$ be a strictly convex normed vector space, C be a closed subset of Y and $u \notin C$. Assume that $P_C(u) \neq \emptyset$. Then for any $p \in P_C(u)$ and any $t \in]0, 1]$, one has $P_C(u + t(p - u)) = \{p\}$.

Proof. The case $t = 1$ being obvious, we may suppose $t \in]0,1[$. Putting $u_t := u + t(p-u)$ we have

$$\|u_t - p\| = \|u + t(p-u) - p\| = (1-t)\|u-p\| = (1-t)d_C(u).$$

Further, for any $y \in C$,

$$\begin{aligned} \|u_t - y\| = \|u + t(p-u) - y\| &\geq \|u-y\| - t\|u-p\| \\ &\geq d_C(u) - t\|u-p\| \\ &= (1-t)d_C(u) \\ &= \|u_t - p\|, \end{aligned}$$

and hence $p \in P_C(u_t)$.
Suppose that there exists $p_t \neq p$ with $p_t \in P_C(u_t)$. Then setting $y = p_t$ in the above sequence of inequalities we obtain

$$\begin{aligned} d_C(u_t) = \|u_t - p_t\| = \|u - p_t + t(p-u)\| &\geq \|u-p_t\| - t\|u-p\| \\ &\geq d_C(u) - t\|u-p\| \\ &= \|u_t - p\| = d_C(u_t). \end{aligned}$$

All the last inequalities are then equalities and hence

$$\|u - p_t\| = d_C(u) \tag{1.12}$$

and

$$\|u - p_t\| = \|u - p_t + t(p-u)\| + \|t(u-p)\|.$$

Further, obviously $t(u-p) \neq 0$, and one also has $u - p_t + t(p-u) \neq 0$ since $\|u - p_t + t(p-u)\| \geq (1-t)d_C(u)$. So, because of the rotundity, the last equality above entails (see (1.11)) that there exists $\mu > 0$ with

$$u - p_t + t(p-u) = \mu t(u-p),$$

that is,

$$u - p_t = t(\mu+1)(u-p). \tag{1.13}$$

Using (1.12) and taking the norm of both members of (1.13) yield $d_C(u) = t(\mu+1)d_C(u)$ and since $d_C(u) > 0$ we have $t(\mu+1) = 1$. Putting this value in (1.13) gives $p_t = p$, which completes the proof. □

For a set $C \subset X$ we will denote by $\mathrm{cl}\, C$ its norm closure in X. A vector $p \in X$ is said to be a *primal proximal normal vector* (shortly *proximal normal vector*) to

C at $x \in \operatorname{cl} C$ (see Borwein and Strójwas [31]) if there are $u \notin \operatorname{cl} C$ and $r > 0$ such that $p = r^{-1}(u - x)$ and $\|u - x\| = d_C(u)$. It is known according to Lau theorem (see Lau [108]) that in any reflexive Banach space endowed with a Kadec norm, the set of those points which have a nearest point to any fixed closed subset is a dense set. In the appropriately renormed space X we are working in, the above property holds and hence there are proximal normal vectors at any point of some dense subset of the boundary of C. Observe that the proximal normality of a non-zero $p \in X$ to C at $x \in \operatorname{cl} C$ corresponds to the existence of some $r > 0$ such that $x \in P_{\operatorname{cl} C}(x + rp)$. The cone of all such vectors p, together with the origin, will be denoted by $PN_C(x)$ and called *primal proximal normal cone* of C at x.

The concept is local in the sense that for any $u \notin \operatorname{cl} C$ and any closed ball $V := B[x, \beta]$ centered at $x \in \operatorname{cl} C$ such that $\|u - x\| = d_{C \cap V}(u)$ one has $u - x \in PN_C(x)$. Indeed, put $\rho := d_{C \cap V}(u) > 0$ and $u_t := x + t(u - x)$ for any fixed positive $t < \min\left(1, \frac{\beta}{2\rho}\right)$. Observe first that $u_t \in \operatorname{int} V$ according to the inequality $t < \frac{\beta}{2\rho}$ and hence $u_t \notin \operatorname{cl} C$ because otherwise one would get $u_t \in \operatorname{cl}(C \cap V)$ and

$$\|u_t - u\| = (1 - t)\|u - x\| < d_{C \cap V}(u)$$

which would be a contradiction. Further $\|u_t - x\| = t\rho$ and, on the one hand, for any $y \in C \cap V$ one can write

$$\|u_t - y\| = \|u + (1 - t)(x - u) - y\| \geq \|u - y\| - (1 - t)\|u - x\| \geq t\rho = \|u_t - x\|.$$

On the other hand, for any $y \in C \setminus V$ one has

$$\|u_t - y\| \geq \|y - x\| - \|u_t - x\| > \beta - t\|u - x\| = \beta - t\rho > t\rho = \|u_t - x\|,$$

the last inequality being due to the choice of t. So, $\|u_t - x\| = d_C(u_t)$ and hence by the definition of $PN_C(x)$ we have

$$u - x = t^{-1}(u_t - x) \in PN_C(x). \tag{1.14}$$

A continuous linear functional $p^* \in X^*$ is said to be a *proximal normal functional* to C at $x \in \operatorname{cl} C$ (see Borwein and Strójwas [31]) if there are $u \notin \operatorname{cl} C$, $r > 0$ such that $p^* = r^{-1}J(u - x)$ and $\|u - x\| = d_C(u)$. Or, equivalently, a non-zero $p^* \in X^*$ is a proximal normal functional to C at $x \in \operatorname{cl} C$ if there exists $r > 0$ such that $x \in P_{\operatorname{cl} C}(x + rJ^*(p^*))$. The cone of all such functionals p^*, together with the origin, will be denoted by $N_C^P(x)$. One easily verifies that if $p \in PN_C(x)$, then $J(p) \in N_C^P(x)$, and that if $p^* \in N_C^P(x)$, then $J^*(p^*) \in PN_C(x)$. Hence, $PN_C(x)$ and $N_C^P(x)$ completely determine each other.

A functional $x^* \in X^*$ is said to be a *Fréchet normal functional* (see Borwein and Strójwas [31]) to C at x if for any $\varepsilon > 0$ there exists a neighbourhood U_ε of x such that the inequality $\langle x^*, x' - x\rangle - \varepsilon\|x' - x\| \leq 0$ holds for all $x' \in C \cap U_\varepsilon$.

As the norm of the space X we work in is Fréchet differentiable away from the origin, it is not difficult to verify that for any closed subset $C \subset X$ and any $x \in C$, any proximal normal functional to C at x is also a Fréchet normal functional to C at x (see Borwein and Strójwas [31, Corollary 3.1]).

Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous function. By definition, the effective *domain* of f is the set $\operatorname{dom} f := \{x \in X : f(x) < +\infty\}$ and the *epigraph* of f is the set $\operatorname{epi} f := \{(x, r) \in X \times \mathbb{R} : f(x) \le r\}$. It is clear that $\operatorname{epi} f$ is non-empty whenever f is proper, and $\operatorname{epi} f$ is closed exactly when f is lower semicontinuous.

Let $x \in \operatorname{dom} f$. We say that $p^* \in X^*$ is a *proximal subgradient* of f at x if $(p^*, -1)$ is a proximal normal functional to the epigraph of f at $(x, f(x))$. The *proximal subdifferential* of f at x, denoted by $\partial_p f(x)$, consists of all such functionals. Thus, we have $p^* \in \partial_p f(x)$, if and only if, $(p^*, -1) \in N^P_{\operatorname{epi} f}(x, f(x))$. The functional $x^* \in X^*$ is said to be a *Fréchet subgradient* of f at x if $(x^*, -1)$ is a Fréchet normal functional to the epigraph of f at $(x, f(x))$. The *Fréchet subdifferential* of f at x, denoted by $\partial_F f(x)$, consists of all such functionals. If $x \notin \operatorname{dom} f$ then all subdifferentials of f at x are empty, by convention. It is known that for a lower semicontinuous function f on a reflexive Banach space with a Kadec and Fréchet differentiable norm (in particular, on X), the set $\operatorname{dom} \partial_p f$ is dense in $\operatorname{dom} f$ (see Borwein and Strójwas [32, Theorem 7.1]). Moreover, from what we saw above, $\partial_p f(x) \subset \partial_F f(x)$ for all $x \in X$. The Fréchet subgradients are known (see Ioffe [89]) to have an analytical characterization in the sense that $x^* \in \partial_F f(x)$, if and only if, $\liminf_{y \to x} \frac{f(y) - f(x) - \langle x^*, y - x\rangle}{\|y - x\|} \ge 0$. When $\partial_F f(x) \neq \emptyset$, one says that f is Fréchet subdifferentiable at the point x.

As usual, we will denote by ψ_C the *indicator function* of a closed set $C \subset X$, i.e., $\psi_C(y) = 0$ if $y \in C$ and $\psi_C(y) = +\infty$ otherwise. It is easily checked that $\partial_p \psi_C(x) = N^P_C(x)$ for any $x \in C$.

Like for the proximal normal cone in Hilbert space (see Clarke, Stern and Wolenski [46] and Bounkhel and Thibault [33]) one can express, in our uniformly convex space X, the proximal normal functional cone to C in terms of the proximal subdifferential of d_C. We denote by B^* the closed unit ball of X^* and by $C(\rho)$ the ρ-enlargement of the set C, i.e., $C(\rho) := \{u \in X : d_C(u) \le \rho\}$.

Proposition 1.3.2. For any closed subset C of X and any $x \in C$,

$$\partial_p d_C(x) = N^P_C(x) \cap B^*.$$

Proof. The inclusion $x^* \in \partial_p d_C(x)$ means $(x^*, -1) \in N^P_{\operatorname{epi} d_C}(x, 0)$, or equivalently, for any $t > 0$ small enough,

$$\inf_{(y,\lambda) \in \operatorname{epi} d_C} \{\|x + tv - y\|^2 + (-t - \lambda)^2\} = t^2\|v\|^2 + t^2, \tag{1.15}$$

where $v = J^*(x^*)$. This entails that

$$\inf_{y \in C} \{\|x + tv - y\|^2\} = t^2\|v\|^2,$$

hence $v \in PN_C(x)$. Further (1.15) ensures for all $y \in X$ that

$$\|x + tv - y\|^2 + 2td_C(y) + d_C^2(y) \geq t^2\|v\|^2$$

and since $-2tJ(v) = \nabla(-\|\cdot\|^2)(tv)$, for each $\varepsilon > 0$ there exists some positive number $r < \varepsilon$ such that for all $y \in B[x, r]$

$$2t\langle -J(v), x - y\rangle \leq \varepsilon\|x - y\| + 2td_C(y) + d_C^2(y)$$

and taking the inequality $d_C(y) \leq \|x - y\|$ into account we see that

$$2t\langle -J(v), x - y\rangle \leq (\varepsilon + 2t + \|x - y\|)\|x - y\| \leq (2\varepsilon + 2t)\|x - y\|.$$

This easily yields $\|v\| = \| - J(v)\| \leq 1$ and hence $x^* \in N_C^P(x) \cap B^*$.

Conversely, take $x^* \in N_C^P(x) \cap B^*$. Put $v = J^*(x^*)$ and choose $t > 0$ small enough that $d_C^2(x + tv) = t^2\|v\|^2$. Then

$$\inf_{y\in X}\{\|x + tv - y\|^2 + (t + d_C(y))^2\} = \inf_{\rho\geq 0} h(\rho),$$

where $h(\rho) := \inf_{y\in C(\rho)}\{\|x + tv - y\|^2 + (t + \rho)^2\}$. Obviously we have the equality $h(\rho) = d_{C(\rho)}^2(x + tv) + (t + \rho)^2$. Consider two cases:

— If $d_C(x + tv) \leq \rho$, then $h(\rho) = (t + \rho)^2 \geq t^2 + d_C^2(x + tv)$.

— If $d_C(x + tv) > \rho$, using Bounkhel and Thibault [33, Lemma 3.1] or Lemma 1.3.28 in forthcoming Subsection 1.3.4, we have $d_{C(\rho)}(x + tv) = d_C(x + tv) - \rho$ and thus,

$$h(\rho) = d_C^2(x + tv) + t^2 + 2\rho[\rho + t - d_C(x + tv)] \geq d_C^2(x + tv) + t^2.$$

The last estimation comes from the fact that $\|v\| \leq 1$.

Finally, making use of the above inequalities and of the equality $d_C^2(x+tv) = t^2\|v\|^2$ we obtain

$$\inf_{y\in X}\{\|x + tv - y\|^2 + (t + d_C(y))^2\} = t^2\|v\|^2 + t^2,$$

which entails that $(v, -1)$ is in $PN_{\text{epi } d_C}(x, 0)$ or, in other words, that $x^* \in \partial_p d_C(x)$. □

1.3.1 Prox-regular sets

The concept of prox-regularity was introduced for functions from $\mathbb{R}^n$ into $\mathbb{R} \cup \{+\infty\}$ by Poliquin and Rockafellar in [137], extending the class of primal lower nice functions previously considered by Poliquin in [136]. The introduction of prox-regular functions in Poliquin and Rockafellar [137] has been motivated by the study of second-order properties of some non-convex functions. A subset of $\mathbb{R}^n$ is defined to be prox-regular in Poliquin and Rockafellar [137] when its indicator function is prox-regular. The concept of prox-regularity of sets then has been studied and developed in Hilbert space in [138] by Poliquin, Rockafellar and Thibault who showed in particular its rich geometric implications.

The prox-regularity property for a closed set C is, like for functions, *local* and *directional*, concerning a point $\overline{x} \in C$ and a direction $\overline{p} \in PN_C(\overline{x})$. If the property holds for all possible proximal normal vectors to C at $\overline{x}$, the set is said to be prox-regular at $\overline{x}$. Considering the prox-regularity at a point, Poliquin, Rockafellar and Thibault [138] showed it to be a localization of Federer's positive reach concept (see Federer [73]) or proximal smoothness property of Clarke, Stern and Wolenski [46]. The localization of the mentioned behavior is made clear by the fact that the authors showed in Poliquin, Rockafellar and Thibault [138] the equivalence of the prox-regularity of a set C, of the local single valuedness and continuity of the metric projection mapping P_C, and of the local C^1 regularity of the square distance function d_C^2 to C among other characterizations. Their approach allowed them to retrieve also the important global level results of Clarke, Stern and Wolenski [46].

Before extending the above concepts to uniformly convex spaces, we must point out that, in addition to Federer's study of Steiner polynomial formula (see Federer [73]) and Canino's work related to geodesics, the first strong applications of proximal smoothness of sets to Control theory has been provided by Clarke, Ledyaev, Stern and Wolenski [45]. For other recent applications to evolution problems with moving sets putting in light the amenability of prox-regular and proximally smooth sets, we refer to Edmond and Thibault [68] and Thibault [158].

Our extension of the definition of a prox-regular set to our setting will use the duality mapping as follows, in order to find striking characterizations of prox-regularity like in the Hilbert space setting.

Definition 1.3.3. A closed set $C \subset X$ is called *prox-regular* at $\overline{x} \in C$ for $\overline{p}^* \in N_C^P(\overline{x})$ if there exist $\varepsilon > 0$ and $r > 0$ such that for all $x \in C$ and for all $p^* \in N_C^P(x)$ with $\|x - \overline{x}\| < \varepsilon$ and $\|p^* - \overline{p}^*\| < \varepsilon$ the point x is a nearest point of $\{x' \in C : \|x' - \overline{x}\| < \varepsilon\}$ to $x + rJ^*(p^*)$. The set C is prox-regular at $\overline{x}$ if this property holds for all $\overline{p}^* \in N_C^P(\overline{x})$.

The following proposition shows that the prox-regularity concept for subsets of X in fact does not depend on any direction. The first part of its proof reproduces ideas of the proof of Proposition 1.2 in Poliquin, Rockafellar and Thibault [138].

Proposition 1.3.4. A closed set $C \subset X$ is prox-regular at $\overline{x}$, if and only if, it is prox-regular at $\overline{x}$ for $\overline{p}^* = 0$. If the closed set C is prox-regular at $\overline{x}$ for $\overline{p}^* = 0$ with ε and r, then for all $x \in C$ with $\|x - \overline{x}\| < \varepsilon$ and for all $p^* \in N_C^P(x)$ with $\|p^*\| \leq \varepsilon$,

$$0 \geq \langle J[J^*(p^*) - r^{-1}(x' - x)], x' - x\rangle, \quad \forall x' \in C \text{ with } \|x' - \overline{x}\| < \varepsilon. \tag{1.16}$$

Proof. Obviously, if C is prox-regular at $\overline{x}$ for all $\overline{p}^* \in N_C^P(\overline{x})$ then it is so for $\overline{p}^* = 0$. To establish the converse, let us assume that C is prox-regular at $\overline{x}$ for $\overline{p}^* = 0$ with $\varepsilon > 0$ and $r > 0$. Take any $\overline{p}^* \in N_C^P(\overline{x})$ with $\overline{p}^* \neq 0$, and set $\varepsilon' := \min\{\varepsilon/2, \|\overline{p}^*\|/2\}$. For $x \in C$ and $p^* \in N_C^P(x)$ with $\|x - \overline{x}\| < \varepsilon'$ and $\|p^* - \overline{p}^*\| < \varepsilon'$ we have that

$$\frac{\varepsilon}{2\|\overline{p}^*\|}\|p^*\| \leq \frac{\varepsilon}{2\|\overline{p}^*\|}[\|\overline{p}^* - p^*\| + \|\overline{p}^*\|] \leq \frac{\varepsilon}{2\|\overline{p}^*\|}\varepsilon' + \frac{\varepsilon}{2} \leq \frac{\varepsilon}{4} + \frac{\varepsilon}{2} < \varepsilon.$$

We may rewrite the latter as $\left\|\frac{\varepsilon p^*}{2\|\overline{p}^*\|} - 0\right\| < \varepsilon$. By prox-regularity of C at $\overline{x}$ for 0, we have that x is a nearest point of $\{x' \in C : \|x' - \overline{x}\| < \varepsilon\}$ to $x + rJ^*\left(\frac{\varepsilon p^*}{2\|\overline{p}^*\|}\right) = x + \frac{r\varepsilon}{2\|\overline{p}^*\|}J^*(p^*)$. This means that C is prox-regular at $\overline{x}$ for $\overline{p}^*$ with constants ε' and $r' = \frac{r\varepsilon}{2\|\overline{p}^*\|}$.

To prove the second claim, we suppose that C is prox-regular at $\overline{x}$ for $\overline{p}^* = 0$ with ε and r. Fix any $x \in C$ with $\|x - \overline{x}\| < \varepsilon$ and any $p^* \in N_C^P(x)$ with $0 < \|p^*\| < \varepsilon$. By definition, x is a nearest point of $\{x' \in C : \|x' - \overline{x}\| < \varepsilon\}$ to $x + rJ^*(p^*)$, that is,

$$\|x + rJ^*(p^*) - x\| \leq \|x + rJ^*(p^*) - x'\|, \quad \forall x' \in C \text{ with } \|x' - \overline{x}\| < \varepsilon.$$

Setting $p := J^*(p^*)$ we rewrite the latter as

$$r\|p\| \leq \|x - x' + rp\|, \quad \forall x' \in C \text{ with } \|x' - \overline{x}\| < \varepsilon. \tag{1.17}$$

Since $J(u)$ is the derivative of $\frac{1}{2}\|.\|^2$ at u, we have for all $t > 0$ and all x' that

$$\gamma := \langle J[J^*(p^*) - r^{-1}(x' - x)], x' - x\rangle = \langle J(p - r^{-1}(x' - x)), x' - x\rangle \leq$$

$$[2t]^{-1}\{\left\|p - r^{-1}(x' - x) + t(x' - x)\right\|^2 - \left\|p - r^{-1}(x' - x)\right\|^2\}.$$

In particular for $t = r^{-1}$

$$\gamma \leq \frac{r}{2}\left\{\|p\|^2 - \|p - r^{-1}(x' - x)\|^2\right\} = \frac{r}{2}\left\{\|p\|^2 - r^{-2}\|x - x' + rp\|^2\right\} =$$

$$\frac{1}{2r}\left\{(r\|p\|)^2 - \|x - x' + rp\|^2\right\}.$$

Taking (1.17) into account we deduce that $\gamma \leq 0$ for all $x' \in C$ with $\|x' - \overline{x}\| < \varepsilon$, which is (1.16). The case $\|p^*\| = \varepsilon$ is obtained via a limit process and hence the proof is complete. □

In what follows, saying that C is prox-regular at $\overline{x}$ with ε and r we will mean that the constants ε and r are taken from prox-regularity of C at $\overline{x}$ for $\overline{p}^* = 0$. It is clear that if the closed set C is prox-regular at $\overline{x}$ for $\overline{p}^* = 0$ with some positive constants ε and r then it is so for any constants $0 < \varepsilon' \leq \varepsilon$ and $0 < r' \leq r$.

The notion corresponding to the inequality (1.16) in the case of functions is introduced in the following Definition 1.3.5. Let us note that another definition is considered in Bernard and Thibault [14], using the "proximal-type" estimation with the square of the norm as in Poliquin and Rockafellar [137] instead of (1.18). In the Hilbert setting, functions satisfying the definition given below, are primal lower-nice functions (see Poliquin and Rockafellar [137] or Bernard and Thibault [14]).

We will see in the next subsection that the J-plr concept introduced in Definition 1.3.5 below for functions also yields, concerning their Moreau envelopes, various important properties which have their own interest. Recall that the context here is broader than in Bernard and Thibault [14] since the power of the modulus of convexity is any $q \geq 2$. These properties applied to the indicator functions of sets will be among the keys of the development of our study.

Definition 1.3.5. A lower semicontinuous function $f : X \to \mathbb{R} \cup \{+\infty\}$ is *J-primal lower regular* (*J-plr* in short) at $\overline{x} \in \operatorname{dom} f$ if there exist positive constants ε and r such that

$$f(y) \geq f(x) + \langle J[J^*(p^*) - t(y - x)], y - x\rangle \tag{1.18}$$

for all $x, y \in B(\overline{x}, \varepsilon)$, all $p^* \in \partial_p f(x)$, and all t such that $\|p^*\| \leq \varepsilon r t$.

It is easily seen that if f is J-plr at $\overline{x}$ with some positive constants ε and r then it is so for any constants $0 < \varepsilon' \leq \varepsilon$ and $0 < r' \leq r$. If the lower semicontinuous function f is J-plr at $\overline{x} \in \operatorname{dom} f$ with positive constants ε and r, one can derive that

$$\langle J[J^*(p^*) - t(y - x)] - J[J^*(q^*) - t(x - y)], y - x\rangle \leq 0 \tag{1.19}$$

for all $x, y \in B(\overline{x}, \varepsilon)$, for all $p^* \in \partial_p f(x)$, $q^* \in \partial_p f(y)$, and all t such that $\max\{\|p^*\|, \|q^*\|\} \leq \varepsilon r t$. This is the analog of the hypomonotonicity of certain truncations of $\partial_p f$, that is intrinsic to primal lower-nice functions in Hilbert spaces: see Poliquin [136], Levi, Poliquin and Thibault [111], Bernard, Thibault and Zagrodny [16] and the references therein. The hypomonotonicity is no more appropriate in our setting and therefore we introduce the following closely related concept, that we call J-hypomonotonicity.

Definition 1.3.6. A set-valued mapping $T : X \rightrightarrows X^*$ is said to be *J-hypomonotone* of degree $t \geq 0$ if for any $(x_i, x_i^*) \in \operatorname{gph} T := \{(x, x^*) \in X \times X^* : x^* \in T(x)\}$, $i = 1, 2$, one has

$$\langle J[J^*(x_1^*) - t(x_2 - x_1)] - J[J^*(x_2^*) - t(x_1 - x_2)], x_2 - x_1\rangle \leq 0.$$

For a set-valued mapping $T : X \rightrightarrows X^*$, we denote by $\operatorname{Dom} T$ its domain, i.e. $\operatorname{Dom} T := \{x \in X : T(x) \neq \emptyset\}$. We will also use the following concept of truncation of a set-valued mapping, as in Bernard, Thibault and Zagrodny [16].

Let $T : X \rightrightarrows X^*$ be a set-valued mapping and let $\varepsilon > 0$ and $t \geq 0$. Then its *ε, t-truncation* at a point $\overline{x} \in X$ is the set-valued mapping $T_{\overline{x},\varepsilon,t}$ defined by

$$\operatorname{gph} T_{\overline{x},\varepsilon,t} := \{(x, x^*) \in \operatorname{gph} T : \|x - \overline{x}\| < \varepsilon,\ \|x^*\| \leq t\}.$$

Without ambiguity, $T_{\overline{x},\varepsilon,t}$ will be denoted simply by T_t.

So, by (1.19) we see that if f is J-plr at $\overline{x}$ with ε and r, then $(\partial_p f)_{\overline{x},\varepsilon,\varepsilon rt}$ is J-hypomonotone of degree t for any $t \geq 0$. We are not far from sets since the next proposition shows that the prox-regularity of a set C entails the J-plr property of its indicator function ψ_C. The equivalence will be obtained later in Theorem 1.3.25, as well as the equivalence with the J-hypomonotonicity of a certain truncation of the normal cone N_C^P.

It will be convenient, for any $\sigma \geq 0$, to denote below by $N_C^{P\sigma}$ the set-valued mapping N_C^P truncated with σB^*, i.e.,

$$N_C^{P\sigma}(x) := N_C^P(x) \cap \sigma B^* \quad \text{for all } x \in X. \tag{1.20}$$

Proposition 1.3.7. If the closed set $C \subset X$ is prox-regular at $\overline{x} \in C$ with ε and r, then the indicator function ψ_C of C is J-plr at $\overline{x}$, and hence (1.19) yields, for any $t \geq 0$, the J-hypomonotonicity of degree t on $B(\bar{x}, \varepsilon)$ of the set-valued mapping $N_C^{P\sigma}$, where $\sigma := \varepsilon rt$.

Proof. As C is prox-regular at $\overline{x}$ for $\overline{p}^* = 0$ with $\varepsilon > 0$ and $r > 0$, from Proposition 1.3.4 we have

$$\psi_C(x') \geq \psi_C(x) + \langle J[J^*(p^*) - r^{-1}(x' - x)], x' - x\rangle,$$

whenever $x' \in B(\overline{x}, \varepsilon)$, $x \in C \cap B(\overline{x}, \varepsilon)$, and $\|p^*\| \leq \varepsilon$ with $p^* \in N_C^P(x)$. We have already noticed that $p^* \in N_C^P(x)$ if and only if $p^* \in \partial_p \psi_C(x)$. If $p^* \in N_C^P(x)$ and $\|p^*\| \leq \varepsilon rt$, then $r^{-1}t^{-1}p^* \in N_C^P(x)$ with $\|r^{-1}t^{-1}p^*\| \leq \varepsilon$, hence

$$\psi_C(x') \geq \psi_C(x) + \langle J[J^*(r^{-1}t^{-1}p^*) - r^{-1}(x' - x)], x' - x\rangle,$$

$$\psi_C(x') \geq \psi_C(x) + \langle J[J^*(p^*) - t(x' - x)], x' - x\rangle,$$

which means that the function ψ_C is J-plr at $\overline{x}$. The proof is complete. □

We will need another result concerning J-hypomonotone set-valued mappings. It will be one of the key steps of our development of the proof of Theorem 1.3.13. First recall that a set-valued mapping $T : X \rightrightarrows X^*$ is bounded when its range $T(X) := \cup_{x \in X} T(x)$ is a bounded set in X^*.

Lemma 1.3.8. Let $T : X \rightrightarrows X^*$ be a bounded set-valued mapping which is J-hypomonotone of degree $\overline{r}$. Then for any $r > 2\overline{r}$ we have that $(I + r^{-1}J^* \circ T)^{-1}$ is a single-valued mapping on its domain which is $\frac{1}{q}$-Hölder continuous on its intersection with any bounded subset.

Proof. Let us denote by μ any upper bound of $\{\|z\| : z \in T(x), x \in \operatorname{Dom} T\}$. For any $r > 2\bar{r}$ and $\rho > 0$, take $x_i \in \operatorname{Dom}(I + r^{-1}J^* \circ T)^{-1}$ with $\|x_i\| \leq \rho$, $i = 1, 2$. Choose any $y_i \in (I + r^{-1}J^* \circ T)^{-1}(x_i)$, i.e., $J[r(x_i - y_i)] \in T(y_i)$, $i = 1, 2$. Hence $\|x_i - y_i\| \leq \mu/r$. By assumption, for any $t \geq \bar{r}$,

$$\begin{array}{rcl} 0 & \geq & \langle J\{J^*(J[r(x_1 - y_1)]) - t(y_2 - y_1)\} - J\{J^*(J[r(x_2 - y_2)]) - t(y_1 - y_2)\}, y_2 - y_1\rangle \\ 0 & \geq & \langle J(rx_1 - ty_2 + (t - r)y_1) - J(rx_2 - ty_1 + (t - r)y_2), y_2 - y_1\rangle. \end{array}$$

Now for any $\lambda \in]0, 1[$ such that $\frac{\lambda r}{2} > \bar{r}$, replacing t in the above inequality by rt_λ where $t_\lambda := \lambda/2$ we obtain

$$\begin{array}{rcl} 0 & \geq & \langle J(x_1 - t_\lambda y_2 - (1 - t_\lambda)y_1) - J(x_2 - t_\lambda y_1 - (1 - t_\lambda)y_2), x_1 - x_2 + (1 - 2t_\lambda)(y_2 - y_1)\rangle \\ & & +\langle J(x_1 - t_\lambda y_2 - (1 - t_\lambda)y_1) - J(x_2 - t_\lambda y_1 - (1 - t_\lambda)y_2), x_2 - x_1\rangle := \text{(I)} + \text{(II)}. \end{array}$$

Note also that

$$\begin{array}{rcl} \|x_1 - t_\lambda y_2 - (1 - t_\lambda)y_1\| & \leq & (1 - t_\lambda)\|x_1 - y_1\| + t_\lambda\|x_1 - y_2\| \\ & \leq & (1 - t_\lambda)\|x_1 - y_1\| + t_\lambda(\|x_1 - x_2\| + \|x_2 - y_2\|) \\ & \leq & (1 - t_\lambda)\frac{\mu}{r} + t_\lambda\left(2\rho + \frac{\mu}{r}\right) \leq \gamma, \end{array}$$

where $\gamma := \rho + \frac{\mu}{r}$, and similarly $\|x_2 - t_\lambda y_1 - (1 - t_\lambda)y_2\| \leq \gamma$. A first estimation of (I) is obtained by using (1.8), that is,

$$\text{(I)} \geq K_\gamma\|x_1 - x_2 + (1 - 2t_\lambda)(y_2 - y_1)\|^q.$$

To proceed further in the estimation, we need to consider two cases.

The **first case** is when $(1 - 2t_\lambda)\|y_1 - y_2\| > \|x_1 - x_2\|$. In that case, we need to estimate below $\|a - b\|^q$ when $\|a\| > \|b\|$. Since $\|a - b\| \geq \|a\| - \|b\| > 0$, we derive $\|a - b\|^q \geq [\|a\| - \|b\|]^q = \|a\|^q\left[1 - \frac{\|b\|}{\|a\|}\right]^q$. This leads us to consider the real-valued function $g(s) = [1 - s]^q + qs$ on the interval $s \in [0, 1[$. As the derivative $g'(s) = -q[1-s]^{q-1}+q$ is non-negative on this interval, the function g is non-decreasing on $[0, 1[$ and then $g(s) \geq g(0) = 1$ for all $s \in [0, 1[$. Finally, $[1 - s]^q \geq 1 - qs$ for $s \in [0, 1[$. We conclude that

$$\|a - b\|^q \geq \|a\|^q\left[1 - \frac{\|b\|}{\|a\|}\right]^q \geq \|a\|^q\left[1 - q\frac{\|b\|}{\|a\|}\right] = \|a\|^q - q\|a\|^{q-1}\|b\|.$$

Using the latter we continue to estimate (I) by

$$\begin{array}{rcl} \text{(I)} & \geq & K_\gamma\left[(1 - 2t_\lambda)^q\|y_1 - y_2\|^q - q(1 - 2t_\lambda)^{q-1}\|y_1 - y_2\|^{q-1}\|x_1 - x_2\|\right] \\ & \geq & \gamma_1\|y_1 - y_2\|^q - \gamma_2\|x_1 - x_2\|, \end{array}$$

where γ_1, γ_2 are some nonnegative constants, depending on λ. On the other hand,

$$\begin{aligned}(\mathrm{II}) &\geq -\|J(x_1 - t_\lambda y_2 - (1-t_\lambda)y_1) - J(x_2 - t_\lambda y_1 - (1-t_\lambda)y_2)\|.\|x_2 - x_1\| \\ &\geq -K'_\gamma \|x_1 - x_2 + (1-2t_\lambda)(y_2 - y_1)\|^{s-1}.\|x_2 - x_1\| \\ &\geq -\gamma_3 \|x_2 - x_1\|,\end{aligned}$$

where we used (1.9) for the second estimation, and γ_3 is some nonnegative constant depending on λ. Finally, $0 \geq (\mathrm{I}) + (\mathrm{II}) \geq \gamma_1\|y_1 - y_2\|^q - (\gamma_2 + \gamma_3)\|x_1 - x_2\|$ and hence for some constant $\gamma' > 0$ that depends on λ,

$$\|y_1 - y_2\| \leq \gamma' \|x_1 - x_2\|^{1/q}.$$

The **second case** is when $(1 - 2t_\lambda)\|y_1 - y_2\| \leq \|x_1 - x_2\|$. Observing that $\|x_1 - x_2\| \leq (2\rho)^{1-\frac{1}{q}}\|x_1 - x_2\|^{\frac{1}{q}}$, we see that in both cases we have that $\|y_1 - y_2\| \leq \gamma''\|x_1 - x_2\|^{1/q}$ for some constant $\gamma'' > 0$, so $\left(I + r^{-1}J^* \circ T\right)^{-1}$ is a single-valued mapping on its domain and it is $\frac{1}{q}$-Hölder continuous on the intersection of its domain with the set ρB. □

1.3.2 Local Moreau envelopes

Several properties of d_C^2 and P_C will be derived from corresponding ones (with their own interest) concerning the so-called local Moreau envelope of a function. Here we will give the definition and properties of local Moreau envelopes. Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous function and $W \subset X$ be a nonempty closed subset where f is bounded from below and finite at some point. The *local Moreau envelope of index* $\lambda > 0$ *of* f (relative to W), is defined as

$$e_{\lambda,W} f(x) := \inf_{y \in W} \left\{ f(y) + \frac{1}{2\lambda}\|x - y\|^2 \right\}. \tag{1.21}$$

We fix W with the above property and we will write, when there is no risk of confusion, $e_\lambda f$ instead of $e_{\lambda,W} f$. Note that the infimum in (1.21) may be seen as taken over all X for the function $\tilde{f}$ given by $\tilde{f}(x) = f(x)$ if $x \in W$ and $\tilde{f}(x) = +\infty$ otherwise. It is easy to see that the functions $e_\lambda f$ are everywhere defined and Lipschitz on bounded subsets. As usual we will consider the set

$$P_\lambda f(x) := \left\{ y \in W : e_\lambda f(x) = f(y) + \frac{1}{2\lambda}\|x - y\|^2 \right\}.$$

Whenever there exists some $p_\lambda(x) \in P_\lambda f(x)$ one has by Correa, Jofre and Thibault [51]

$$\partial_F e_\lambda f(x) \subset \{\lambda^{-1} J(x - p_\lambda(x))\} \cap \partial_F \tilde{f}(p_\lambda(x)), \tag{1.22}$$

hence $P_\lambda f(x)$ is either empty or a singleton thanks to the one-to-one property of the mapping J. We know by Theorem 11 of Borwein and Giles [25] that the infimum is attained whenever x is a point of Fréchet subdifferentiability of $e_\lambda f$. Denoting by G_λ the subset of X where $e_\lambda f$ is Fréchet subdifferentiable, we obtain for any $x \in G_\lambda$ that $P_\lambda f(x) = \{p_\lambda(x)\}$ and $\partial_F e_\lambda f(x) = \{\lambda^{-1}J(x - p_\lambda(x))\}$. Note that G_λ is dense in X according to the result in Mordukhovich and Shao [126] and Preiss [139] concerning the density of subdifferentiability points.

When $P_\lambda f(x) = \{p_\lambda(x)\}$ is a singleton, we make no difference between will write sometimes $P_\lambda f(x)$ and $p_\lambda(x)$.

In the proof of the next lemma we follow an idea due to Borwein and Giles from [25]. The lemma will be used in the proof of Proposition 1.3.19.

Lemma 1.3.9. Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous function bounded from below on W and with $W \cap \operatorname{dom} f \neq \emptyset$. Then the following assertions are equivalent:

(a) $\partial_F e_\lambda f(x) \neq \emptyset$;

(b) $e_\lambda f$ is Fréchet differentiable at x.

Further, in the case of (a) or (b), $\nabla^F e_\lambda f(x) = \lambda^{-1}J(x - p_\lambda(x))$.

Proof. Obviously (b) implies (a). Now suppose that (a) holds. As we have seen above, $P_\lambda f(x)$ is single-valued with a unique element $p_\lambda(x)$ and $\partial_F e_\lambda f(x) = \{\lambda^{-1}J(x-p_\lambda(x))\}$. Then, for any $\varepsilon > 0$ there exists some $\delta > 0$ such that for any $t \in]0, \delta[$ and for any $y \in B$

$$\langle \lambda^{-1}J(x - p_\lambda(x)), ty\rangle \leq e_\lambda f(x + ty) - e_\lambda f(x) + \varepsilon t,$$

hence,

$$t^{-1}[e_\lambda f(x + ty) - e_\lambda f(x)] - \langle \lambda^{-1}J(x - p_\lambda(x)), y\rangle \geq -\varepsilon. \tag{1.23}$$

At the same time, taking δ smaller if necessary and using the definition of $e_\lambda f$ and the fact that the function $\frac{1}{2}\|\cdot\|^2$ is Fréchet differentiable with

$$J(x - p_\lambda(x)) = \nabla^F\left(\frac{1}{2}\|\cdot\|^2\right)(x - p_\lambda(x)),$$

we have for any $y \in B$

$$\begin{aligned} e_\lambda f(x + ty) - e_\lambda f(x) &\leq f(p_\lambda(x)) + \frac{1}{2\lambda}\|x + ty - p_\lambda(x)\|^2 - f(p_\lambda(x)) - \frac{1}{2\lambda}\|x - p_\lambda(x)\|^2 \\ &\leq \langle \lambda^{-1}J(x - p_\lambda(x)), ty\rangle + \varepsilon t, \end{aligned}$$

i.e.,

$$t^{-1}[e_\lambda f(x + ty) - e_\lambda f(x)] - \langle \lambda^{-1}J(x - p_\lambda(x)), y\rangle \leq \varepsilon. \tag{1.24}$$

Combining (1.23) and (1.24) we obtain the Fréchet differentiability of $e_\lambda f$ at x as well as the equality $\nabla^F e_\lambda f(x) = \lambda^{-1}J(x - p_\lambda(x))$. □

The differentiability of the Moreau envelopes is also related to their regularity. This connection given by the equivalence between assertions (a) and (b) of the next lemma will be needed in Proposition 1.3.19. The lemma also establishes, in view of Theorem 1.3.13, the differentiability of $e_\lambda f$ under the single valuedness and continuity of $P_\lambda f$. Before giving its statement, let us recall that a function f is Fréchet regular at x provided $\partial_F f(x) = \partial_C f(x)$, where ∂_C stands for the Clarke subdifferential.

Lemma 1.3.10. Under the assumptions of Lemma 1.3.9, for any open subset U of X, the equivalences (a)$\Leftrightarrow$(b) and (c)$\Leftrightarrow$(d) hold for the following properties:
(a) $e_\lambda f$ is Fréchet regular on U;
(b) $e_\lambda f$ is Fréchet differentiable on U and its Fréchet derivative $\nabla^F e_\lambda f : U \to X^*$ is norm-to-weak* continuous;
(c) $e_\lambda f$ is continuously Fréchet differentiable on U (and hence (a) and (b) hold);
(d) $P_\lambda f$ is a single-valued norm-to-norm continuous mapping on U.

In any one of these cases, $e_\lambda f$ is Fréchet differentiable on U with $\nabla^F e_\lambda f(x) = \lambda^{-1} J(x - P_\lambda f(x))$.

Proof. (a)$\Rightarrow$(b): If $\partial_F e_\lambda f = \partial_C e_\lambda f$ on U, then we have that $\partial_F e_\lambda f(x) \neq \emptyset$ for any $x \in U$ and, hence, $e_\lambda f$ is Fréchet differentiable on U from Lemma 1.3.9. Moreover, $\partial_C e_\lambda f(x) = \{\nabla^F e_\lambda f(x)\}$ for any $x \in U$, and by the norm-to-weak* upper semicontinuity of $\partial_C e_\lambda f$ we have that $\nabla^F e_\lambda f$ is norm-to-weak* continuous.
(b)$\Rightarrow$(a): Conversely, if $e_\lambda f$ is Fréchet differentiable and norm-to-weak* continuous on U, we have that $\partial_F e_\lambda f(x) = \{\nabla^F e_\lambda f(x)\} = \partial_L e_\lambda f(x)$ for any $x \in U$, where for a locally Lipschitz continuous function $g : X \to \mathbb{R}$ the *limiting subdifferential* $\partial_L g$ is defined as the weak* sequential outer limit

$$w^* - \limsup_{y \to x} \partial_F g(y) := \{w^* - \lim x_n^* : \ x_n^* \in \partial_F g(x_n), x_n \to x\}. \tag{1.25}$$

By Mordukhovich and Shao [126], we know that $\partial_C g(x) = \overline{\text{co}}^* \partial_L g(x)$, where $\overline{\text{co}}^*$ denotes the weak* closed convex hull in X^*. Thus, we obtain that $\partial_F e_\lambda f(x) = \partial_C e_\lambda f(x)$ for $x \in U$, which is the assertion (a).
(c)$\Rightarrow$(d): The continuous differentiability of $e_\lambda f$ on U implies via Lemma 1.3.9 the single valuedness and norm-to-norm continuity of $P_\lambda f$, i.e., the implication holds.
(d)$\Rightarrow$(c): Assume now that $P_\lambda f$ is a single-valued norm-to-norm continuous mapping on U. This continuity property along with (1.22) and (1.25) gives $\partial_C e_\lambda f(x) = \{\lambda^{-1} J(x - P_\lambda f(x))\}$ which entails that $e_\lambda f$ is Gâteaux differentiable on U with $\nabla^G e_\lambda f(x) = \lambda^{-1} J(x - P_\lambda f(x))$. The norm-to-norm continuity of $P_\lambda f$ once again yields the existence of $\nabla^F e_\lambda f$ as well as its norm-to-norm continuity on U. The proof of the lemma is then complete. □

In the remainder of this subsection, we fix a point $\overline{x} \in \operatorname{dom} f$ and $\rho > 0$ such that f is *bounded from below* over $B[\overline{x}, 4\rho]$ and hence we also fix $W = B[\overline{x}, 4\rho]$. Note that according to the lower semicontinuity of f one always has some $\rho > 0$ with the desired property. So it is natural to write $e_{\lambda,\rho,\overline{x}}f(x)$ in place of $e_{\lambda,W}f(x)$ and when $\overline{x}$ and ρ with the above mentioned properties are fixed, it will be convenient to keep as above for index only λ since this will not cause any confusion.

By the useful localization lemma (see Thibault and Zagrodny [160, Lemma 4.2]), there exists some $\lambda_0 > 0$ such that for all $\lambda \in]0, \lambda_0]$

$$P_\lambda f(x) \subset B(\overline{x}, 3\rho) \quad \text{for all} \quad x \in U := B(\overline{x}, \rho). \tag{1.26}$$

So, for any $x \in U \cap G_\lambda$ the unique element $p_\lambda(x)$ of $P_\lambda f(x)$ belongs to $B(\overline{x}, 3\rho)$ and then by (1.22)

$$\nabla^F e_\lambda f(x) = \lambda^{-1} J(x - p_\lambda(x)) \in \partial_F f(p_\lambda(x)) \quad \forall x \in U \cap G_\lambda \tag{1.27}$$

and moreover

$$\|x - p_\lambda(x)\| \le \|x - \overline{x}\| + \|\overline{x} - p_\lambda(x)\| < \rho + 3\rho = 4\rho. \tag{1.28}$$

In fact, we can make (1.27) more precise by proving in the following lemma that the stronger inclusion $\nabla^F e_\lambda f(x) \in \partial_p f(p_\lambda(x))$ holds for $x \in U \cap G_\lambda$.

Lemma 1.3.11. For any $\lambda \in]0, \lambda_0]$, $x \in U \cap \operatorname{Dom} P_\lambda$, and $p_\lambda(x) \in P_\lambda f(x)$, we have that $\lambda^{-1} J(x - p_\lambda(x)) \in \partial_p f(p_\lambda(x))$. In other words, for any $x \in U$ and any $\lambda \in]0, \lambda_0]$,

$$P_\lambda f(x) \subset (I + \lambda J^* \circ \partial_p f)^{-1}(x).$$

Proof. Fix any $0 < \lambda \le \lambda_0$, $x \in U \cap \operatorname{Dom} P_\lambda$, and $p_\lambda(x) \in P_\lambda f(x)$. Then,

$$f(p_\lambda(x)) + (2\lambda)^{-1}\|x - p_\lambda(x)\|^2 \le f(y) + (2\lambda)^{-1}\|x - y\|^2, \quad \forall y \in W,$$

that is,

$$(2\lambda)^{-1}\|x - p_\lambda(x)\|^2 - (2\lambda)^{-1}\|x - y\|^2 \le f(y) - f(p_\lambda(x)), \quad \forall y \in W.$$

Since $p_\lambda(x) \in B[\overline{x}, 3\rho]$ the last inequality holds true in particular for all $y \in B[p_\lambda(x), \rho]$. Let us set $p := \lambda^{-1}(x - p_\lambda(x))$. From the last inequality we have

$$2^{-1}\lambda\|p\|^2 - (2\lambda)^{-1}\|x - y\|^2 \le f(y) - f(p_\lambda(x)), \quad \forall y \in B[p_\lambda(x), \rho],$$

which entails

$$\frac{\lambda^2}{2}\|p\|^2 - \frac{1}{2}\|\lambda p + p_\lambda(x) - y\|^2 \le \lambda[f(y) - f(p_\lambda(x))], \quad \forall y \in B[p_\lambda(x), \rho],$$

or,

$$\lambda^2\|p\|^2 - \|\lambda p + p_\lambda(x) - y\|^2 \le 2\lambda[\beta - f(p_\lambda(x))], \quad \forall (y, \beta) \in \operatorname{epi} f \text{ with } y \in B[p_\lambda(x), \rho].$$

Adding λ^2 to both sides yields

$$\lambda^2\|p\|^2-\|\lambda p+p_\lambda(x)-y\|^2+\lambda^2 \leq 2\lambda[\beta-f(p_\lambda(x))]+\lambda^2,\ \forall(y,\beta)\in \operatorname{epi} f \text{ with } y\in B[p_\lambda(x),\rho],$$

and using the inequality $2\lambda[\beta - f(p_\lambda(x))] + \lambda^2 \leq [\beta - f(p_\lambda(x)) + \lambda]^2$ we obtain that

$$\lambda^2\|p\|^2+\lambda^2 \leq \|\lambda p+p_\lambda(x)-y\|^2+[\beta-f(p_\lambda(x))+\lambda]^2, \quad \forall(y,\beta)\in \operatorname{epi} f \text{ with } y\in B[p_\lambda(x),\rho].$$

So, we obtain that for all $(y,\beta)\in \operatorname{epi} f$ with $y\in B[p_\lambda(x),\rho]$

$$|||\lambda(p,-1)|||\leq|||(p_\lambda(x), f(p_\lambda(x)))+\lambda(p,-1)-(y,\beta)|||.$$

By (1.14) this inequality entails that $(p,-1) \in PN_{\operatorname{epi} f}(p_\lambda(x), f(p_\lambda(x)))$, which gives (see Subsection 1.3.1) that $J(p) \in \partial_p f(p_\lambda(x))$. This means that

$$\lambda^{-1}J(x - p_\lambda(x)) \in \partial_p f(p_\lambda(x)),$$

which entails the inclusion of the lemma. □

Remark 1.3.12. (a) In the case when the lower semicontinuous function f is the indicator function ψ_C of a non-empty closed set C, the above conclusions hold for $W = X$ and any $\lambda > 0$ or for $\rho = +\infty$, any $\overline{x} \in C$, and any $\lambda > 0$. Further with $\rho = +\infty$ one has $e_\lambda f(x) = \frac{1}{2\lambda}d_C^2(x)$ and $P_\lambda f(x) = P_C(x)$ for all $x \in X$.
(b) Still with $f = \psi_C$, for any $\overline{x} \in C$, any $\rho \in [0,+\infty]$, and any $\lambda > 0$ one has $e_\lambda f(x) = \frac{1}{2\lambda}d_{C\cap W}^2(x)$ and $P_\lambda f(x) = P_{C\cap W}(x)$ for all $x \in X$.

But for any $x \in B[\overline{x}, 2\rho]$ and any $y \in C \setminus W$,

$$\|x - y\| \geq \|y - \overline{x}\| - \|x - \overline{x}\| > 4\rho - 2\rho = 2\rho \geq \|x - \overline{x}\| \geq d_{C\cap W}(x).$$

Hence, for $x \in B[\overline{x}, 2\rho]$, $d_{C\cap W}(x) = d_C(x)$ and $P_{C\cap W}(x) = P_C(x)$. Therefore, $e_\lambda f(x) = \frac{1}{2\lambda}d_C^2(x)$ and $P_\lambda f(x) = P_C(x)$ for any $x \in U := B(\overline{x},\rho)$.

Recall that a function g is of class $C^{1,\alpha}$ on an open set $U \subset X$ when it is differentiable on U and the derivative ∇g is locally α-Hölder continuous on U.

Theorem 1.3.13. Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous function which is J-plr at $\overline{x} \in \operatorname{dom} f$ with positive real numbers ε and r, such that $\varepsilon < 1 < \frac{1}{r}$. Let $\rho \in \left]0, \frac{r\varepsilon}{16}\right]$ be fixed in such a way that f is bounded from below on $B[\overline{x}, 4\rho]$. Then there exists $\lambda_0 > 0$ such that for any $\lambda \in]0, \lambda_0]$ the map $x \mapsto P_\lambda f(x)$ is single-valued on $U := B(\overline{x},\rho)$ with, for some constant $\gamma \geq 0$,

$$\|p_\lambda(x) - p_\lambda(x')\| \leq \gamma\|x - x'\|^{\frac{1}{q}}, \quad \forall x, x' \in U, \tag{1.29}$$

where p_λ is given by $P_\lambda f(y) = \{p_\lambda(y)\}$ for any $y \in U$. Moreover, for each $\lambda \in]0, \lambda_0]$ the function $e_\lambda f$ is of class $C^{1,\alpha}$ on U with $\alpha := q^{-1}(s-1)$ and $\nabla^F e_\lambda f(x) = \lambda^{-1}J(x-p_\lambda(x))$ for all $x \in U$.

Proof. Let λ_0 be given by the analysis of (1.26) above for ρ fixed as in the statement of the theorem. We will work with arbitrary fixed $\lambda \in]0, \lambda_0]$. Put $c := r\varepsilon$ and $T_{ct} := (\partial_p f)_{\overline{x},\varepsilon,ct}$ for any $t \geq 0$. The proof is divided in three steps.

Step 1. Let us prove that $P_\lambda f$ is $\frac{1}{q}$-Hölder continuous on $U \cap \operatorname{Dom} P_\lambda f$.
We have from Definition 1.3.6 and from (1.19) that the set-valued mapping T_{ct} is J-hypomonotone of degree t for any $t \geq 0$. Hence, for $t_\lambda := c/(4\lambda)$, the set-valued mapping T_{t_λ} is J-hypomonotone of degree $\overline{r} := 1/(4\lambda)$. As $\lambda^{-1} > 2\overline{r}$, Lemma 1.3.8 entails that $(I + \lambda J^* \circ T_{t_\lambda})^{-1}$ is a single-valued mapping on its domain and this mapping is $\frac{1}{q}$-Hölder continuous on the intersection of its domain with any bounded subset of X. From Lemma 1.3.11, we have that $P_\lambda f(x) \subset (I + \lambda \partial_p f)^{-1}(x)$ for any $x \in U$. We claim that we even have

$$P_\lambda f(x) \subset (I + \lambda J^* \circ T_{t_\lambda})^{-1}(x) \quad \text{for any } x \in U. \tag{1.30}$$

Indeed, fixing any $x \in U \cap \operatorname{Dom} P_\lambda f$ and $p_\lambda(x) \in P_\lambda f(x)$, we know that $p_\lambda(x) \in B[\overline{x}; 3\rho]$. So $\|p_\lambda(x) - \overline{x}\| < \varepsilon$, and

$$\|\lambda^{-1} J(x - p_\lambda(x))\| \leq \lambda^{-1}\|x - p_\lambda(x)\| \leq \lambda^{-1}(\|x - \overline{x}\| + \|\overline{x} - p_\lambda(x)\|) \leq 4\rho\lambda^{-1} \leq (c\lambda^{-1})/4 = t_\lambda.$$

Hence, as also $\lambda^{-1} J(x - p_\lambda(x)) \in \partial_p f(p_\lambda(x))$, we have that $\lambda^{-1} J(x - p_\lambda(x)) \in T_{t_\lambda}(p_\lambda(x))$, which proves the claim. Therefore we obtain that $P_\lambda f$ is a single-valued $\frac{1}{q}$-Hölder continuous mapping on $U \cap \operatorname{Dom} P_\lambda f$, that is, there exists some constant $\gamma \geq 0$ such that for all $x, x' \in U \cap \operatorname{Dom} P_\lambda f$

$$\|p_\lambda(x) - p_\lambda(x')\| \leq \gamma \|x - x'\|^{\frac{1}{q}}. \tag{1.31}$$

Step 2. Let us prove that $U \subset \operatorname{Dom} P_\lambda f$.
The proof now is similar to that of Bernard and Thibault [15, 14]. Take any $x \in U$ and fix some integer $k \geq 1$ with $B(x, 1/k) \subset U$. According to the density of the Fréchet subdifferentiability points of $e_\lambda f$, for any integer $n \geq k$ there exists some point $x_n \in G_\lambda \cap B(x, 1/n)$. By (1.31), for any integers $n, m \geq k$,

$$\|p_\lambda(x_n) - p_\lambda(x_m)\| \leq \gamma \|x_n - x_m\|^{\frac{1}{q}}.$$

Hence, $(p_\lambda(x_n))_n$ is a Cauchy sequence. Denote by z_λ its limit. By the definition of $p_\lambda(x_n)$ we have

$$f(p_\lambda(x_n)) + \frac{1}{2\lambda}\|x_n - p_\lambda(x_n)\|^2 \leq f(y) + \frac{1}{2\lambda}\|x_n - y\|^2, \quad \forall y \in W.$$

Since f is lower semicontinuous, the latter implies

$$f(z_\lambda) + \frac{1}{2\lambda}\|x - z_\lambda\|^2 \leq f(y) + \frac{1}{2\lambda}\|x - y\|^2, \quad \forall y \in W.$$

This means that $z_\lambda \in P_\lambda f(x)$, which yields $U \cap \operatorname{Dom} P_\lambda f = U$. Hence (1.31) holds for all $x, x' \in U$ and, through Step 1, $P_\lambda f$ is a single-valued $\frac{1}{q}$-Hölder continuous mapping on U. Then by Lemma 1.3.10, the envelope $e_\lambda f$ is continuously Fréchet differentiable on U with $\nabla^F e_\lambda f(x) = \lambda^{-1} J(x' - p_\lambda(x))$ for any $x \in U$.

Step 3. We will prove that $e_\lambda f$ is of class $C^{1,\alpha}$ on U with $\alpha = q^{-1}(s-1)$. Let us take any $x, x' \in U$ and, for $r := 4\rho$, use (1.9) to estimate

$$\|\nabla^F e_\lambda f(x) - \nabla^F e_\lambda f(x')\| = \lambda^{-1}\|J(x - p_\lambda(x)) - J(x' - p_\lambda(x'))\| \le$$

$$\lambda^{-1} K'_r \|x - p_\lambda(x) - x' + p_\lambda(x')\|^{s-1} \le \lambda^{-1} K'_r [\|x - x'\| + \|p_\lambda(x) - p_\lambda(x')\|]^{s-1} \le$$

$$\lambda^{-1} K'_r [\|x - x'\| + \gamma \|x - x'\|^{\frac{1}{q}}]^{s-1} \le \lambda^{-1} K'_r (1+\gamma)^{s-1} \|x - x'\|^{\frac{s-1}{q}},$$

where the third inequality is due to (1.29) and the last one to the fact that $\|x - x'\| < 1$. The proof of the theorem is then complete. □

We now state in the next proposition the relation obtained between the proximal mappings and some truncations of the subdifferential of a J-plr function.

Proposition 1.3.14. Under the assumptions of Theorem 1.3.13, one has for all $\lambda \in]0, \lambda_0]$ and $x \in U$

$$P_\lambda f(x) = (I + \lambda^{-1} J^* \circ T_{t_\lambda})^{-1}(x),$$

where $t_\lambda := \varepsilon r/(4\lambda)$ and $T_{t_\lambda} := (\partial_P f)_{\overline{x},\varepsilon,t_\lambda}$.

Proof. The inclusion of the first member in the second one is (1.30) of Theorem 1.3.13, and the reverse one follows from the inclusion $U \subset \operatorname{Dom} P_\lambda f$ established in Step 2 of the same proof since $(I + \lambda^{-1} J^* \circ T_{t_\lambda})^{-1}$ is at most single-valued as we saw in Step 1 above. □

Corollary 1.3.15. Let $C \subset X$ be a non-empty closed set such that its indicator function ψ_C is J-plr at $\overline{x} \in C$ with positive real numbers ε and r satisfying $\varepsilon < 1 < \frac{1}{r}$. Then for $\rho = \frac{\varepsilon r}{16}$ the mapping $x \mapsto P_C(x)$ is single-valued on $U = B(\overline{x}, \rho)$ and for some constant $\gamma \ge 0$

$$\|P_C(x) - P_C(x')\| \le \gamma \|x - x'\|^{\frac{1}{q}}, \quad \forall x, x' \in U. \tag{1.32}$$

Further, the function d_C^2 is of class $C^{1,\alpha}$ on U with $\alpha = q^{-1}(s-1)$ and $\nabla^F d_C^2(x) = 2J(x - P_C(x))$ for all $x \in U$.

Corollary 1.3.16. Under the assumptions of Corollary 1.3.15, we have that

$$P_C(x) = (I + J^* \circ N_C^{P\sigma})^{-1}(x) \quad \forall x \in U,$$

where $\sigma := \varepsilon r/4$ and the set-valued mapping $N_C^{P\sigma}$ is defined by (1.20).

Proof. Taking $\rho = \frac{\varepsilon r}{16}$ as in the statement of Corollary 1.3.15, since $r < 1$ it is easily checked for $T := \partial_p \psi_C$ that for all $x \in U = B(\overline{x}, \rho)$

$$\left(I + J^* \circ T_{\overline{x},\varepsilon,\frac{\varepsilon r}{4}}\right)^{-1}(x) = \left(I + J^* \circ N_C^{P\,\varepsilon r/4}\right)^{-1}(x).$$

Soq it suffices to apply Proposition 1.3.14 with $\lambda = 1$ to $f = \psi_C$, keeping in mind Remark 1.3.12(b). □

1.3.3 Characterizations of prox-regular sets

In this subsection we will give different characterizations of prox-regularity of a set.

Lemma 1.3.17. Let $C \subset X$ be a closed subset. Then the single valuedness and norm-to-weak continuity of the projection mapping P_C over an open set U imply its norm-to-norm continuity on U.

Proof. Let $u_n \xrightarrow[n\to\infty]{\|\cdot\|} u$ and $P_C(u_n) \xrightarrow[n\to\infty]{w} P_C(u)$. From the Lipschitz continuity of the distance function, $\|u_n - P_C(u_n)\| = d_C(u_n) \xrightarrow[n\to\infty]{} d_C(u) = \|u - P_C(u)\|$. By the Kadec property of the norm, $P_C(u_n) \xrightarrow[n\to\infty]{\|\cdot\|} P_C(u)$. □

The following proposition establishes that the continuity of the metric projection mapping to a closed set C is equivalent to the continuous differentiability of the distance function d_C, as shown in the Hilbert setting in Poliquin, Rockafellar and Thibault [138], where it is proved that those properties characterize the prox-regularity of a set in a Hilbert space. Its proof follows directly from Lemma 1.3.10 with $f = \psi_C$.

Proposition 1.3.18. Let $C \subset X$ be a closed set and $U \subset X$ be an open set. Then the following are equivalent:
(a) P_C is single-valued and norm-to-norm continuous on U;
(b) d_C^2 is of class C^1 on U.

In fact these properties are equivalent to the only Fréchet subdifferentiability of the distance function as we can see from the following proposition.

Proposition 1.3.19. For any closed set $C \subset X$ and any open set U of X the following are equivalent:
(a) d_C is continuously differentiable on $U \setminus C$;
(b) $\partial_F d_C(x)$ is non-empty for all $x \in U$;
(c) $\partial_F d_C^2(x)$ is non-empty at all points x in U;

(d) d_C is Fréchet differentiable on $U \setminus C$;
(e) d_C is Fréchet regular on $U \setminus C$;
(f) d_C is Gâteaux differentiable on $U \setminus C$ with $\|\nabla^G d_C(x)\| = 1$ for all $x \in U \setminus C$.

Proof. (a) ⇒ (b) is obvious since one always has $0 \in \partial_F d_C(u)$ for any $u \in C$.
(b) ⇒ (c) follows from the fact that for any $x^* \in \partial_F d_C(x)$, one has $2d_C(x)x^* \in \partial_F d_C^2(x)$ according to Lemma 3.9 in Poliquin, Rockafellar and Thibault [138].
(c) ⇒ (d): By Lemma 1.3.9 and Remark 1.3.12(a) we have that d_C^2 is Fréchet differentiable on U, hence so is d_C on $U \setminus C$.
(d) ⇒ (a): It is clear that (d) entails (b) and hence (c). From Lemma 1.3.9 and Remark 1.3.12(a) we get the Fréchet differentiability of d_C^2 on U. The latter implies that P_C is single-valued on U and that

$$\nabla^F d_C^2(x) = 2J(x - P_C(x)) \quad \text{for any } x \in U. \tag{1.33}$$

So, it remains to prove the norm-to-norm continuity of P_C over U and we will obtain that of $\nabla^F d_C^2$. Take any $x_0 \in U$, and $U \ni x_n \xrightarrow[n\to\infty]{\|\cdot\|} x_0$. Then we also have that

$$\|x_n - P_C(x_n)\| = d_C(x_n) \xrightarrow[n\to\infty]{} d_C(x_0) = \|x_0 - P_C(x_0)\|,$$

which entails that the sequence $(P_C(x_n))_n$ is bounded. Taking if necessary a subsequence, we may suppose that $P_C(x_n) \xrightarrow[n\to\infty]{w} z$ for some $z \in X$. As $\|x_0 - P_C(x_n)\| \xrightarrow[n\to\infty]{} \|x_0 - P_C(x_0)\|$, having in mind the Kadec property of the norm, it suffices to prove that

$$\|x_0 - z\| = \|x_0 - P_C(x_0)\| \tag{1.34}$$

to get that $P_C(x_n) \xrightarrow[n\to\infty]{\|\cdot\|} z$, and hence that $z \in C$. Then, using (1.34) and the fact that $P_C(x_0)$ is single-valued, we will have that $P_C(x_0) = \{z\}$, so P_C is norm-to-norm continuous at x_0. To this end, following an idea of Borwein and Giles from [25], let us set $t_n^2 := \|x_0 - P_C(x_n)\|^2 - d_C^2(x_0)$. If $t_n = 0$ then $P_C(x_n) = P_C(x_0)$ due to the single-valuedness of P_C, and there would be nothing to prove if this equality holds for infinitely many n, so we may suppose that $t_n > 0$ for any integer n. Fix any $\varepsilon > 0$. From the Fréchet differentiability of d_C^2 at x_0 it follows that

$$\begin{aligned}
\langle \nabla^F d_C^2(x_0), P_C(x_n) - x_0 \rangle &\le \frac{d_C^2(x_0 + t_n(P_C(x_n) - x_0)) - d_C^2(x_0)}{t_n} + \frac{\varepsilon}{4} \\
&\le \frac{\|x_0 + t_n(P_C(x_n) - x_0) - P_C(x_n)\|^2 - d_C^2(x_0)}{t_n} + \frac{\varepsilon}{4} \\
&\le \frac{(1 - t_n)^2 \|x_0 - P_C(x_n)\|^2 - d_C^2(x_0)}{t_n} + \frac{\varepsilon}{4}
\end{aligned}$$

and hence for n large enough

$$\begin{aligned}\langle \nabla^F d_C^2(x_0), P_C(x_n) - x_0\rangle &\le \frac{\|x_0 - P_C(x_n)\|^2 - d_C^2(x_0)}{t_n} - 2\|x_0 - P_C(x_n)\|^2 + \frac{\varepsilon}{2} \\ &= t_n - 2\|x_0 - P_C(x_n)\|^2 + \frac{\varepsilon}{2} \\ &\le -2d_C^2(x_0) + \varepsilon.\end{aligned}$$

Passing to the limit, we obtain $\langle \nabla^F d_C^2(x_0), x_0 - z\rangle \ge 2d_C^2(x_0)$, or,

$$\langle 2J(x_0 - P_C(x_0)), x_0 - z\rangle \ge 2\|x_0 - P_C(x_0)\|^2,$$

which entails that $\|x_0 - z\| \ge \|x_0 - P_C(x_0)\|$. Since $\|x_n - P_C(x_n)\| \le \|x_n - P_C(x_0)\|$, one has that $\liminf_{n\to\infty} \|x_n - P_C(x_n)\| \le \lim_{n\to\infty} \|x_n - P_C(x_0)\|$, and hence, by the weak lower semicontinuity of the norm, one gets $\|x_0 - z\| \le \|x_0 - P_C(x_0)\|$. Finally, $\|x_0 - z\| = \|x_0 - P_C(x_0)\|$, that is, (1.34), and hence the implication (d) $\Rightarrow$ (a) holds.
The implications (e) $\Rightarrow$ (d) and (a) $\Rightarrow$ (e) follow from Lemma 1.3.10 and the equivalence (d) $\Leftrightarrow$ (f) is a direct consequence of Theorem 2.4 of Fitzpatrick [75]. The proof of the proposition is then complete. □

In addition to Proposition 1.3.19 we state the following theorem providing a weak derivability condition on d_C under which P_C is continuous, supposing it is nonempty-valued. It is a direct consequence of Theorem 2.4 in Fitzpatrick [75] as observed in Corollary 2 of Borwein, Fitzpatrick and Giles [24]. Note that in those works, the framework is beyond the uniform convexity. Further, Fitzpatrick's important condition of Fréchet differentiability concerns general functions. The result of the theorem will be used in the proof of Theorem 1.3.25.

Recall that, for $v \in X$, a function $f : X \to \mathbb{R}$ has a *Gâteaux directional derivative* at a point x in the full direction v provided that the limit $\lim_{t\to 0} t^{-1}[f(x + tv) - f(x)]$ exists and is finite.

Theorem 1.3.20. Let $C \subset X$ be a closed set and $x \in X \setminus C$ be such that $P_C(x) \neq \emptyset$. If d_C has a Gâteaux directional derivative at x in the full direction $x - p(x)$ for some $p(x) \in P_C(x)$, then d_C is Fréchet differentiable at x.

Another interest of this result will appear in the proof of Theorem 1.3.35.

We now proceed to establish two lemmas. The first one is a key result proved in Lemma 3.3 of Poliquin, Rockafellar and Thibault [138] in the Hilbert context. The proof is valid in our setting, and we sketch the main parts below.

Lemma 1.3.21. Let C be a closed subset of X. Assume that d_C is Fréchet differentiable on a neighbourhood of a point $\overline{u} \notin C$. Then there exists $\delta > 0$ such that whenever $u \in B(\overline{u}, \delta)$ and $P_C(u) = x$, there exists some $t > 0$ such that the point $u_t := u + t(u - x)$ likewise has $P_C(u_t) = x$.

Proof. By Proposition 1.3.19 and Proposition 1.3.18, there exists $\varepsilon > 0$ such that P_C is single-valued and norm-to-norm continuous on $B(\overline{u}, 2\varepsilon)$, with d_C continuously Fréchet differentiable on this ball as well. For each $u \in B(\overline{u}, \varepsilon)$ and each $t > 0$ put $u_t := u + t(u - P_C(u))$. Following the proof of Lemma 3.3 in Poliquin, Rockafellar and Thibault [138], we find out some positive numbers $\delta < \varepsilon$ and $s < 1$ such that for all $u \in B(\overline{u}, \delta)$ one has $d_C(u) \geq \delta$, $sd_C(u) < \delta$ and $d_C(u_s) > d_C(u)$. Fix now $u \in B(\overline{u}, \delta)$ and consider the closed set $D := \{w \in X : d_C(w) \geq d_C(u_s)\}$. As $u \notin D$, according to Lau's theorem (see Lau [108]) there is a sequence $D \not\ni y_n \xrightarrow[n\to\infty]{\|\cdot\|} u$ with $P_D(y_n) \neq \emptyset$. Choosing $w_n \in P_D(y_n)$ we have $d_C(w_n) = d_C(u_s)$ (because w_n is a boundary point of D). For all n large enough, $w_n \in B(\overline{u}, 2\delta)$ since

$$\|y_n - w_n\| = d_D(y_n) \leq \|y_n - u_s\| \underset{n\to\infty}{\longrightarrow} \|u - u_s\| = s\|u - P_C(u)\| = sd_C(u) < \delta. \tag{1.35}$$

Consequently d_C is Fréchet differentiable at w_n and by (1.33) we have

$$\nabla^F d_C(w_n) = J(w_n - P_C(w_n))/d_C(w_n) \quad \text{and} \quad \|\nabla^F d_C(w_n)\| = 1.$$

Therefore, the half-space $E := \{v \in X : \langle -\nabla^F d_C(w_n), v\rangle \leq 0\}$ gives the Clarke tangent cone to D at w_n (see Clarke [44]) and hence its negative polar cone $-[0, \infty[\, \nabla^F d_C(w_n)$ is the Clarke normal cone to D at w_n. The nonzero functional $J(y_n - w_n)$ being a proximal normal functional to D at w_n, it belongs to the Clarke normal cone to D at w_n. Consequently, there exists some $\lambda_n > 0$ such that $J(y_n - w_n) = -\lambda_n \nabla^F d_C(w_n)$ which entails

$$y_n - w_n = -\lambda_n(w_n - P_C(w_n))/d_C(w_n) \quad \text{and} \quad \lambda_n = \|y_n - w_n\|.$$

For n large enough, we have by (1.35) that $\lambda_n < \delta$ and hence

$$\lambda_n < \delta \leq d_C(u) < d_C(u_s) = d_C(w_n).$$

It follows that for $\alpha_n := \lambda_n / d_C(w_n)$ we have $\alpha_n \in]0, 1[$ and $y_n = (1 - \alpha_n)w_n + \alpha_n P_C(w_n)$. Hence, $P_C(y_n) = P_C(w_n)$ and

$$\lambda_n = \|y_n - w_n\| = d_C(w_n) - d_C(y_n) = d_C(u_s) - d_C(y_n).$$

Putting $t_n := \dfrac{\alpha_n}{1 - \alpha_n} = \dfrac{d_C(u_s) - d_C(y_n)}{d_C(y_n)}$, we obtain $w_n = y_n + t_n(y_n - P_C(y_n))$. As $(t_n)_n$ converges to $t := (d_C(u_s) - d_C(u))/d_C(u) > 0$, we have $w_n \xrightarrow[n\to\infty]{\|\cdot\|} u_t$ and $u_t \in B(\overline{u}, 2\delta)$ by (1.35) and by the inclusion $u \in B(\overline{u}, \delta)$. So, by continuity of P_C over $B(\overline{u}, 2\delta)$ we get $P_C(w_n) \xrightarrow[n\to\infty]{\|\cdot\|} P_C(u_t)$. But we also have $P_C(w_n) = P_C(y_n) \xrightarrow[n\to\infty]{\|\cdot\|} P_C(u)$. Finally, for this number t we have $P_C(u_t) = P_C(u)$ and hence the proof is complete. □

The next lemma follows from the previous one.

Lemma 1.3.22. Let $C \subset X$ be a closed subset of the space X and $\overline{x} \in C$. If the mapping P_C is single-valued and norm-to-norm continuous in a neighbourhood U of $\overline{x}$, then there exists some $\varepsilon > 0$ such that for all $x \in C \cap B(\overline{x}, \varepsilon)$ and all $p \in PN_C(x)$ with $p \neq 0$ the equality $P_C\left(x + \varepsilon\frac{p}{\|p\|}\right) = x$ holds.

Proof. Let P_C be single-valued and norm-to-norm continuous in $B(\overline{x}, \delta)$. Take $\varepsilon < \delta/2$ and consider any non-zero $p \in PN_C(x)$ with $\|x-\overline{x}\| < \varepsilon$. By definition of the proximal normal cone and by Lemma 1.3.1, there exists $\lambda > 0$ such that $P_C(x + \lambda p) = x$. Set $\lambda_s := \sup\left\{\lambda \leq \varepsilon : P_C\left(x + \lambda\frac{p}{\|p\|}\right) = x\right\}$. By the continuity of P_C on $B(\overline{x}, \delta)$ we have that $P_C\left(x + \lambda_s\frac{p}{\|p\|}\right) = x$. Suppose that $\lambda_s < \varepsilon$. As $x + \lambda_s\frac{p}{\|p\|}$ belongs to the open set $B(\overline{x}, \delta)$ where d_C is Fréchet differentiable according to Proposition 1.3.18, by Lemma 1.3.21 there exists $\eta > 0$ with $\lambda_s + \eta \leq \varepsilon$ such that $P_C\left(x + (\lambda_s + \eta)\frac{p}{\|p\|}\right) = x$. This gives a contradiction with the definition of λ_s. Hence $\lambda_s = \varepsilon$. □

Lemma 1.3.22 allows us to establish the following proposition which prepares the theorem on characterizations of prox-regularity.

Proposition 1.3.23. Let $C \subset X$ be a closed subset of the space X. The following assertions are equivalent:
(a) C is prox-regular at $\overline{x} \in C$;
(b) there exists $\varepsilon > 0$ such that the condition $\left.\begin{array}{l} x=P_C(u), x \neq u \\ 0 < \|u-\overline{x}\| < \varepsilon \end{array}\right\}$ implies that $x=P_C(u')$ for $u':=x+\varepsilon\frac{u-x}{\|u-x\|}$;
(c) there exists $\varepsilon > 0$ such that $p \in PN_C(x)$ with $x \in B(\overline{x}, \varepsilon)$ and $p \neq 0$ imply that

$$P_C\left(x + \varepsilon\frac{p}{\|p\|}\right) = x.$$

Proof. (a) $\Rightarrow$ (b): If C is prox-regular at $\overline{x}$, then P_C is single valued and Hölder continuous in a neighbourhood U of $\overline{x}$ according to Corollary 1.3.15. From the continuity of P_C, for the positive number ε of Lemma 1.3.22, there exists $\varepsilon' \in]0, \varepsilon[$ such that $\left.\begin{array}{l} x = P_C(u), x \neq u \\ 0 < \|u - \overline{x}\| < \varepsilon' \end{array}\right\}$ implies that $\|x-\overline{x}\| = \|P_C(u)-P_C(\overline{x})\| < \varepsilon$. Lemma 1.3.22 and Lemma 1.3.1 ensure for $p = u - x$ that $P_C\left(x + \varepsilon'\frac{u-x}{\|u-x\|}\right) = x$.
(b) $\Rightarrow$ (c): Let us suppose that (b) holds with some $\varepsilon > 0$. If $\|x - \overline{x}\| < \varepsilon/2$ and $p \in PN_C(x)$ with $p \neq 0$, then by definition of $PN_C(x)$ and by Lemma 1.3.1 there exists some $\eta \in]0, \varepsilon/2[$ such that $x = P_C(u)$, where $u := x + \eta\frac{p}{\|p\|}$. We have

$$\|u - \overline{x}\| \leq \|u - x\| + \|x - \overline{x}\| < \varepsilon/2 + \varepsilon/2,$$

and from (b), $P_C\left(x + \varepsilon\frac{u-x}{\|u-x\|}\right) = x$. Hence, one obtains (c) with $\varepsilon/2$.
(c) $\Rightarrow$ (a): We suppose that (c) holds with some $\varepsilon > 0$. Let $0 \neq p^* \in N_C^P(x)$ with

$\|x - \overline{x}\| < \varepsilon$ and $\|p^*\| < \varepsilon$. There exists $u \notin C$ such that $\left\langle \frac{p^*}{\|p^*\|}, u - x \right\rangle = \|u - x\| = d_C(u)$. Thus, $u - x \in PN_C(x)$ and $\frac{p^*}{\|p^*\|} = J\left(\frac{u-x}{\|u-x\|}\right)$. We have by (c) that $P_C\left(x + \varepsilon\frac{u-x}{\|u-x\|}\right) = x$, so $P_C\left(x + \frac{\varepsilon}{\|p^*\|}J^*(p^*)\right) = x$. Now, for all $s \leq 1$ (since $1 \leq \frac{\varepsilon}{\|p^*\|}$), we have that $P_C(x + sJ^*(p^*)) = x$. By Definition 1.3.3 and by Proposition 1.3.4, the set C is prox-regular at $\overline{x}$. □

Now, following Poliquin and Rockafellar [137] and Poliquin, Rockafellar and Thibault [138] we give a subdifferential characterization of the prox-regularity of a set in terms of the truncated cone (see (1.20)) of proximal normal functionals.

Proposition 1.3.24. A set $C \subset X$ is prox-regular at $\overline{x} \in C$, if and only if, for some $\varepsilon, \rho > 0$, the set-valued mapping $N_C^{P\varepsilon} : X \rightrightarrows X^*$ that assigns to each $x \in X$ the truncated cone of proximal normal functionals $N_C^{P\varepsilon}(x)$ is J-hypomonotone of degree ρ on $B(\overline{x}, \varepsilon)$.

Proof. By Proposition 1.3.7, if C is prox-regular at $\overline{x}$ then, for some $\varepsilon > 0$ and $\rho > 0$, the truncated normal functional cone mapping $N_C^{P\varepsilon}$ is J-hypomonotone of degree ρ on $B(\overline{x}, \varepsilon)$. Conversely, suppose that $N_C^{P\varepsilon}$ is J-hypomonotone of degree ρ. Then the argument of Theorem 1.3.13 or Corollary 1.3.15 works as well (since it only makes use of the J-hypomonotonicity of the truncation of $\partial_P f$), to get that P_C is single-valued and continuous on a neighbourhood of $\overline{x}$. It just remains to invoke Lemma 1.3.22 and Proposition 1.3.23 to conclude. □

Now we can state the theorem giving several characterizations of the prox-regularity of a set. Recall first that the (lower) *Dini subdifferential* of a locally Lipschitz continuous function $f : X \to \mathbb{R}$ at a point x is defined by

$$\partial^- f(x) := \left\{ x^* \in X^* : \langle x^*, h \rangle \leq \liminf_{t \downarrow 0} t^{-1}[f(x + th) - f(x)], \ \forall h \in X \right\}.$$

Theorem 1.3.25. Let $C \subset X$ be a closed set. The following are equivalent:
(a) C is prox-regular at $\overline{x}$;
(b) P_C is single-valued and norm-to-norm $\frac{1}{q}$-Hölder continuous on some neighbourhood U of the point $\overline{x}$;
(c) P_C is single-valued and norm-to-weak continuous on some neighbourhood U of $\overline{x}$;
(d) there exists $\varepsilon > 0$ such that $p \in PN_C(x)$ with $x \in B(\overline{x}, \varepsilon)$ and $p \neq 0$ implies that $P_C\left(x + \varepsilon\frac{p}{\|p\|}\right) = x$;
(e) there exists $\varepsilon{>}0$ such that the condition $\left.\begin{array}{c} x{=}P_C(u), x \neq u \\ 0 < \|u{-}\overline{x}\| < \varepsilon \end{array}\right\}$ implies that $x{=}P_C(u')$ for $u'{=}x{+}\varepsilon\frac{u-x}{\|u-x\|}$;
(f) d_C^2 is of class $C^{1,\alpha}$ on some neighbourhood U of $\overline{x}$ with $\alpha = q^{-1}(s - 1)$;
(g) d_C is Fréchet differentiable on $U \setminus C$ for some neighbourhood U of $\overline{x}$;

(h) for some neighbourhood U of $\overline{x}$, the function d_C is Gâteaux differentiable on $U \setminus C$ with $\|\nabla^G d_C(x)\| = 1$ for all $x \in U \setminus C$;
(i) d_C is Fréchet subdifferentiable on U for some neighbourhood U of $\overline{x}$;
(j) d_C is Fréchet regular on $U \setminus C$ for some neighbourhood U of $\overline{x}$;
(k) P_C is nonempty-valued on U and d_C is Dini subdifferentiable on U for some neighbourhood U of $\overline{x}$;
(l) the indicator function ψ_C is J-plr at $\overline{x}$;
(m) there exist $\varepsilon, \rho > 0$ such that the truncated normal functional cone mapping $N_C^{P\varepsilon}$ is J-hypomonotone of degree ρ on $B(\overline{x}, \varepsilon)$.

If C is weakly closed, one has one more equivalent condition
(n) P_C is single-valued on some neighbourhood U of $\overline{x}$.

Proof. First we will establish all the equivalences without specifying the Hölder character of the continuity in (b) and (f). The proof follows the scheme:

$$\begin{array}{ccccccccccccccccccc}
(m) & \Leftrightarrow & (a) & \Rightarrow & (l) & \Rightarrow & (f) & \Leftrightarrow & (j) & \Leftrightarrow & (h) & \Leftrightarrow & (i) & \Leftrightarrow & (g) & \Leftrightarrow & (k) \\
 & & & & & & \Updownarrow & & & & & & & & & & \\
 & & & & (c) & \Leftrightarrow & (b) & \Rightarrow & (d) & \Leftrightarrow & (e) & \Leftrightarrow & (a). & & & &
\end{array}$$

(m) $\Leftrightarrow$ (a) is Proposition 1.3.24.
(a) $\Rightarrow$ (l) is established in Proposition1.3.7.
(l) $\Rightarrow$ (f) follows from Corollary 1.3.15.
(f) $\Leftrightarrow$ (j) $\Leftrightarrow$ (h) $\Leftrightarrow$ (i) $\Leftrightarrow$ (g) is Proposition 1.3.19.
(g) $\Rightarrow$ (k): Assume that (g) holds. This obviously ensures the Dini subdifferentiability of d_C on $U \setminus C$ and since one always has $0 \in \partial^- d_C(x)$ for all $\in C$, we obtain that d_C is Dini subdifferentiable on U. The nonvacuity of P_C on U follows from the above implication (g) $\Rightarrow$ (f) and from Proposition 1.3.18.
(k) $\Rightarrow$ (g): For any $x \in U \setminus C$, there exists some x^* in the Dini subdifferential $\partial^- d_C(x)$ of d_C at x. Choose $p(x) \in P_C(x)$. By the definition of Dini subdifferential, for any $\varepsilon > 0$, there is some $\delta > 0$ such that, for any $t \in]0, \delta[$, one has

$$\langle x^*, p(x) - x \rangle \le t^{-1}[d_C(x + t(p(x) - x)) - d_C(x)] + \varepsilon \tag{1.36}$$

and hence

$$\begin{aligned}
\langle x^*, x - p(x) \rangle &\ge t^{-1}[d_C(x) - d_C(x + t(p(x) - x))] - \varepsilon \\
&\ge t^{-1}[\|x - p(x)\| - \|x + t(p(x) - x) - p(x)\|] - \varepsilon.
\end{aligned}$$

Then, for any $\varepsilon > 0$, using the equality $\nabla^F(\|\cdot\|^2)(x - p(x)) = 2J(x - p(x))$ and taking some $t > 0$ small enough, we obtain

$$\begin{aligned}
\langle x^*, x - p(x) \rangle &\ge \left\langle \frac{J(x - p(x))}{\|x - p(x)\|}, x - p(x) \right\rangle - 2\varepsilon \\
&= \|x - p(x)\| - 2\varepsilon,
\end{aligned}$$

the last inequality being due the fact that $\langle J(y), y\rangle = \|y\|^2$. Therefore,

$$\|x - p(x)\| \le \langle x^*, x - p(x)\rangle \le \liminf_{t\downarrow 0} t^{-1}[d_C(x + t(x - p(x)) - d_C(x)] \le$$

$$\limsup_{t\downarrow 0} t^{-1}[d_C(x + t(x - p(x)) - d_C(x)] \le \|x - p(x)\|$$

(the second inequality coming from $x^* \in \partial^- d_C(x)$), so

$$\lim_{t\downarrow 0} t^{-1}[d_C(x + t(x - p(x)) - d_C(x)] = \|x - p(x)\|.$$

Further observe (as in the proof of Corollary 2 in Borwein, Fitzpatrick and Giles [24]) that for each $t \in [-1, 0[$ one has $p(x) \in P_C(x + t(x - p(x))$ and hence

$$t^{-1}[d_C(x + t(x - p(x))) - d_C(x)] = \|x - p(x)\|.$$

So, $\lim_{t\to 0} t^{-1}[d_C(x + t(x - p(x)) - d_C(x)] = \|x - p(x)\|$, that is, d_C has a Gâteaux directional derivative in the full direction $x - p(x)$. Consequently, (g) follows from Theorem 1.3.20.
(f) $\Leftrightarrow$ (b) is Proposition 1.3.18.
(b) $\Leftrightarrow$ (c) is Lemma 1.3.17.
(b) $\Rightarrow$ (d) is Lemma 1.3.22.
(d) $\Leftrightarrow$ (e) $\Leftrightarrow$ (a) is Proposition 1.3.23.

Under the additional assumption, to see that we have (n) $\Leftrightarrow$ (c) we need to prove the implication (n) $\Rightarrow$ (c).
Let us take any $u \in U$ and $u_n \xrightarrow[n\to\infty]{\|\cdot\|} u$. By weak compactness, taking if necessary a subsequence, we may suppose that $P_C(u_n) \xrightarrow[n\to\infty]{w} v$. From the weak lower semicontinuity of the norm, we have $\|v - u\| \le \liminf_{n\to\infty} \|u_n - P_C(u_n)\|$ and hence $\|v - u\| \le d_C(u)$. Therefore, as C is weakly closed, $v \in C$ and $P_C(u) = v$. Thus, $P_C(u_n) \xrightarrow[n\to\infty]{w} P_C(u)$, which gives the norm-to-weak continuity of P_C on U.

To conclude, we apply Corollary 1.3.15 to obtain (a) $\Leftrightarrow$ (b) $\Leftrightarrow$ (f) but now with the Hölder character of the continuity that holds on (possibly smaller) neighbourhood of $\overline{x}$. The proof is then complete. □

1.3.4 Characterizations of uniformly prox-regular sets

In this final subsection we proceed to the study of the global setting of positively reached or proximally smooth set C, corresponding (see Clarke, Stern and Wolenski [46]) to the continuous Fréchet differentiability of the distance function d_C over all an open tube of uniform thickness around the set C. We still have most of the characterizations for these sets given in Federer [73] for finite dimensional space and in Clarke, Stern and Wolenski [46] and Poliquin, Rockafellar and Thibault [138] in Hilbert space. We also add the characterizations (c), (e), and (g).

Definition 1.3.26. Following our definition of prox-regular set (see Definition 1.3.3) and modifying slightly Definition 2.4 of Poliquin, Rockafellar and Thibault [138], we will say that the closed set C is *uniformly r-prox-regular* if whenever $x \in C$ and $p^* \in N_C^P(x)$ with $\|p^*\| < 1$, then x is the unique nearest point of C to $x + rJ^*(p^*)$.

For the subset C of X, let us first recall the definitions of
the *r-enlargement* of C

$$C(r) := \{x \in X : \ d_C(x) \le r\},$$

the open *r-tube* around C

$$U_C(r) := \{x \in X : \ 0 < d_C(x) < r\},$$

and let us define the set of *r-distance points* to C

$$D_C(r) := \{x \in X : \ d_C(x) = r\}.$$

Theorem 1.3.27. Let $C \subset X$ be a closed set and $r > 0$. The following are equivalent:
(a) C is uniformly r-prox-regular;
(b) d_C is continuously differentiable on $U_C(r) \setminus C$;
(c) d_C is Fréchet regular on $U_C(r) \setminus C$;
(d) d_C is Fréchet differentiable on $U_C(r) \setminus C$;
(e) d_C is Gâteaux differentiable on $U_C(r) \setminus C$ with $\|\nabla^G d_C(x)\| = 1$ for all $x \in U_C(r) \setminus C$;
(f) $\partial_F d_C$ is nonempty-valued at all points in $U_C(r)$;
(g) P_C and $\partial^- d_C$ are nonempty-valued at all points in $U_C(r)$;
(h) d_C^2 is C^1 on $U_C(r)$ with locally Hölder continuous derivative mapping;
(i) P_C is single-valued and locally Hölder continuous on $U_C(r)$;
(j) P_C is single-valued and norm-to-weak continuous on $U_C(r)$;
(k) For any non-zero $p \in PN_C(x)$ with $x \in C$ one has $x \in P_C\left(x + r\frac{p}{\|p\|}\right)$;
(l) If $u \in U_C(r)$ and $x = P_C(u)$, then $x \in P_C(u')$ for $u' = x + r\frac{u-x}{\|u-x\|}$.

If C is weakly closed, one has the additional equivalent condition
(m) P_C is single-valued on $U_C(r)$.

Proof. We will follow the scheme:

$$\begin{array}{ccccccccccc} & & (a) & \Leftrightarrow & (k) & \Leftrightarrow & (l) & \Rightarrow & (i) & \Rightarrow & (h) \\ & & & & & & \Uparrow & & & & \Downarrow \\ & & & & (j) & \Leftrightarrow & (b) & \Leftrightarrow & (c) & \Leftrightarrow & (d) & \Leftrightarrow & (e) & \Leftrightarrow & (f). \\ & & & & & & & & & & \Updownarrow \\ & & & & & & & & & & (g) \end{array}$$

(k) $\Leftrightarrow$ (l) is obvious and here below both will be considered.
(k) $\Rightarrow$ (a) follows from Lemma 1.3.1.
(a) $\Rightarrow$ (k): Fix any nonzero $p \in PN_C(x)$ and take $p_n := p/\left(\|p\| + \frac{1}{n}\right)$. For any $x' \in C$, (a) entails $\|x' - (x + rp_n)\| \geq r\|p_n\|$ which yields after taking the limit $\left\|x' - (x + r\frac{p}{\|p\|})\right\| \geq r$. The latter means that $x \in P_C\left(x + r\frac{p}{\|p\|}\right)$, which is (k).
(b) $\Leftrightarrow$ (c) $\Leftrightarrow$ (d) $\Leftrightarrow$ (e) $\Leftrightarrow$ (f) follows from Proposition 1.3.19.
(d) $\Leftrightarrow$ (g) results from Theorem 1.3.20 like in the proof of the similar equivalence in Theorem 1.3.25.
(b) $\Leftrightarrow$ (j) is a consequence of Proposition 1.3.18 and Lemma 1.3.17.
(b) $\Rightarrow$ (l): Let $u \in U_C(r)$ and $x = P_C(u)$. Since (b) holds, according to Lemma 1.3.21 there exists some $t_0 > 0$ such that $P_C(u_t) = x$ for all $u_t := u + t(u - x)/\|u - x\|$ with $0 < t < t_0$. As in Poliquin, Rockafellar and Thibault [138] one may consider the number λ_0 given by the supremum over all $t \in [0, r - d_C(u)]$ such that $x \in P_C(u_t)$. Using the equivalence (note that $x \in C$ and $\|x - u_t\| = d_C(u) + t$)

$$x \in P_C(u_t) \Leftrightarrow \forall x' \in C,\ \|x' - u_t\| \geq d_C(u) + t,$$

it is easily seen that the supremum λ_0 is attained. We now claim that $\lambda_0 = r - d_C(u)$. Assume the contrary, i.e., $\lambda_0 < r - d_C(u)$. Then one would have on the one hand $x \in U_C(r)$ and on the other hand $x = P_C(u_{\lambda_0})$ because of the assumptions (b) and Proposition 1.3.18. Applying Lemma 1.3.21 again one would obtain a contradiction with the supremum property of λ_0. So the equality $\lambda_0 = r - d_C(u)$ holds. As u_t can be written in the form $u_t = x + (d_C(u) + t)(u - x)/\|u - x\|$, taking $t = \lambda_0$ gives (l).
(l) $\Rightarrow$ (i): We will proceed in five steps.

Step 1. For any $x^* \in N_C^{P\,r}(x) = N_C^P(x) \cap rB^*$ with $x \in C$ and any $\alpha \in]0,1]$, from (k) $\Leftrightarrow$ (l) one has $P_C\left(x + r\alpha\frac{J^*x^*}{\|x^*\|}\right) \ni x$. This means that, for any $x' \in C$,

$$\left\|x + r\alpha\frac{J^*x^*}{\|x^*\|} - x\right\| \leq \left\|x + r\alpha\frac{J^*x^*}{\|x^*\|} - x'\right\|.$$

Besides, as $J = \nabla^F\left(\frac{1}{2}\|\cdot\|^2\right)$, one also has

$$\frac{1}{2}\left\|x + r\alpha\frac{J^*x^*}{\|x^*\|} - x'\right\|^2 + \left\langle J\left(x - x' + r\alpha\frac{J^*x^*}{\|x^*\|}\right), x' - x\right\rangle \leq \frac{1}{2}\left\|r\alpha\frac{J^*x^*}{\|x^*\|}\right\|^2.$$

So, $\left\langle J\left(J^*(x^*) - \frac{\|x^*\|}{r\alpha}(x'-x)\right), x'-x\right\rangle \leq 0$. It is possible to take any α in $\left]0, \frac{\|x^*\|}{r}\right]$. Hence, whenever $x_i \in C, x_i^* \in N_C^{P\,r}(x_i)$, $i = 1, 2$, and $t \geq 1$, one has

$$\begin{aligned} \langle J(J^*(x_1^*) - t(x_2 - x_1)), x_2 - x_1\rangle &\leq 0 \\ \text{and} \quad \langle J(J^*(x_2^*) - t(x_1 - x_2)), x_1 - x_2\rangle &\leq 0. \end{aligned}$$

By adding, one obtains

$$\langle J(J^*(x_1^*) - t(x_2 - x_1)) - J(J^*(x_2^*) - t(x_1 - x_2)), x_1 - x_2\rangle \leq 0$$

which is the J-hypomonotonicity of $N_C^{P\,r}$ of degree t for any $t \geq 1$.

Step 2. For any $\alpha \in]0, 1/2[$, we have by Step 1 and by Lemma 1.3.8 that the set-valued mapping $(I + J^* \circ N_C^{P\,\alpha r})^{-1}$ is $\frac{1}{q}$-Hölder continuous on the intersection of its domain with any bounded subset. We also have, for any $r' > 0$,

$$P_C(x) \subset (I + J^* \circ N_C^{P\,r'})^{-1}(x) \quad \text{for any } x \in U_C(r'). \tag{1.37}$$

Indeed, for any $x \in U_C(r')$, the inclusion $y \in P_C(x)$ entails that $J(x-y) \in N_C^P(y)$, and $\|y - x\| < r'$ so $J(x-y) \in N_C^{P\,r'}(y)$. So we have that for any $\alpha \in]0, 1/2[$, P_C is $\frac{1}{q}$-Hölder continuous on the intersection of any bounded set with $\operatorname{Dom} P_C \cap U_C(\alpha r)$. Then by the arguments of Step 2 in the proof of Theorem 1.3.13, P_C is also nonempty, single-valued on $U_C(\alpha r)$. As α can be made as close as one wants to $\frac{1}{2}$, P_C is nonempty, single-valued, locally $\frac{1}{q}$-Hölder continuous on $U_C(r/2)$.

Step 3 corresponds to the two following lemmas. The first one completes the result of Lemma 3.1 of Bounkhel and Thibault [33]. As usual the line segment between two points $u, v \in X$ will be denoted by $[u, v]$, that is, $[u, v] := \{tu + (1-t)v \,:\, t \in [0, 1]\}$.

Lemma 1.3.28. Let C be a nonempty closed subset of a normed vector space $(Y, \|\cdot\|)$. Let $\rho > 0$ and $u \notin C(\rho)$. Then the following hold:
(a) $d_C(u) = \rho + d_{C(\rho)}(u) = \rho + d_{D_C(\rho)}(u)$;
(b) If $u_0 \in P_C(u)$ and $y_0 \in [u_0, u] \cap D_C(\rho)$, then $y_0 \in P_{C(\rho)}(u)$;
(c) If $y \in P_{C(\rho)}(u)$ and $z \in P_C(y)$, then $z \in P_C(u)$. Further, if $P_{C(\rho)}(u) = \{y\}$ and $z \in P_C(y)$, then $y \in [z, u]$ and $P_C(u) = \{z\}$.

Proof. (a) For all $y \in C(\rho)$

$$d_C(u) \leq d_C(y) + \|u - y\| \leq \rho + \|u - y\|,$$

hence

$$d_C(u) \leq \rho + d_{C(\rho)}(u) \leq \rho + d_{D_C(\rho)}(u). \tag{1.38}$$

Fix any $\varepsilon > 0$ and choose $u_\varepsilon \in C$ with $\|u - u_\varepsilon\| \le d_C(u) + \varepsilon$. Since $d_C(u_\varepsilon) = 0$ and $d_C(u) > \rho$ we may choose $y_\varepsilon \in [u_\varepsilon, u] \cap D_C(\rho)$ and hence

$$d_C(u) + \varepsilon \ge \|u - u_\varepsilon\| = \|u - y_\varepsilon\| + \|y_\varepsilon - u_\varepsilon\| \ge d_{D_C(\rho)}(u) + d_C(y_\varepsilon) = d_{D_C(\rho)}(u) + \rho.$$

Combining this with (1.38) we obtain

$$d_C(u) = \rho + d_{C(\rho)}(u) = \rho + d_{D_C(\rho)}(u). \tag{1.39}$$

(b) Assume now that $u_0 \in P_C(u)$ and $y_0 \in [u_0, u] \cap D_C(\rho)$. Then $u_0 \in P_C(y_0)$ and by (1.39) we have

$$\|u - y_0\| + \rho = \|u - y_0\| + d_C(y_0) = \|u - y_0\| + \|y_0 - u_0\| = \|u - u_0\| = \rho + d_{C(\rho)}(u)$$

and hence $\|u - y_0\| = d_{C(\rho)}(u)$, i.e., $y_0 \in P_{C(\rho)}(u)$.
(c) Assume that $y \in P_{C(\rho)}(u)$ and $z \in P_C(y)$. Then

$$\|u - z\| \le \|u - y\| + \|y - z\| = d_{C(\rho)}(u) + d_C(y) \le d_{C(\rho)}(u) + \rho = d_C(u)$$

(the last equality being due to (1.39)) and hence $z \in P_C(u)$.

Assume now that $P_{C(\rho)}(u) = \{y\}$. Taking $y' \in [z, u] \cap D_C(\rho) \ne \emptyset$, we have by (b) that $y' \in P_{C(\rho)}(u)$ and hence $y = y' \in [z, u]$. If there exists $z' \ne z$ with $z' \in P_C(u)$, then one sees that $z' \notin u + [0, +\infty[(z - u)$ and by (b) for $y'' \in [z', u] \cap D_C(\rho) \ne \emptyset$ (hence $y'' \ne u$) one would have $y'' \in P_{C(\rho)}(u)$ and hence $y'' = y \in [z, u]$ which would contradict $z' \notin u + [0, +\infty[(z - u)$. This completes the proof of the lemma. □

Lemma 1.3.29. If C satisfies the assertion (l) of Theorem 1.3.27 with parameter r and P_C is nonempty, single-valued on $U_C(\alpha r)$ for some $\alpha \in]0, 1]$, then for any $\alpha' \in]0, \alpha[$, the set $C(\alpha' r) := \{x \in X : d_C(x) \le \alpha' r\}$ satisfies (l) with parameter $r(1 - \alpha')$.

Proof. Take $u \in U_{C(\alpha' r)}(r(1 - \alpha'))$ and put $r' := d_C(u)$. Note that $0 < d_{C(\alpha' r)}(u) < r(1 - \alpha')$ and hence by (a) of Lemma 1.3.28 one has $\alpha' r < d_C(u) < r$, which implies in particular $u \in U_C(r)$. Suppose that $P_{C(\alpha' r)}(u) = \{y\}$. We have to prove that $y \in P_{C(\alpha' r)}\left(y + r(1 - \alpha')\frac{u-y}{\|u-y\|}\right)$. Observing that $y \in U_C(\alpha r)$, we may put $z := P_C(y)$ according to the assumption on P_C over $U_C(\alpha r)$. By (c) of Lemma 1.3.28, we have $z \in P_C(u)$. Since $d_C(z) = 0$ and $d_C(u) > \alpha' r$, we may take $y_1 \in [z, u] \cap D_C(\alpha' r) \ne \emptyset$. The assertion (b) of Lemma 1.3.28 says that $y_1 \in P_C(\alpha' r)(u)$ and hence $y_1 = y$. Now, since $y_1 \in U_C(r)$ and $z = P_C(y_1)$ we have by (l) that $P_C(u') \ni z$ for

$$u' := z + r\frac{y_1 - z}{\|y_1 - z\|} = z + r\frac{u - z}{\|u - z\|}.$$

This entails by (b) of Lemma 1.3.28 again that $P_{C(\alpha' r)}(u') \ni y$ since (recall that $y = y_1$)

$$y \in [z, u] \cap D_C(\alpha' r) \subset [z, u'] \cap D_C(\alpha' r).$$

Further, since $\|u' - y\| = \|u' - z\| - \|y - z\| = r - \alpha' r$ (the second equality being due to the definition of u' and to the inclusion $y \in D_C(\alpha' r)$) we see that

$$u' = y + (r - \alpha' r)\frac{u - y}{\|u - y\|}.$$

So, we obtain that $C(\alpha' r)$ satisfies (l) with parameter $r(1 - \alpha')$, and the proof of the lemma is complete. □

Step 4. For $\alpha \in]0, 1]$, let us consider the property

$$\mathcal{P}(\alpha) \begin{cases} C \text{ satisfies (l) with parameter } r \text{ and} \\ P_C \text{ is single-valued, locally Hölder continuous on } U_C(\alpha r). \end{cases}$$

We claim that $\mathcal{P}(\alpha) \Rightarrow \mathcal{P}\left(\frac{\alpha+1}{2}\right)$.
Suppose that $\mathcal{P}(\alpha)$ holds. From Steps 1, 2 and 3, we have that for any $\alpha' \in]0, \alpha[$,

$$(1.40) \qquad P_{C(\alpha' r)} \text{ is single-valued, locally Hölder continuous on } U_{C(\alpha' r)}\left(\frac{r(1 - \alpha')}{2}\right).$$

Take any $u \in U_C\left(\alpha' r + \frac{r(1-\alpha')}{2}\right)$ such that $r' := d_C(u) > \alpha' r$. By (a) of Lemma 1.3.28 we have $d_{C(\alpha' r)}(u) = r' - \alpha' r$, and so

$$(1.41) \qquad u \in U_{C(\alpha' r)}\left(\frac{r(1 - \alpha')}{2}\right)$$

since $r' - \alpha' r < \frac{r(1-\alpha')}{2}$ because of the inclusion $u \in U_C\left(\alpha' r + \frac{r(1-\alpha')}{2}\right)$. We may then put $y := P_{C(\alpha' r)}(u)$ according to (1.40) and put $z := P_C(y)$ according to the second assumption in $\mathcal{P}(\alpha)$. Then by (c) of Lemma 1.3.28 we have $z = P_C(u)$ and so $P_C(u) = z = P_C \circ P_{C(\alpha' r)}(u)$. So for any $\alpha' \in]0, \alpha[$, (1.40), (1.41), and $\mathcal{P}(\alpha)$ ensure that P_C is single-valued and locally Hölder continuous on $U_C\left(\alpha' r + \frac{r(1-\alpha')}{2}\right) \setminus C(\alpha' r)$. By assumption it is also locally Hölder continuous on $U_C(\alpha r)$. So P_C is single valued, locally Hölder continuous on $U_C\left(\alpha' r + \frac{r(1-\alpha')}{2}\right)$ for any $\alpha' \in]0, \alpha[$ and hence also on $U_C\left(\alpha r + \frac{r(1-\alpha)}{2}\right) = U_C\left(\frac{\alpha+1}{2} r\right)$. This establishes the claim and finishes the proof of Step 4.

Step 5. Define $(\alpha_n)_n$ by $\alpha_0 = 1/2$, $\alpha_{n+1} = (\alpha_n + 1)/2$. We have $\alpha_n \to 1$ and by Step 4, $\mathcal{P}(\alpha_n) \Rightarrow \mathcal{P}(\alpha_{n+1})$. As $\mathcal{P}(\alpha_0)$ is true by Step 1, we have $\mathcal{P}(1)$, that entails (i).
(i) ⇒ (h): This implication follows from Lemma 1.3.10 and Remark 1.3.12.
(h) ⇒ (d) is obvious.

Since the additional equivalence (m) can be established like in Theorem 1.3.25, the proof is now complete. □

The following proposition gives the expression of the metric projection mapping in terms of the normal cone.

Proposition 1.3.30. Under the assumptions of Theorem 1.3.27, one has for all $x \in U_C(r)$

$$\nabla^F d_C(x) = J(x - P_C(x))/d_C(x) \quad \text{and} \quad P_C(x) = (I + J^* \circ N_C^{P\,r})^{-1}(x).$$

Proof. The equality concerning the derivative follows easily from the theorem above and from the expression of $\nabla^F e_\lambda f(x)$ in Lemma 1.3.9. Let us prove the equality concerning the metric projection mapping. We know by (1.37) that $P_C(x) \subset (I + J^* \circ N_C^{P\,r})^{-1}(x)$ for all $x \in U_C(r)$, so it is enough to check that $(I + J^* \circ N_C^{P\,r})^{-1}$ is at most single-valued on $U_C(r)$. Suppose that y_1, y_2 are in $(I + J^* \circ N_C^{P\,r})^{-1}(x)$ with $x \in U_C(r)$, that is, $J(x - y_i) \in N_C^{P\,r}(y_i)$, $i = 1, 2$. Then $P_C(x) = y_1 = y_2$ by (k). □

The uniform prox-regularity also entails the J-hypomonotonicity of the truncated normal functional cone.

Proposition 1.3.31. Under the assumptions of Theorem 1.3.27, we also have
(n) The truncated normal functional cone mapping $N_C^{P\,r}$ is J-hypomonotone of degree t for any $t \geq 1$.
Conversely, (n) entails the assertions of Theorem 1.3.27 with parameter $r/2$ instead of r.

Proof. See Steps 1 and 2 in the proof of Theorem 1.3.27. □

The following corollary gives several characterizations of convex sets. The condition (k) in the corollary is Vlasov's extension (see Vlasov [167]) to Banach space with rotund dual of the known result proved by Asplund in Hilbert space (see Asplund [1]). The characterization (f) is exactly Theorem 18 of Borwein, Fitzpatrick and Giles [24]. Theorem 18 in Borwein, Fitzpatrick and Giles [24] was established in the more general case where the dual space is merely rotund and it was, in some sense, a generalization of Theorem 3.6 of Fitzpatrick [75] where both the norm and the dual norm were assumed to be Fréchet differentiable away from the origins. Note also that the characterization (h) was explicitly given by Borwein, Fitzpatrick and Giles [24, Theorem 17] (with directional Gâteaux derivability as in Theorem 1.3.20 instead of the Dini subdifferentiability), in the larger context of Banach space Y with rotund dual Y^*. Both Theorem 17 and Theorem 18 in Borwein, Fitzpatrick and Giles [24] are derived by the authors by proving that any closed subset of a Banach space Y such that

$$\limsup_{\|y\| \to 0} \frac{d_C(x + y) - d_C(x)}{\|y\|} = 1 \quad \text{for all} \quad x \in Y \setminus C$$

is almost convex and hence, according to a result due to Vlasov [167], convex whenever Y^* is rotund for its dual norm.

In our corollary, the requirement (X5) is imposed for a large part because of characterizations (b), (i), (j), (m), (n).

Corollary 1.3.32. Let $C \subset X$ be a closed set. The following are equivalent:
(a) C is convex;
(b) C is uniformly ∞-prox-regular, i.e., uniformly r-prox-regular for any real number $r > 0$;
(c) d_C is continuously differentiable on $X \setminus C$;
(d) d_C is Fréchet regular on $X \setminus C$;
(e) d_C is Fréchet differentiable on $X \setminus C$;
(f) d_C is Gâteaux differentiable on $X \setminus C$ with $\|\nabla^G d_C(x)\| = 1$ for all $x \in X \setminus C$;
(g) $\partial_F d_C(x) \neq \emptyset$ for all $x \in X$;
(h) $P_C(x) \neq \emptyset$ and $\partial^- d_C(x) \neq \emptyset$ for all $x \in X$;
(i) d_C^2 is C^1 on X with locally Hölder continuous derivative mapping on X;
(j) P_C is single-valued and locally Hölder continuous on X;
(k) P_C is single-valued and norm-to-norm continuous on X;
(l) P_C is single-valued and norm-to-weak continuous on X;
(m) For any $p \in PN_C(x)$ with $x \in C$ one has $x \in P_C(x + p)$;
(n) If $u \in X \setminus C$ and $x = P_C(u)$, then $x \in P_C(u')$ for $u' = x + r(u - x)$ and any $r > 0$.

Proof. The equivalence between all the assertions from (b) to (n) is easily seen to follow from Theorem 1.3.27. The implication (a) $\Rightarrow$ (g) is obvious according to the convexity of the continuous function d_C under (a). By Proposition 1.3.31, the condition (g) entails that $N_C^{P\,r}$ is J-hypomonotone of degree 1 on $U_C(r)$ for any $r > 0$. Fix any $x, y \in C$ and any $x^* \in N_C^P(x)$, $y^* \in N_C^P(y)$. By definition of J-hypomonotonicity of degree 1 for all $s > 0$ large enough

$$\langle J[J^*(sx^*) - (y - x)] - J[J^*(sy^*) - (x - y)], x - y\rangle \geq 0,$$

or

$$\left\langle J\left[J^*(x^*) - \frac{1}{s}(y - x)\right] - J\left[J^*(y^*) - \frac{1}{s}(x - y)\right], x - y\right\rangle \geq 0.$$

Using the continuity of J and J^* and passing to the limit when s goes to $+\infty$ we obtain

$$\langle x^* - y^*, x - y\rangle \geq 0.$$

This means that N_C^P is monotone and hence by Correa, Gajardo and Thibault [49] (see also Correa, Jofre and Thibault [51, 52]) the set C is convex, that is, (a). The proof of the corollary is then complete. □

In the remainder of the subsection, we will give, for an r-prox-regular set C, some properties of the ρ-enlargement $C(\rho)$ with $\rho \in]0, r[$ and of the set of ρ-external points to C

$$E_C(\rho) := \{u \in X \,:\, d_C(u) \geq \rho\}.$$

The first result concerns the proximal normal cone to the ρ-enlargement of C. In its statement we will denote by N_S^{Cl} the Clarke normal cone to a set S and by $\mathbb{R}_+$ the set of all non negative real numbers.

Proposition 1.3.33. Assume that C is uniformly r-prox-regular. Then, for any $\rho \in]0, r[$ and any $y \in D_C(\rho)$,

$$N^P_{C(\rho)}(y) = N^{Cl}_{C(\rho)}(y) = \mathbb{R}_+ \nabla^F d_C(y) \subset N^P_C(P_C(y)).$$

Proof. Fix $y \in D_C(\rho)$. From Lemma 1.3.29 and Theorem 1.3.27, the set $C(\rho)$ is uniformly $(r-\rho)$-prox-regular and $P_{C(\rho)}$ is single-valued on $U_C(r)$. To see that $\nabla^F d_C(y)$ actually belongs to $N^P_{C(\rho)}(y)$, put

$$u := y + \varepsilon(y - P_C(y)), \tag{1.42}$$

where ε is small enough that $d_C(u) < r$. As $P_C(u) = P_C(y)$ by Theorem 1.3.27 (l), we have

$$d_C(u) = \|u - P_C(y)\| = \|y + \varepsilon(y - P_C(y)) - P_C(y)\| = (1 + \varepsilon)\|y - P_C(y)\| = (1 + \varepsilon)\rho > \rho.$$

Then by Lemma 1.3.28 (b), since $P_{C(\rho)}$ is single-valued on $U_C(r)$, one has $y = P_{C(\rho)}(u)$. So $u - y \in PN_{C(\rho)}(y)$ and hence $y - P_C(y) \in PN_{C(\rho)}(y)$ which entails $\nabla^F d_C(y) \in N^P_{C(\rho)}(y)$. Further, as $C(\rho) = \{x \in X \,:\, d_C(x) \leq d_C(y)\}$ because $y \in D_C(\rho)$, we know by Clarke [44, Theorem 2.4.7 and Corollary 1] that $N^{Cl}_{C(\rho)} = \mathbb{R}_+ \nabla^F d_C(y)$ and hence

$$\mathbb{R}_+ \nabla^F d_C(y) \subset N^P_{C(\rho)}(y) \subset N^{Cl}_{C(\rho)}(y) = \mathbb{R}_+ \nabla^F d_C(y).$$

So it remains to see that the inclusion $\mathbb{R}_+ \nabla^F d_C(y) \subset N^P_C(P_C(y))$ follows from $\nabla^F d_C(y) = \dfrac{y - P_C(y)}{\|y - P_C(y)\|} \in N^P_C(P_C(y))$. □

Before establishing the second result, let us prove the following lemma which has its own interest.

Lemma 1.3.34. Assume that C is uniformly r-prox-regular. Then for any $\rho \in]0, r]$ and $y \in U_C(\rho)$ one has

$$d_C(y) + d_{E_C(\rho)}(y) = \rho. \tag{1.43}$$

Proof. Fix $y \in U_C(\rho)$ and put $x := P_C(y)$ according to (i) of Theorem 1.3.27. Then for $u := x+\rho\frac{y-x}{\|y-x\|}$ one has $x \in P_C(u)$ by Theorem 1.3.27 (k) and hence $u \in D_C(\rho) \subset E_C(\rho)$ and

$$d_C(y) + d_{E_C(\rho)}(y) \le \|y - x\| + \|y - u\| = \|u - x\| = \rho. \tag{1.44}$$

For any $z \in E_C(\rho)$ we have

$$\|z - y\| \ge \|z - P_C(y)\| - \|y - P_C(y)\| \ge d_C(z) - d_C(y) \ge \rho - d_C(y)$$

hence

$$d_{E_C(\rho)}(y) \ge \rho - d_C(y). \tag{1.45}$$

It follows from (1.44) and (1.45) that $d_C(y) + d_{E_C(\rho)}(y) = \rho$. □

From this lemma we see, on the one hand, through Theorem 1.3.27 (d) that if C is uniformly r-prox-regular, then for any $\rho \in]0, r[$, the set $E_{C(\rho)}$ is *uniformly ρ-prox-regular*, because $d_{E_C(\rho)}(.) = \rho - d_C(.)$ and $U_C(\rho) = U_{E_C(\rho)}(\rho)$. On the other hand, the lemma allows us to retrieve the following characterization of uniform r-prox-regularity given by Clarke, Stern and Wolenski [46] in the context of Hilbert space. Their proof is different and it relies on the analysis of an appropriate infimum value function. In our characterization below, the additional nonvacuity of $P_C(y)$ is not required (compare with Clarke, Stern and Wolenski [46, Theorem 4.1 (c)]).

Theorem 1.3.35. A closed set C is uniformly r-prox-regular for some $r > 0$, if and only if,

$$d_C(y) + d_{E_C(r)}(y) = r \quad \text{for all } y \in U_C(r). \tag{1.46}$$

Proof. The fact that (1.46) is implied by the uniform r-prox-regularity of C follows from Lemma 1.3.34 with $\rho = r$. Assume now that (1.46) holds and consider any $y \in U_C(r)$ for which $P_C(y)$ is a singleton, say $P_C(y) = x$. Then, for any $y' \in]x, y[$, one has $P_C(y') = x$. This yields for any non zero $t \in \left]-1, \frac{\|y-y'\|}{\|y'-x\|}\right[$

$$t^{-1}[d_C(y' + t(y' - x))) - d_C(y')] = t^{-1}[\|y' + t(y' - x) - x\| - \|y' - x\|] = \|y' - x\|,$$

which entails that d_C has a Gâteaux directional derivative in the full direction $y' - x$ and hence by Theorem 1.3.20 the function d_C is Fréchet differentiable at y'. Thus by (1.46) the function $d_{E_C(r)}$ is also differentiable at y'. We then deduce (see the formula for $\nabla^F e_\lambda f$ in Lemma 1.3.9) that $P_{E_C(r)}(y')$ is a singleton that will be denoted by u'. Using successively the inclusion $u' \in D_C(r)$ and (1.46) we obtain

$$r \le \|x - u'\| \le \|y' - x\| + \|y' - u'\| = r,$$

so $y' \in]x, u'[$ (see (1.11)) and $P_C(u') \ni x$. By Theorem 1.3.27 (l), this means that C is uniformly r-prox-regular. □

1.4 Prox-regular sets and epigraphs in uniformly convex Banach spaces: various regularities and other properties

In this section we continue the study of prox-regular sets that we began in the previous section in the setting of uniformly convex Banach space endowed with a norm both uniformly smooth and uniformly convex (like L^p, $W^{m,p}$ spaces).

Here in uniformly convex Banach space we prove normal and tangential regularity properties for these sets, and in particular the equality between Mordukhovich and proximal normal cones for such sets. We also compare in this setting the proximal normal cone with different Hölderian normal cones depending on the power types s, q of moduli of smoothness and convexity of the norm. In the case of sets that are epigraphs of functions, we show that J-primal lower regular functions have prox-regular epigraphs and we compare these functions with Poliquin's primal lower nice functions depending on the power types s, q of the moduli. The preservation of prox-regularity for the intersection of finitely many sets and of the inverse image are obtained under a calmness assumption. A conical derivative formula for the metric projection mapping of prox-regular set is also established. Among other results of the section it is proved that the Attouch-Wets convergence preserves the uniform r-prox-regularity property and that the metric projection mapping is continuous with respect to this convergence for such sets.

The results from this section are published by Bernard, Thibault and Zlateva in [19].

In this section we base our considerations on those in the previous one where we extended in some sense the study made in Poliquin, Rockafellar and Thibault [138] of prox-regular sets C in Hilbert spaces. Here we continue our study mainly considering normal cones and subdifferentials. In the context of Hilbert space the coincidence for a prox-regular set C at $\bar{x}$ between the Mordukhovich normal cone and the proximal normal one follows directly from the fact (due to the Hilbert structure) that in such a case there exists some non negative number γ and some neighbourhood U of $\bar{x}$ such that for any proximal normal x^* of C at a point $x \in U \cap C$ with $\|x^*\| \leq 1$ one has

$$\langle x^*, y - x \rangle \leq \gamma \|y - x\|^2 \quad \text{for all } y \in U \cap C.$$

The case of uniformly convex Banach space is not so obvious. Our aim is on the one hand to show that this important property of equality between the Mordukhovich normal cone and the proximal normal one still holds for prox-regular sets of uniformly convex Banach spaces and on the other hand to take advantage of this property to provide in the same context several new results including in particular the conical derivative of the metric projection mapping to C and the preservation of prox-regularity under the Attouch-Wets convergence.

In Subsection 1.4.1 we show the normal and tangential regularity properties of prox-regular sets of a uniformly convex Banach space X and deduce a proximal normal formula for these sets.

In Subsection 1.4.2 we prove that the epigraphs of the J-primal lower regular functions are prox-regular, and compare these functions with Poliquin's primal lower nice functions.

In Subsection 1.4.3, we draw a comparison between the prox-regularity concept considered in the previous section and taken from Poliquin, Rockafellar and Thibault [138], and another one from Poliquin and Rockafellar [137] and adapted by Bernard and Thibault in [14] to the general Banach context. We also compare the prox-regularity notions when the norm varies in certain families.

Subsection 1.4.4 concerns some relationships between different normal cones (of closed sets) related to the moduli of power type of the norm of X.

Subsection 1.4.5 deals with the preservation of prox-regularity under the intersection and the inverse image. Under a calmness qualification condition, it is shown that the intersection of finitely many prox-regular sets and the inverse image of a prox-regular set by a $C^{1,1}$ mapping inherit the prox-regularity property.

In Subsection 1.4.6 we take advantage of the result of tangential regularity of Subsection 1.4.1 to establish a conical derivative formula for the metric projection mapping of prox-regular set.

Finally Subsection 1.4.7 studies the behaviour of the metric projection mapping for a family $(C_t)_t$ of uniformly r-prox-regular closed sets of X which converges in the sense of Attouch-Wets to a closed set C. It is proved that C inherits the uniform r-prox-regularity property and that, for any x_0 with $d(x_0, C) < r$, one has $P_{C_t}(x_0) \underset{t}{\longrightarrow} P_C(x_0)$, for the metric projection mapping P_C onto the closed set C.

We will use most of the notations and preliminaries from the previous section. However, we will recall or refine some of them.

Throughout this section we work in an uniformly convex Banach space $(X, \|\cdot\|)$ whose norm $\|\cdot\|$ is both uniformly convex and uniformly smooth. In some statements, the moduli of uniform convexity and uniform smoothness of the norm $\|\cdot\|$ will be required to be of power type q, and of power type s, respectively and this will be explicitly stated when needed. One knows that such a renorm of the uniformly convex space always exists.

Some of the properties of uniformly convex Banach spaces were discussed in the beginning of the previous section, see also Diestel [58], Beauzami [12], Deville, Godefroy and Zizler [57].

For any real number $\sigma > 1$, consider the mapping $J_\sigma : X \to X^*$ defined by

$$J_\sigma(x) = \{x^* \in X^* : \langle x^*, x\rangle = \|x^*\|.\|x\|, \quad \|x^*\| = \|x\|^{\sigma-1}\}.$$

For $\sigma = 2$, J_2 will be simply denoted by J and it is generally called the *normalized duality mapping* associated with the norm $\|\cdot\|$. As X is reflexive we have that J is surjective. The mapping J_σ is the subdifferential of the convex function $\frac{1}{\sigma}\|\cdot\|^\sigma$, i.e., $J_\sigma = \partial\left(\frac{1}{\sigma}\|\cdot\|^\sigma\right)$. With the additional uniformly convex property of X and by the choice of the norm that we made, for any $\sigma > 1$, J_σ is single-valued, bijective and norm-to-norm continuous. The inverse mapping J^{-1} (of J) will be denoted by J^*, it is the normalized duality mapping for the dual norm on X^*. Observe also (by the definitions of J_σ and $J := J_2$) that for $\mathbb{R}_+ := [0, +\infty[$

$$J_\sigma(x) \in \mathbb{R}_+ J(x) \quad \text{and} \quad J_\sigma(tx) = t^{\sigma-1}J_\sigma(x) \quad \text{for all } x \in X \text{ and } t \geq 0. \tag{1.47}$$

We recall (see Deville, Godefroy and Zizler [57]) that the mapping J is uniformly continuous over each bounded subset of X (in fact, this property characterizes the uniform smoothness of the norm).

Moreover, from Xu and Roach [169, (2.16) p. 201 with p=2], there is a constant $K' > 0$ such that for all nonzero pairs $(x, y) \in X \times X$

$$\langle J(x) - J(y), x - y\rangle \geq K'(\max\{\|x\|, \|y\|\})^2 \delta_{\|\cdot\|}\left(\frac{\|x - y\|}{2\max\{\|x\|, \|y\|\}}\right) \tag{1.48}$$

and by $(3.1)'$ of Xu and Roach [169, p. 208] one also has some constant $L' > 0$ such that for all pairs $(x, y) \in X \times X$ with $x \neq y$

$$\|J(x) - J(y)\| \leq L' \frac{(\max\{\|x\|, \|y\|\})^2}{\|x - y\|} \rho_{\|\cdot\|}\left(\frac{\|x - y\|}{\max\{\|x\|, \|y\|\}}\right). \tag{1.49}$$

When the modulus of uniform convexity of the norm $\|\cdot\|$ is of power type q, from Xu and Roach [169, $(2.17)'$ p. 202] again, there exists some constant $K > 0$ such that for every $(x, y) \in X \times X$,

$$\|x + y\|^q \geq \|x\|^q + q\langle J_q(x), y\rangle + K\|y\|^q. \tag{1.50}$$

Similarly, whenever the modulus of smoothness of the norm $\|\cdot\|$ is of power type s, there exists, according to Xu and Roach [169, Remark 5, p. 208], some constant $L > 0$ such that, for all $x, y \in X$,

$$\|x + y\|^s \leq \|x\|^s + s\langle J_s(x), y\rangle + L\|y\|^s. \tag{1.51}$$

Observe that (1.50) and (1.51) respectively entail that for all $x, y \in X$,

$$\langle J_q(x) - J_q(y), x - y\rangle \geq \frac{2K}{q}\|x - y\|^q \tag{1.52}$$

and

$$\langle J_s(x) - J_s(y), x - y\rangle \leq \frac{2L}{s}\|x - y\|^s. \tag{1.53}$$

Inequalities in the lines of (1.52) and (1.53) also hold with the normalized duality mapping J in place of J_q or J_s when they are restricted to bounded subsets of X. Indeed, if the norm $\|\cdot\|$ has modulus of convexity (resp. smoothness) of power type q (resp. s), then for any $r > 0$ according to (1.48) (resp. (1.49)) there exist some positive constant K_r (resp. L_r) such that

$$\langle J(x) - J(y), x - y\rangle \geq K_r\|x - y\|^q, \quad \forall x, y \in rB, \text{ (see (1.8))} \tag{1.54}$$

(resp.

$$\|J(x) - J(y)\| \leq L_r\|x - y\|^{s-1}, \quad \forall x, y \in rB, \text{ see (1.9)).} \tag{1.55}$$

As in the previous section, the space $X \times \mathbb{R}$ will be considered with the norm $|||\cdot|||$ given by $|||(x, r)||| = \sqrt{\|x\|^2 + r^2}$. So, for the normalized duality mapping $J_{X\times\mathbb{R}}$: $X \times \mathbb{R} \to X^* \times \mathbb{R}$ associated with the norm $|||\cdot|||$, one has the equality

$$J_{X\times\mathbb{R}}(x, r) = (J(x), r). \tag{1.56}$$

When there is no risk of confusion, $J_{X\times\mathbb{R}}$ will be simply denoted by J.

Recall that for a closed set $C \subset X$, a nonzero vector $p \in X$ is said to be a *primal proximal normal vector* to C at $x \in C$ if there are $u \notin C$ and $r > 0$ such that $p = r^{-1}(u - x)$ and $\|u - x\| = d_C(u)$. Equivalently, a nonzero $p \in X$ is a primal proximal normal vector to C at $x \in C$ if there exists $r > 0$ such that $x \in P_C(x + rp)$. We also take by convention the origin of X as a primal proximal normal vector to C at x. The cone of all primal proximal normal vectors to C at x is denoted by $PN_C(x)$ and called the *primal proximal normal cone* of C at x. The *concept is local* in the sense that

$$\begin{cases} \text{for any } u \notin C \text{ and any closed ball } V := B[x, \beta] \text{ centred at } x \in C \\ \text{and such that } \|u - x\| = d_{C\cap V}(u), \text{ one has } u - x \in PN_C(x). \end{cases} \tag{1.57}$$

A continuous linear functional $p^* \in X^*$ is said to be a *proximal normal functional* to C at $x \in C$ if $J^*(p^*) \in PN_C(x)$. This means for $p^* \neq 0$ that there are $u \notin C$, $r > 0$ such that $p^* = r^{-1}J(u - x)$ and $\|u - x\| = d_C(u)$. Or, equivalently, a nonzero $p^* \in X^*$ is a proximal normal functional to C at $x \in C$ if there exists $r > 0$ such that

$x \in P_C(x + rJ^*(p^*))$. The cone of all proximal normal functionals to C at x will be denoted by $N^P_C(x)$.

One easily verifies that if $p \in PN_C(x)$, then $J(p) \in N^P_C(x)$, and that if $p^* \in N^P_C(x)$, then $J^*(p^*) \in PN_C(x)$ (keep in mind that $J^* = J^{-1}$ is the normalized duality mapping for X^* endowed with the dual norm of $\|\cdot\|$). Hence, $PN_C(x)$ and $N^P_C(x)$ completely determine each other.

We also recall the concept of Fréchet normal cone $N^F_C(x)$. A continuous linear functional $x^* \in X^*$ is said to be a *Fréchet normal functional* to C at $x \in C$ if for any $\varepsilon > 0$ there exists a neighbourhood U of x such that the inequality $\langle x^*, x' - x\rangle \leq \varepsilon\|x' - x\|$ holds for all $x' \in C \cap U$.

Since the norm of the space X we work in is Fréchet differentiable away from the origin, for any closed subset $C \subset X$ and any $x \in C$, any proximal normal functional to C at x is also a Fréchet normal functional to C at x.

We will also use the β-Hölder normal cone $N^\beta_C(\cdot)$ defined at $x \in C$ by $x^* \in N^\beta_C(x)$ when there exist $\varepsilon, \gamma > 0$ such that for all $x' \in B(x, \varepsilon) \cap C$, $\langle x^*, x' - x\rangle \leq \gamma\|x' - x\|^\beta$.

In the context where the norm $\|\cdot\|$ is associated with an inner product $(\cdot\,|\,\cdot)$, that is, $(X, \|\cdot\|)$ is a Hilbert space, then it is straightforward to verify that $x \in P_C(x + rp)$ if and only if

$$(p|y - x) \leq (2r)^{-1}\|y - x\|^2 \quad \text{for all } y \in C.$$

The same description says that for a nonzero vector p there exists some positive r satisfying the above property if and only if there are positive ε and γ such that

$$\langle Jp, y - x\rangle = (p|y - x) \leq \gamma\|y - x\|^2 \quad \text{for all } y \in B(x, \varepsilon) \cap C,$$

that is, for $\beta = 2$ one has $Jp \in N^\beta_C(x)$, and hence $N^P_C(x) = N^\beta_C(x)$. Since $N^\beta_C(x)$ is of course independent of any equivalent norm to $\|\cdot\|$, so is the cone $N^P_C(x)$ in the Hilbert setting.

The above notions can be translated in the context of functions.

Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous function. By definition, the effective *domain* of f is the set $\operatorname{dom} f := \{x \in X : f(x) < +\infty\}$ and the *epigraph* of f is the set $\operatorname{epi} f := \{(x, r) \in X \times \mathbb{R} : f(x) \leq r\}$. Let $x \in \operatorname{dom} f$.

We say that $p^* \in X^*$ is a *proximal subgradient* of f at x if $(p^*, -1)$ is a proximal normal functional to the epigraph of f at $(x, f(x))$. The *proximal subdifferential* of f at x, denoted by $\partial_P f(x)$, consists of all such functionals. Thus, we have $p^* \in \partial_P f(x)$, if and only if, $(p^*, -1) \in N^P_{\operatorname{epi} f}(x, f(x))$.

The functional $x^* \in X^*$ is said to be a *Fréchet subgradient* of f at x if $(x^*, -1)$ is a Fréchet normal functional to the epigraph of f at $(x, f(x))$. The *Fréchet subdifferential* of f at x, denoted by $\partial_F f(x)$, consists of all such functionals.

If $x \notin \operatorname{dom} f$ then all subdifferentials of f at x are empty, by convention. It is known that for a lower semicontinuous function f on a reflexive Banach space with a Fréchet differentiable Kadec norm (in particular, on the space X we work in), the set $\operatorname{dom}\partial_P f$ is dense in $\operatorname{dom} f$. Moreover, from what we saw above, $\partial_P f(x) \subset \partial_F f(x)$ for all $x \in X$. When $\partial_F f(x) \neq \emptyset$, one says that f is Fréchet subdifferentiable at the point x.

Similarly, a *β-Hölder subgradient* of f at x is any functional $x^* \in X^*$ such that $(x^*, -1) \in N_C^\beta(x, f(x))$. Denoting by $\partial_\beta f(x)$ the set of such subgradients, we have $x^* \in \partial_\beta f(x)$ if there exists $\gamma, \varepsilon > 0$ such that for all $y \in B(x, \varepsilon)$, $f(y) \geq f(x) + \langle x^*, y - x\rangle - \gamma\|y - x\|^\beta$.

It is easily checked that $\partial_P \psi_C(x) = N_C^P(x)$ for any $x \in C$. The Fréchet and β-Hölder normal cones of C are related to the indicator function ψ_C in a similar way.

We continue to study prox-regularity property of a set in the context of uniformly convex Banach spaces we began in Section 1.3.

First of all we will refine the notion of prox-regularity of sets. Taking into account Definition 1.3.3 and Proposition 1.3.4, it is easy to get

Proposition 1.4.1. For a closed subset $C \subset X$ the following are equivalent:
(a) C is prox-regular at $\overline{x} \in C$;
(b) there exist $\varepsilon > 0$ and $r > 0$, such that for all $x \in C$ with $\|x - \overline{x}\| < \varepsilon$ and for all $p^* \in N_C^P(x)$ such that $\|p^*\| \leq \varepsilon$,

$$0 \geq \langle J[J^*(p^*) - r^{-1}(x' - x)], x' - x\rangle, \quad \forall x' \in C \text{ such that } \|x' - \overline{x}\| < \varepsilon.$$

(c) for any $\Theta > 0$ there exist $\varepsilon > 0$, $r > 0$ such that for all $x \in C$ with $\|x - \overline{x}\| < \varepsilon$ and for all $p^* \in N_C^P(x)$ such that $\|p^*\| \leq \Theta$,

$$0 \geq \langle J[J^*(p^*) - r^{-1}(x' - x)], x' - x\rangle, \quad \forall x' \in C \text{ such that } \|x' - \overline{x}\| < \varepsilon.$$

Proof. The equivalence (a) $\Leftrightarrow$ (b) is proved in Proposition 1.3.4.

We will prove that (b) $\Leftrightarrow$ (c).

Let (b) holds with $\varepsilon' > 0$ and $r' > 0$. Fix $\Theta > 0$. Let $p^* \in N_C^P(x)$ be such that $\|x - \overline{x}\| < \varepsilon'$ and $\|p^*\| \leq \Theta$. Then $\frac{\varepsilon'}{\Theta}p^* \in N_C^P(x)$ is such that $\left\|\frac{\varepsilon'}{\Theta}p^*\right\| \leq \varepsilon'$ and form (b) we have

$$0 \geq \left\langle J\left[J^*\left(\frac{\varepsilon'}{\Theta}p^*\right) - (r')^{-1}(x' - x)\right], x' - x\right\rangle, \quad \forall x' \in C \text{ such that } \|x' - \overline{x}\| < \varepsilon',$$

or equivalently

$$0 \geq \langle J[J^*(p^*) - (r'\varepsilon'\Theta^{-1})^{-1}(x' - x)], x' - x\rangle, \quad \forall x' \in C \text{ such that } \|x' - \overline{x}\| < \varepsilon',$$

which means that (c) holds for Θ with $\varepsilon = \varepsilon'$ and $r = r'\varepsilon'\Theta^{-1}$.

Let now (c) holds for some $\Theta > 0$ with $\varepsilon' > 0$ and $r > 0$. Let $\varepsilon > 0$ be such that $\varepsilon \leq \min\{\varepsilon', \Theta\}$. Let $p^* \in N_C^P(x)$ be such that $\|x - \overline{x}\| < \varepsilon$ and $\|p^*\| \leq \varepsilon \leq \Theta$. From the choice of $\varepsilon \leq \varepsilon'$ from (c) it holds that

$$0 \geq \langle J[J^*(p^*) - r^{-1}(x' - x)], x' - x\rangle, \quad \forall x' \in C \text{ such that } \|x' - \overline{x}\| < \varepsilon,$$

which means that (b) holds with ε and r. □

To the rest of this section as prox-regularity of a set C at a point $\overline{x}$ will be considered Proposition 1.4.1 (c) for $\Theta = 1$, namely the property

$$(1.58)\quad \begin{cases} \text{there exist } \varepsilon > 0, r > 0 \text{ such that for all } x \in C \text{ with } \|x - \overline{x}\| < \varepsilon \\ \text{and for all } p^* \in N_C^P(x) \text{ with } \|p^*\| \le 1, \\ 0 \ge \langle J[J^*(p^*) - r^{-1}(x' - x)], x' - x\rangle, \quad \forall x' \in C \text{ such that } \|x' - \overline{x}\| < \varepsilon. \end{cases}$$

In this regard, we refine the definition of prox-regularity as follows

Definition 1.4.2. A closed set $C \subset X$ is called (*metrically*) *prox-regular* (with respect to the uniformly convex norm $\|\cdot\|$) or $\|\cdot\|$-*prox-regular* at $\overline{x} \in C$ provided there exist $\varepsilon > 0$ and $r > 0$ such that for all $x \in C$ and for all $p^* \in N_C^P(x)$ with $\|x - \overline{x}\| < \varepsilon$ and $\|p^*\| < 1$ the point x is a nearest point of $\{x' \in C : \|x' - \overline{x}\| < \varepsilon\}$ to $x + rJ^*(p^*)$.

The metrical aspect is due to the fact that the proximal normal cone $N_C^P(\cdot)$ is related to the norm $\|\cdot\|$ and depends on it in general when the latter is not a Hilbert norm. Whenever there is no ambiguity concerning either the norm $\|\cdot\|$ or the involvement of the proximal normal cone $N_C^P(\cdot)$, we will merely say that C is prox-regular at $\overline{x}$.

The crucial fact which needs to be emphasized here is that the real number r of the definition (for which the closed ball $B[x+rJ^*(p^*), r\|p^*\|]$ touches the set $C \cap B(\overline{x}, \varepsilon)$ at the point x, when x is a boundary point of C with $\|x - \overline{x}\| < \varepsilon$) does not depend on either the neighbouring point x or the proximal normal functional $p^* \in N_C^P(x)$ with $\|p^*\| < 1$.

Let us recall that for prox-regular sets in uniformly convex Banach space with norm with moduli of uniform convexity and smoothness of power type a list of equivalent characterizations is given in Theorem 1.3.25.

A set-valued mapping $T : X \rightrightarrows X^*$ is said to be J-*hypomonotone* of degree $t \ge 0$ on a subset $U \subset X$ (see Definition 1.3.6) if for any $(x_i, x_i^*) \in \operatorname{gph} T := \{(x, x^*) \in U \times X^* : x^* \in T(x)\}$, $i = 1, 2$, one has

$$\langle J[J^*(x_1^*) - t(x_2 - x_1)] - J[J^*(x_2^*) - t(x_1 - x_2)], x_2 - x_1\rangle \le 0.$$

We know that prox-regular property (1.58) means that the indicator function ψ_C of C is J-primal lower regular at $\overline{x}$. Another characterization can be given in terms of the mapping J_σ in place of the normalized duality mapping J.

Proposition 1.4.3. Let $\sigma > 1$ be a real number. The prox-regular property (1.58) is equivalent to any one of the following:
(i_σ) there exist $\varepsilon > 0$ and $r > 0$ such that for all $x \in C$ and all $p^* \in N_C^P(x)$ with $\|x - \overline{x}\| < \varepsilon$ and $\|p^*\| \le 1$

$$0 \ge \langle J_\sigma[J_\sigma^{-1}(p^*) - r^{-1}(x' - x)], x' - x\rangle \quad \forall x' \in C \text{ with } \|x' - \overline{x}\| < \varepsilon;$$

(i'_σ) there exist $\varepsilon > 0$ and $r > 0$ such that for all $x \in C$, $t \geq 0$, and $p^* \in N^P_C(x)$ with $\|x - \overline{x}\| < \varepsilon$ and $\|p^*\| \leq r^{\sigma-1} t^{\sigma-1}$

$$0 \geq \langle J_\sigma[J_\sigma^{-1}(p^*) - t(x' - x)], x' - x\rangle \quad \forall x' \in C \text{ with } \|x' - \overline{x}\| < \varepsilon.$$

Proof. Note first that property (1.58) is equivalent to the following one:
for any $x, x' \in C \cap B(\overline{x}, \varepsilon)$, any $t \geq 0$, $p^* \in N^P_C(x)$ with $\|p^*\| \leq rt$,

$$0 \geq \langle J[J^*(p^*) - t(x' - x)], x' - x\rangle. \tag{1.59}$$

Indeed, taking any $t > 0$, $x, x' \in C \cap B(\overline{x}, \varepsilon)$ and $p^* \in N^P_C(x)$ with $\|p^*\| \leq rt$ we have by (1.58)

$$0 \geq \langle J[J^*(r^{-1}t^{-1}p^*) - r^{-1}(x' - x)], x' - x\rangle,$$

which is easily seen to be equivalent to (1.59) according to the positive homogeneity of J and J^*. Further it is obvious that (1.59) still holds for $t = 0$. The implication from (1.58) to (1.59) is then established. The converse follows from similar arguments.

With this new formulation, we are able to prove that (1.58) is equivalent to (i'_σ).

Suppose that (i'_σ) holds. Observe first that for any non zero $p^* \in X^*$, one has $J_\sigma^{-1}(p^*) = J^*\left(\|p^*\|^{\frac{2-\sigma}{\sigma-1}} p^*\right)$ and for any non zero $u \in X$, one has $J_\sigma(u) = J(\|u\|^{\sigma-2}u)$. Hence, putting $\sigma' := \frac{2-\sigma}{\sigma-1}$, for any non zero $p^* \in X^*$ and any $t \geq 0$ one has the equivalences

$$\begin{aligned} \langle J[J^*(p^*) - t(x' - x)], x' - x\rangle \leq 0 \quad &\Leftrightarrow \\ \langle J[\|p^*\|^{-\sigma'} J_\sigma^{-1}(p^*) - t(x' - x)], x' - x \leq 0 \quad &\Leftrightarrow \\ \langle J[J_\sigma^{-1}(p^*) - t\|p^*\|^{\sigma'}(x' - x)], x' - x\rangle \leq 0 \quad &\Leftrightarrow \\ \langle J_\sigma[J_\sigma^{-1}(p^*) - t\|p^*\|^{\sigma'}(x' - x)], x' - x\rangle \leq 0. & \end{aligned} \tag{1.60}$$

Fix now any $t > 0$, any $x, x' \in C \cap B(\overline{x}, \varepsilon)$, and any non zero $p^* \in N^P_C(x)$ such that $rt \geq \|p^*\|$. The latter inequality ensures that $rt\|p^*\|^{\sigma'} \geq \|p^*\|^{\frac{1}{\sigma-1}}$, i.e., $\|p^*\| \leq r^{\sigma-1}t'^{\sigma-1}$ for $t' := t\|p^*\|^{\sigma'}$, thus (i'_σ) with $t' = t\|p^*\|^{\sigma'}$ in place of t yields (1.60). By the above equivalences we obtain the inequality (1.59).

Conversely, suppose that (1.59) is fulfilled. Fix any $t \geq 0$, $x, x' \in C \cap B(\overline{x}, \varepsilon)$, and any non zero $p^* \in N^P_C(x)$ such that $\|p^*\| \leq r^{\sigma-1}t^{\sigma-1}$. One has $\|p^*\|^{\frac{1}{\sigma-1}} \leq rt$ and hence $\|p^*\| \leq rt\|p^*\|^{1-\frac{1}{\sigma-1}}$, which by (1.59), with $t' := t\|p^*\|^{1-\frac{1}{\sigma-1}} = t\|p^*\|^{-\sigma'}$ in place of t (where as above $\sigma' := \frac{2-\sigma}{\sigma-1}$), yields

$$0 \geq \langle J[J^*(p^*) - t'(x' - x)], x' - x\rangle.$$

By (1.60) this is equivalent to

$$0 \geq \langle J_\sigma[J_\sigma^{-1}(p^*) - t'\|p^*\|^{\sigma'}(x' - x)], x' - x\rangle,$$

which is exactly

$$0 \geq \langle J_\sigma[J_\sigma^{-1}(p^*) - t(x' - x)], x' - x\rangle.$$

The latter being still true for $p^* = 0$, we obtain (i'_σ).

The equivalence between (i'_σ) and (i_σ) can be argued like in the first part of the proof. □

Anticipating Definition 1.4.4, we call property (i'_σ) J_σ-primal lower regularity of the function ψ_C.

In the line of Definition 1.4.2 and Definition 1.3.26, the concept of prox-regularity on a *global* point of view is defined as follows: a closed subset C of X is (*metrically*) *uniformly r-prox-regular* or r-uniformly prox-regular if whenever $x \in C$ and $p^* \in N_C^P(x)$ with $\|p^*\| < 1$, then x is the unique nearest point of C to $x + rJ^*(p^*)$.

Let us recall that for uniformly prox-regular sets in uniformly convex Banach space with norm with moduli of uniform convexity and smoothness of power type a list of equivalent characterizations is given in Theorem 1.3.27.

In Definition 1.3.5 we introduced the J-plr concept for functions. Here we slightly extend this concept keeping the name. Now it reduces in the Hilbert setting to the Poliquin's primal lower nice concept, see Poliquin [136], and Levi, Poliquin and Thibault [111].

Definition 1.4.4. A lower semicontinuous function $f : X \to \mathbb{R} \cup \{+\infty\}$ is *J-primal lower regular* (*J-plr*) at $\overline{x} \in \operatorname{dom} f$ if there exist positive constants ε, r and Θ such that

$$f(y) \geq f(x) + \langle J[J^*(p^*) - t(y - x)], y - x\rangle \tag{1.61}$$

for all $x, y \in B(\overline{x}, \varepsilon)$, all $p^* \in \partial_p f(x)$, and all $t \geq \Theta$ such that $\|p^*\| \leq rt$.

It is easily seen that a J-plr function according to Definition 1.3.5 is a J-plr function according to Definition 1.4.4 (with arbitrary $\Theta > 0$). Further, Definition 1.3.5 and Definition 1.4.4 are equivalent for indicator functions of closed sets (see the proof of Proposition 1.4.1).

Below we will consider J-plr functions according to Definition 1.4.4. If f is J-plr at $\overline{x}$ with some positive constants ε and r then it is so for any constants $0 < \varepsilon' \leq \varepsilon$ and $0 < r' \leq r$. If the lower semicontinuous function f is J-plr at $\overline{x} \in \operatorname{dom} f$ with positive constants ε, r and Θ, then obviously

$$\langle J[J^*(p^*) - t(y - x)] - J[J^*(q^*) - t(x - y)], y - x\rangle \leq 0 \tag{1.62}$$

for all $x, y \in B(\overline{x}, \varepsilon)$, for all $p^* \in \partial_p f(x)$, $q^* \in \partial_p f(y)$, and all $t \geq \Theta$ such that $\max\{\|p^*\|, \|q^*\|\} \leq rt$. This is the analog of the hypomonotonicity of certain truncations of $\partial_p f$ that characterizes primal lower-nice functions in Hilbert spaces: see Poliquin [136], Poliquin, Rockafellar and Thibault [111], Bernard, Thibault and Zagrodny [16] and the references therein.

1.4.1 Normal and tangential regularity properties of prox-regular sets

For a closed set $C \subset X$, the Mordukhovich limiting (or basic) normal cone $N^L_C(x)$ is defined (see Mordukhovich [124]) as the weak* sequential outer limit

$$(1.63) \qquad N^L_C(x) = {}^{w^*-\mathrm{seq}}\operatorname*{Lim\,sup}_{y \to x} N^P_C(y) := \{w^* - \lim x^*_n : x^*_n \in N^P_C(x_n), x_n \in C \to x\}.$$

The following result appears in Ioffe [89, p. 188] with an approximation in X^* with respect to the weak star topology, but merely under the local uniform convexity of the norm. For completeness we sketch below how Ioffe's arguments also yield to the following approximation with respect to the strong topology.

Proposition 1.4.5. Let C be a closed subset of X with $x \in C$ and let $x^* \in N^F_C(x)$. Then for any $\varepsilon > 0$ there exist $u_\varepsilon \in C$ and $u^*_\varepsilon \in N^P_C(u_\varepsilon)$ such that

$$(1.64) \qquad \|u_\varepsilon - x\| < \varepsilon \quad \text{and} \quad \|u^*_\varepsilon - x^*\| < \varepsilon.$$

In fact the result uses only the Fréchet differentiability outside of zero of the norm $\|\cdot\|$ and of its dual norm.

Proof. We may suppose that $\|x^*\| = 1$. By definition there exists some function ρ from $[0, +\infty[$ into $[0, +\infty[$ with $\lim_{t \downarrow 0} \rho(t) = 0$ and such that

$$(1.65) \qquad \langle x^*, y - x\rangle \le \rho(\|y - x\|)\,\|y - x\| \quad \text{for all } y \in C.$$

As this will not to be confusing for the reader, we denote the dual norm also by $\|\cdot\|$. Putting $h := J^* x^*$ we have

$$(1.66) \qquad \langle x^*, h\rangle = \|x^*\|\|h\| = 1.$$

Further, by (1.65) we see that $x + th \notin C$ for positive t small enough. By Lau theorem (see Lau [108]) for any such t we may choose some $h_t \in X$ such that $\|h_t - h\| < t$ and such that the nearest point of $x + th_t$ in C exists, say $u_t \in C$. Writing u_t in the form $u_t = x + tv_t$ we have

$$(1.67) \qquad t\|v_t - h_t\| = d_C(x + th_t) \le t\|h_t\| \quad \text{and hence} \quad \|v_t\| \le 2\|h_t\| \le 2(1 + t).$$

Then taking (1.65) into account we have

$$\langle x^*, tv_t\rangle \le \rho(t\|v_t\|)\,\|tv_t\| \le 2t\rho(t\|v_t\|)(1 + t)$$

which yields

$$(1.68) \qquad \langle x^*, v_t\rangle \le 2\rho(t\|v_t\|)(1 + t).$$

Now observe by the first inequality of (1.67) that

$$\|h - v_t\| \le \|h - h_t\| + \|h_t - v_t\| \le \|h - h_t\| + \|h_t\| \le \|h\| + 2\|h - h_t\| < 1 + 2t$$

and hence for $w_t := (1 + 2t)^{-1}(h - v_t)$ we have $\|w_t\| \le 1$. Therefore

$$(1 + 2t)^{-1}[1 - 2\rho(t\|v_t\|)(1 + t)] \le \langle x^*, w_t\rangle \le 1,$$

the first inequality being due to (1.66) and (1.68) and the second to the inequality $\|w_t\| \le 1$. Consequently we have

$$\langle x^*, w_t\rangle \to 1 = \langle x^*, h\rangle$$

and then since the dual norm $\|\cdot\|$ of X^* is Fréchet differentiable at the point x^* of the unit sphere of X^* and since $\|w_t\| \le 1$ and $\|h\| = 1$, Šmulian lemma (see, e.g., Fabian, Habala, Hájek, Montesinos, Pelant and Zizler [71, Lemma 8.4]) says that $\|w_t - h\| \to 0$ as $t \downarrow 0$, which is equivalent to $\|v_t\| \to 0$.

For positive t sufficiently small, the functional $u_t^* := \|x + th_t - u_t\|^{-1} J(x + th_t - u_t)$ is by definition a unit functional in $N_C^P(u_t)$ and $\langle u_t^*, x + th_t - u_t\rangle = \|x + th_t - u_t\|$, which is equivalent, by the equality $u_t = x + tv_t$, to $\langle u_t^*, h_t - v_t\rangle = \|h_t - v_t\|$. Since $\|v_t\| \to 0$ and $h_t \to h$, we derive that $\langle u_t^*, h_t\rangle \to \|h\| = 1$ and hence

$$\langle u_t^*, h\rangle \to 1 = \langle x^*, h\rangle.$$

Remembering that $\|u_t^*\| = 1$ and $\|x^*\| = 1$ and that the norm is Fréchet differentiable at the point h of the unit sphere of X, Šmulian lemma again entails that $\|u_t^* - x^*\| \to 0$ as $t \downarrow 0$. The proof is then complete because obviously one also has $\|u_t - x\| \to 0$ as $t \downarrow 0$. □

Taking the latter proposition and (1.63) into account, we see (as in Ioffe [89]) that the Mordukhovich limiting normal cone above coincides with the (sequential) limiting normal cone obtained as above by replacing $N_C^P(\cdot)$ with $N_C^F(\cdot)$, i.e.,

$$N_C^L(x) = {}^{w^*\text{-seq}}\operatorname*{Lim\,sup}_{y \to x} N_C^F(y). \tag{1.69}$$

(We must emphasize that one of the important advantages of the expression of $N_C^L(x)$ in the form of (1.69) is that it makes available the Mordukhovich normal cone in the more general context of Asplund space, see Mordukhovich [124], Mordukhovich and Shao [126]).

We will also need in Subsection 1.4.3 below the strong outer limit

$$N_C^{L,s}(x) := \operatorname*{Lim\,sup}_{y \to x} N_C^P(y) := \{\lim x_n^* : x_n^* \in N_C^P(x_n), x_n \in C \to x\} \tag{1.70}$$

for $x \in C$. From Proposition 1.4.5 we also have $N_C^{L,s}(x) = \operatorname*{Lim\,sup}_{y \to x} N_C^F(y)$.

The concept of tangential regularity (see Clarke [44]) involved below is related to the Bouligand and Clarke tangent cones. A vector v of X is in the *Bouligand tangent cone* (or *contingent cone*) $K_C(x)$ to C at $x \in C$ if there exists a sequence of positive numbers $(t_n)_n$ converging to 0 and a sequence $(v_n)_n$ of X converging to v such that $x + t_n v_n \in C$ for all n. The *Clarke tangent cone* $T_C(x)$ can be also defined in a sequential way. A vector $v \in T_C(x)$ provided that for any sequence $(x_n)_n$ in C converging to x and for any sequence of positive numbers $(t_n)_n$ converging to 0 there exists a sequence $(v_n)_n$ of X converging to v with $x_n + t_n v_n \in C$ for all n. One always has the inclusion $T_C(x) \subset K_C(x)$. When $T_C(x) \equiv K_C(x)$, one says that the set C is *Clarke* or *tangentially regular* at x.

Through the Clarke tangent cone, the *Clarke normal cone* $N_C^{Cl}(x)$ of C at $x \in C$ is defined as the negative polar of the latter, that is,

$$N_C^{Cl}(x) = \{x^* \in X^* : \langle x^*, v \rangle \le 0, \quad \forall v \in T_C(x)\}.$$

In any Asplund space (hence in particular in our context), we have (see Mordukhovich [124], Mordukhovich and Shao [126])

$$N_C^{Cl}(x) = \overline{\text{co}}^*\big({}^{w^*\text{-seq}} \operatorname*{Lim\,sup}_{y \to x} N_C^F(y) \big), \tag{1.71}$$

where $\overline{\text{co}}^*$ denotes the weak star closed convex hull.

When the Mordukhovich limiting normal cone of C at x coincides with the Fréchet (resp. proximal) normal cone at x, one says that C is *normally regular* at x with respect to the Fréchet (resp. the proximal) normal cone. Obviously the normal regularity with respect to the proximal normal cone implies the normal regularity with respect to the Fréchet one (because of the inclusion $N_C^P(x) \subset N_C^F(x)$).

We recall (see Bounkhel and Thibault [33]) that any one of the two above normal regularities entails the (Clarke) tangential regularity. We refer to Bounkhel and Thibault [33] for the development of a detailed comparison between the above concepts of normal and tangential regularities and various others.

The next theorem is one among the results at the heart of this subsection.

Theorem 1.4.6. Assume that the closed set C is prox-regular at $\bar{x} \in C$. Then there exists a neighbourhood U of $\bar{x}$ such that for any $x \in U \cap C$ one has the following normal regularity

$$N_C^P(x) = N_C^F(x) = N_C^L(x) = N_C^{Cl}(x)$$

and hence

$$\partial_P d_C(x) = \partial_F d_C(x) = \partial_L d_C(x) = \partial_C d_C(x),$$

that is, the distance function itself is subdifferentially regular at all points of $U \cap C$.

So, the set C is in particular tangentially regular at $x \in U \cap C$. Further, one has $N^{\beta}_C(x) \subset N^P_C(x)$ for $\beta = 2$.

Proof. By assumption, there exist positive real numbers ε, r with $\varepsilon < 1/2$ such that for every $x \in C \cap B(\overline{x}, \varepsilon)$ and every $x^* \in N^P_C(x)$ with $\|x^*\| \le 1$, we have $\|x + tJ^*(x^*) - x'\| \ge \|tJ^*(x^*)\|$ for any $x' \in C \cap B(\overline{x}, \varepsilon)$ and $t \in]0, r]$. Fix any $x, x' \in B(\overline{x}, \varepsilon) \cap C$ and $x^* \in N^P_C(x)$ with $\|x^*\| \le 1$. Then

$$\|x + tJ^*(x^*) - x'\|^2 \ge \|tJ^*(x^*)\|^2 \quad \text{for any } t \in]0, r]. \tag{1.72}$$

On the other hand for any $t \in]0, r]$, according to the Fréchet differentiability of $\|\cdot\|^2$, we have

$$\begin{aligned} &\|x + tJ^*(x^*) - x'\|^2 \\ = \ &\|tJ^*(x^*)\|^2 + 2\int_0^1 \langle J[tJ^*(x^*) + \theta(x - x')], x - x'\rangle \, d\theta \\ = \ &\|tJ^*(x^*)\|^2 + 2t\langle x^*, x - x'\rangle + 2\int_0^1 \langle J[tJ^*(x^*) + \theta(x - x')] - J[tJ^*(x^*)], x - x'\rangle \, d\theta. \end{aligned}$$

Combining this with (1.72) we obtain

$$\langle x^*, x' - x\rangle \le \frac{1}{t}\|x' - x\| \int_0^1 \|J[tJ^*(x^*) + \theta(x - x')] - J[tJ^*(x^*)]\| \, d\theta. \tag{1.73}$$

In the properties of the duality mapping J recalled in the beginning of the section we saw that J is uniformly continuous over bounded subsets of X. Therefore, denoting by ω_{r+1} the modulus of uniform continuity of J over the bounded set $(r+1)B$, that is,

$$\omega_{r+1}(\tau) := \sup\{\|J(u) - J(u')\| : u, u' \in (r+1)B, \ \|u - u'\| \le \tau\} \quad \text{for } \tau > 0,$$

we have $\omega_{r+1}(\tau) \underset{\tau \downarrow 0}{\longrightarrow} 0$ and (1.73) entails

$$\langle x^*, x' - x\rangle \le \frac{1}{t}\|x' - x\| \, \omega_{r+1}(\|x - x'\|). \tag{1.74}$$

Fix now $x \in C \cap B(\overline{x}, \varepsilon)$ and $x^* \in N^L_C(x)$ and fix also any $\eta > 0$. Let $x^*_n \in N^P_C(x_n)$ (see (1.63)) such that $x_n \in C \underset{n\to\infty}{\longrightarrow} x$ and $x^*_n \underset{n\to\infty}{\overset{w^*}{\longrightarrow}} x^*$. Choose a real number $\gamma > 0$ such that $\|x^*_n\| \le \gamma$ for all integers n and choose a positive $\alpha < \varepsilon - \|x - \overline{x}\|$ such that $\omega_{r+1}(\tau) \le \frac{r\eta}{\gamma}$ for all positive $\tau < \alpha$. Take any $x' \in B(x, \alpha) \cap C$. We have $x' \in B(\overline{x}, \varepsilon)$ and, for n large enough, $x_n \in B(\overline{x}, \varepsilon) \cap C$ and $\|x' - x_n\| < \alpha$. By (1.74) for n large enough we then have

$$\langle x^*_n, x' - x_n\rangle \le \frac{1}{r}\|x^*_n\| \, \|x' - x_n\| \, \omega_{r+1}(\|x' - x_n\|) \le \frac{1}{r}\gamma\|x' - x_n\|\frac{r\eta}{\gamma}$$

and hence passing to the limit for $n \to \infty$ we obtain

$$\langle x^*, x' - x\rangle \leq \eta \|x' - x\|.$$

The latter inequality being true for all $x' \in B(x, \alpha) \cap C$, it means that $x^* \in N_C^F(x)$. So far we have proved that for any $x \in B(\overline{x}, \varepsilon) \cap C$, we have $N_C^L(x) \subset N_C^F(x)$. As the reverse inclusion is also true according to (1.69), we have in fact the equality

$$N_C^L(x) = N_C^F(x) \quad \text{for all } x \in B(\overline{x}, \varepsilon) \cap C. \tag{1.75}$$

Moreover, by (1.71), we know that $N_C^{Cl}(\cdot) = \overline{\text{co}}^*(N_C^L)(\cdot)$, where $\overline{\text{co}}^*$ denotes the weak* closed convex hull in X^*. Since $N_C^F(x)$ is convex and (strongly) closed (see Bounkhel and Thibault [33]), we deduce from the equality (1.75) that we even have

$$N_C^F(x) = N_C^L(x) = N_C^{Cl}(x) \quad \text{for all } x \in B(\overline{x}, \varepsilon) \cap C. \tag{1.76}$$

Let us now prove that the three cones in (1.76) are also equal to the cone of proximal normal functionals to C. In our uniformly convex setting where the norm $\|\cdot\|$ of X is uniformly convex and uniformly smooth, we know according to Proposition 1.4.5 that for any $x \in B(\overline{x}, \varepsilon) \cap C$, $x^* \in N_C^F(x)$ with $\|x^*\| < 1$, there exists $x_n \underset{n\to\infty}{\longrightarrow} x$ with $x_n \in C$, and $x_n^* \in N_C^P(x_n)$ such that $x_n^* \xrightarrow[n\to\infty]{\|\cdot\|} x^*$. For n large enough we have $x_n \in B(\overline{x}, \varepsilon)$ and $\|x_n^*\| < 1$. For any such integer n, for any $t \in]0, r]$ and $x' \in B(\overline{x}, \varepsilon) \cap C$, we have by (1.72)

$$\|x_n + tJ^*(x_n^*) - x'\|^2 \geq \|tJ^*(x_n^*)\|^2$$

which gives, by passing to the limit and by the continuity of J^*,

$$\|x + tJ^*(x^*) - x'\|^2 \geq \|tJ^*(x^*)\|^2,$$

i.e, $x^* \in N_C^P(x)$ thanks to the local character (1.57) of primal proximal normal vector. So for any fixed $x \in B(\overline{x}, \varepsilon) \cap C$ we obtain that $N_C^F(x) \subset N_C^P(x) \subset N_C^L(x)$, which combined with (1.76) gives the equalities

$$N_C^P(x) = N_C^F(x) = N_C^L(x) = N_C^{Cl}(x).$$

These equalities also ensure

$$\partial_C d_C(x) \subset N_C^{Cl}(x) \cap B^* = N_C^P(x) \cap B^* = \partial_p d_C(x),$$

the last equality being due to Proposition 1.3.2 (see also Clarke [44] for the first inclusion). Consequently, we have

$$\partial_p d_C(x) = \partial_F d_C(x) = \partial_L d_C(x) = \partial_C d_C(x).$$

Finally, on the one hand the equality between $N_C^F(x)$ and $N_C^{Cl}(x)$ ensures that C is tangentially regular at x (see Bounkhel and Thibault [33]), and on the other hand, for $\beta = 2$, since one always has $N_C^\beta(x) \subset N_C^F(x)$, the inclusion $N_C^\beta(x) \subset N_C^P(x)$ follows. □

We deduce from Theorem 1.4.6 a proximal normal formula; Lim sup below is like in (1.70) the strong outer limit, i.e., for a multivalued mapping $T : X \rightrightarrows X^*$

$$\operatorname*{Lim\,sup}_{y \xrightarrow{y \in D} x} T(y) := \{\lim y_n^* : y_n^* \in T(y_n), y_n \in D \to x\}.$$

Proposition 1.4.7. Assume that the moduli of uniform convexity and of uniform smoothness of the norm $\|\cdot\|$ of X are of power type. If C is prox-regular at $\overline{x} \in X$, then there exists a neighbourhood U of $\overline{x}$ such that for any $x \in U \cap C$,

$$N_C^P(x) = N_C^{Cl}(x) = \operatorname*{Lim\,sup}_{y \xrightarrow{y \notin C} x} \mathbb{R}_+ \nabla^F d_C(y)$$

or equivalently,

$$\partial_p d_C(x) = \partial_C d_C(x) = [0, 1] \operatorname*{Lim\,sup}_{y \xrightarrow{y \notin C} x} \nabla^F d_C(y),$$

where ∇^F denotes the Fréchet derivative and ∂_C is the Clarke subdifferential.

Proof. Remember that the prox-regularity of C entails the Fréchet differentiability of d_C^2 and the single-valuedness and continuity of P_C on an open neighbourhood U of $\overline{x}$, see Theorem 1.3.25. Fix any $x \in U \cap C$.

Let us prove that $N_C^P(x) \subset \operatorname*{Lim\,sup}_{y \xrightarrow{y \notin C} x} \mathbb{R}_+ \nabla^F d_C(y)$.

Recall first that a nonzero continuous linear functional x^* is in $N_C^P(x)$ if and only if there exists $u \in U \setminus C$ such that $x \in P_C(u)$ and $x^* = \lambda J(u - x)$ for some $\lambda > 0$. Then for any $y \in U \cap \{ tu + (1 - t)x : t \in]0, 1] \}$, one has

$$J(u - x) \in \mathbb{R}_+ J(y - x) = \mathbb{R}_+ \nabla^F d_C(y)$$

and thus for n large enough $x^* \in \mathbb{R}_+ \nabla^F d_C(y_n)$ with $y_n := x + (u - x)/n$, which proves the desired inclusion.

Now for the reverse inclusion, note that for any $y \in U \setminus C$,

$$\mathbb{R}_+ \nabla^F d_C(y) = \mathbb{R}_+ J(y - P_C(y)) \subset N_C^P(P_C(y)).$$

Making $y \to x$, we have by continuity of P_C that $P_C(y) \to x$ and hence, $x^* \in \operatorname*{Lim\,sup}_{y \xrightarrow{y \notin C} x} \mathbb{R}_+ \nabla^F d_C(y)$ entails that $x^* \in \operatorname*{Lim\,sup}_{x' \to x} N_C^P(x') = N_C^P(x)$, the last equality coming from the normal regularity established in Theorem 1.4.6. So the reverse inclusion is proved and of course, by normal regularity with respect to the proximal normal cone (see Theorem 1.4.6) we obtain $N_C^{Cl}(x) = N_C^P(x)$.

The equivalent formulation with the subdifferential of d_C comes from the equality $\partial_p d_C(x) = N_C^P(x) \cap B^*$ (see Proposition 1.3.2) and from the equality $\partial_p d_C(x) = \partial_C d_C(x)$ in Theorem 1.4.6. □

Remark 1.4.8. The proximal normal formula of the above proposition is more precise than the more general one found in Borwein and Giles [25, Theorem 4] for any closed set C in a reflexive Banach space with weaker assumptions on the norm.

1.4.2 Epigraphs of J-primal lower regular functions

The following proposition and its proof essentially reproduces ideas of Bernard and Thibault from [17, Proposition 4.8].

Proposition 1.4.9. Assume that the moduli of uniform convexity and uniform smoothness of the norm $\|\cdot\|$ of X are of power type. If a lower semicontinuous function f is J-plr at $\overline{x} \in \operatorname{dom} f$, then its epigraph $\operatorname{epi} f$ is prox-regular at $(\overline{x}, f(\overline{x}))$.

We will need the following lemma, given in Clarke, Ledyaev, Stern and Wolenski [45, Exercise 2.1 (d)] in the Hilbert setting. The proof is essentially the same in our general context (without the power types of the norm), but the proximal normal cone is no more identical to $N_C^\beta(\cdot)$ with $\beta = 2$.

Lemma 1.4.10. For any $x \in \operatorname{dom} f$ and $\alpha > f(x)$, the implication

$$(x^*, 0) \in N^P_{\operatorname{epi} f}(x, \alpha) \Rightarrow (x^*, 0) \in N^P_{\operatorname{epi} f}(x, f(x))$$

holds true.

Proof. Recall that by convention the square of the norm in $X \times \mathbb{R}$ is defined as $|||(x, r)|||^2 = \|x\|^2 + r^2$. Let $(x^*, 0) \in N^P_{\operatorname{epi} f}(x, \alpha)$. This means that for all positive t close enough to 0,

$$\inf_{(x',\alpha')\in \operatorname{epi} f} \left\{\|x + tJ^*(x^*) - x'\|^2 + (\alpha' - \alpha)^2\right\} = \|tJ^*(x^*)\|^2. \tag{1.77}$$

Choose any $\delta > 0$ with $\delta < \alpha - f(x)$. Fix any $(x', \alpha') \in \operatorname{epi} f \cap B((x, f(x)), \delta)$. We have $\alpha' - f(x) < \delta < \alpha - f(x)$, hence $\alpha' < \alpha$ so $(x', \alpha) \in \operatorname{epi} f$. From (1.77) we obtain

$$\|x + tJ^*(x^*) - x'\|^2 + (\alpha - \alpha)^2 \geq \|tJ^*(x^*)\|^2, \tag{1.78}$$

so

$$\inf_{(x',\alpha')\in \operatorname{epi} f\cap B((x,f(x)),\delta)} \left\{\|x + tJ^*(x^*) - x'\|^2 + (\alpha' - f(x))^2\right\} \geq$$
$$\inf_{(x',\alpha')\in \operatorname{epi} f\cap B((x,f(x)),\delta)} \left\{\|x + tJ^*(x^*) - x'\|^2\right\} \geq \|tJ^*(x^*)\|^2,$$

the last inequality being due to (1.78). Since the concept of proximal normal is local (see (1.57)), we conclude that $(x^*, 0) \in N^P_{\operatorname{epi} f}(x, f(x))$. □

Proof of Proposition 1.4.9. By definition of J-plr function (see Definition 1.4.4), there exist $\Theta, \varepsilon, r > 0$ such that for any $t \geq \Theta$, any $x \in B(\overline{x}, \varepsilon)$ and $x^* \in \partial_p f(x)$ with $\|x^*\| \leq rt$,

$$f(y) \geq f(x) + \langle J[J^*(x^*) - t(y-x)], y - x\rangle \quad \text{for all } y \in B(\overline{x}, \varepsilon). \tag{1.79}$$

Take $(x, \alpha) \in \operatorname{epi} f$ and $(x^*, -\lambda) \in N^P_{\operatorname{epi} f}(x, \alpha)$ where $\|x - \overline{x}\| < \varepsilon$, $|\alpha - f(\overline{x})| < \varepsilon$ and $\|(x^*, -\lambda)\| \leq 1$. Fix any $(x', \alpha') \in \operatorname{epi} f$ with $x' \in B(\overline{x}, \varepsilon)$.
If $\lambda > 0$, then $\alpha = f(x)$ and $\lambda^{-1}x^* \in \partial_p f(x)$ with $\|\lambda^{-1}x^*\| \leq 1/\lambda \leq rt$ for every $t \geq \max(\Theta, 1/(\lambda r))$. Hence by (1.79)

$$\alpha' \geq f(x') \geq f(x) + \langle J[J^*(\lambda^{-1}x^*) - t(x' - x)], x' - x\rangle$$

which entails

$$0 \geq \langle J[J^*(x^*) - \lambda t(x' - x)], x' - x\rangle - \lambda(\alpha' - \alpha) - \lambda t(\alpha' - \alpha)^2.$$

For any $t' \geq \max(\Theta, 1/r)$, since $t' \geq \max(\lambda\Theta, 1/r)$ because $1 \geq \lambda$, the latter inequality with $t = t'/\lambda$, according to (1.56), implies

$$0 \geq \langle J_{X\times\mathbb{R}}[J^*_{X^*\times\mathbb{R}}(x^*, -\lambda) - t'((x', \alpha') - (x, \alpha))], (x', \alpha') - (x, \alpha)\rangle. \tag{1.80}$$

If $\lambda = 0$, we get by Lemma 1.4.10 that $(x^*, 0) \in N^P_{\operatorname{epi} f}(x, f(x))$. Since $N^P(\cdot) \subset N^F(\cdot)$, by the approximation result in Ioffe [89, p. 190] (see also Mordukhovich [124, Lemma 2.37]) there exist sequences $(u_n, f(u_n)) \to (x, f(x))$, $(u_n^*, -\lambda_n) \in N^F_{\operatorname{epi} f}(u_n, f(u_n))$ such that $\lambda_n > 0$ and $\|(u_n^*, -\lambda_n) - (x^*, 0)\| \to 0$. By Proposition 1.4.5, for each integer n choose $(x_n, \alpha_n) \in \operatorname{epi} f$ and then $(y_n^*, -\mu_n) \in N^P_{\operatorname{epi} f}(x_n, \alpha_n)$ such that

$$\|(u_n, f(u_n)) - (x_n, \alpha_n)\| < \lambda_n/2 \quad \text{and} \quad \|(y_n^*, -\mu_n) - (u_n^*, -\lambda_n)\| < \lambda_n/2.$$

The latter inequality ensures in particular $\mu_n > (\lambda_n/2) > 0$ and hence $\alpha_n = f(x_n)$. Consequently for $x_n^* := \mu_n^{-1}y_n^*$ we have $x_n^* \in \partial_P f(x_n)$, $(x_n, f(x_n)) \to (x, f(x))$, $\mu_n \downarrow 0$, and $\|\mu_n x_n^* - x^*\| \to 0$. For n large enough, say $n \geq N$, we have $\|x_n - \overline{x}\| < \varepsilon$ and $\mu_n < 1$. Suppose for a moment that $x^* \neq 0$. Putting

$$t_n := \max\left\{\frac{1}{\mu_n}\Theta, \frac{1}{\|x^*\|}\frac{\|x_n^*\|}{r}\right\}$$

for $n \geq N$, we see that $t_n \geq \max\{\Theta, \|x_n^*\|/r\}$ since $1/\mu_n > 1$ and $1/\|x^*\| \geq 1$, and hence applying (1.79) with $y = x'$ we have

$$f(x') \geq f(x_n) + \langle J[J^*(x_n^*) - t_n(x' - x_n)], x' - x_n\rangle.$$

Multiplying this inequality by μ_n and taking the limit we obtain

$$0 \geq \langle J[J^*(x^*) - \rho(x' - x)], x' - x\rangle$$

for $\rho := \max\{\Theta, 1/r\}$. This yields in particular

$$0 \geq \langle J[J^*(x^*) - \rho(x' - x)], x' - x\rangle - \rho(\alpha' - \alpha)^2,$$

i.e.,

$$0 \geq \langle J_{X\times\mathbb{R}}[J^*_{X^*\times\mathbb{R}}(x^*, 0) - \rho((x', \alpha') - (x, \alpha))], (x', \alpha') - (x, \alpha)\rangle, \tag{1.81}$$

and it is obvious that the inequality continues to hold for $x^* = 0$. Hence both (1.80) and (1.81) hold with $t' = \rho$, which entails that the truncated set-valued mapping $N^{P\,1}_{\operatorname{epi} f}$ is J-hypomonotone of degree ρ on $B(\overline{x}, \varepsilon')$ and hence by Theorem 1.3.25, the set $\operatorname{epi} f$ is prox-regular at $(\overline{x}, f(\overline{x}))$. □

Remembering that the Clarke subdifferential operator ∂_C and the Mordukhovich limiting subdifferential ∂_L satisfy, for any lower semicontinuous function f (see Clarke [44], Mordukhovich [124])

$$\partial_C f(x) = \{x^* : (x^*, -1) \in N^{Cl}_{\operatorname{epi} f}(x, f(x))\} \text{ and } \partial_L f(x) = \{x^* : (x^*, -1) \in N^{L}_{\operatorname{epi} f}(x, f(x))\},$$

from Theorem 1.4.6 we have the following

Corollary 1.4.11. Assume that the moduli of uniform convexity and smoothness of the norm $\|\cdot\|$ of X are of power type. If f is J-plr at $\overline{x}$, then there exists a neighbourhood U of $\overline{x}$ such that for $\beta = 2$ one has

$$\partial_\beta f(x) \subset \partial_p f(x) = \partial_F f(x) = \partial_L f(x) = \partial_C f(x) \quad \text{for all } x \in U.$$

By considering the following property of f being uniformly J-plr, we obtain similarly the *uniform prox-regularity* of $\operatorname{epi} f$. The function f is *uniformly J-plr* on X provided there exist positive constants Θ, r such that for any $t \geq \Theta$, any $x \in X$ and $x^* \in \partial_p f(x)$ with $\|x^*\| \leq rt$,

$$f(y) \geq f(x) + \langle J[J^*(x^*) - t(y - x)], y - x\rangle \qquad \text{for all } y \in X.$$

Proposition 1.4.12. Assume that the moduli of uniform convexity and smoothness of the norm $\|\cdot\|$ of X are of power type. If f is uniformly J-plr on X with parameters r, Θ, then $\operatorname{epi} f$ is uniformly r'-prox-regular for some $r' \geq \frac{1}{2}\min\{1/\Theta, r\}$.

Proof. According to the proof of Proposition 1.4.9, the truncated set-valued mapping $N^{P\,1}_{\operatorname{epi} f}$ is J-hypomonotone of degree $\rho = \max\{\Theta, 1/r\}$. With $\delta = \rho^{-1}$, that fact is equivalent to $N^{P\,\delta}_{\operatorname{epi} f}$ being J-hypomonotone of degree 1 which by Proposition 1.3.31 allows us to conclude. □

That proposition allows us to establish the J-plr property of the basic function equal to the opposite of the square of the norm whenever the latter has moduli of convexity and smoothness of power type.

Proposition 1.4.13. Assume that the moduli of uniform convexity and smoothness of the norm $\|\cdot\|$ of X are of power type. Then the opposite function of the square of the norm, say $-\|\cdot\|^2$, is uniformly J-plr on X and hence in particular $\operatorname{epi}(-\|\cdot\|^2)$ is uniformly r-prox-regular for some $r \geq \frac{1}{4}$.

Proof. For any $x \in X$, $y = x + h \in X$, $x^* = \partial(-\|\cdot\|^2)(x) = -2J(x)$, by the convexity of $\|\cdot\|^2$, we may write

$$\|x\|^2 = \|y - h\|^2 \geq \|y\|^2 + \langle 2J(y), -h\rangle.$$

So, we have

$$-\|y\|^2 \quad \geq \quad -\|x\|^2 + \langle -2J(y), h\rangle$$

which yields, according to the equality $-2y = J^*(x^*) - 2(y - x)$, that

$$-\|y\|^2 \geq -\|x\|^2 + \langle J[J^*(x^*) - 2(y - x)], y - x\rangle.$$

This entails the J-hypomonotonicity of degree 2 of $N_C^{P\,1}$ where $C := \operatorname{epi}(-\|.\|^2)$, hence the conclusion follows from Proposition 1.3.31 and Proposition 1.4.12.

In order to compare J-plr functions with primal lower-nice functions, we suppose in Proposition 1.4.14 and Corollary 1.4.15 below that the modulus of uniform convexity of the norm is of power type $q = 2$. Let us recall the definition of primal lower nice functions, introduced by Poliquin in [136] in $\mathbb{R}^n$ is studied in the Hilbert setting with further developments by Levi, Poliquin and Thibault in [111].

Below we will consider an presubdifferential operator ∂ that associates with each function $f : X \to \mathbb{R} \cup \{+\infty\}$ a multivalued mapping $\partial f : X \rightrightarrows X^*$ and satisfies various properties commonly fulfilled by the usual subdifferentials on appropriate spaces (see Definition 2.1.1 and the discussion after it). Here we will just assume that for any function f from X into $\mathbb{R} \cup \{+\infty\}$, one has

(1.82) $$\partial_P f(x) \subset \partial f(x) \quad \text{for all } x \in X.$$

Let us recall that a lower semicontinuous function f is ∂-pln at $\overline{x} \in \operatorname{dom} f$ (see Definition 1.1.4) if there are $\varepsilon, c, \Theta > 0$ such that whenever $t \geq \Theta$, $x^* \in \partial f(x)$ with $\|x - \overline{x}\| < \varepsilon$ and $\|x^*\| \leq ct$, one has

$$f(x') \geq f(x) + \langle x^*, x' - x\rangle - \frac{t}{2}\|x' - x\|^2 \text{ for all } x' \in B(\overline{x}, \varepsilon).$$

Obviously, for $\beta = 2$ one has $\partial f(\cdot) \cap ctB^* \subset \partial_\beta f(\cdot)$ on $B(\overline{x}, \varepsilon)$ for such a function f.

Examples of subdifferential operators that contain the proximal subdifferential operator ∂_p are given by ∂_p itself, by ∂_F, and hence also by any of the many subdifferential operators that contain ∂_F.

Proposition 1.4.14. Assume that the modulus of uniform convexity of the norm $\|\cdot\|$ of X is of power type $q = 2$. If f is ∂-pln at $\bar{x}$ for $\partial(\cdot) \supset \partial_p(\cdot)$, then f is J-plr at $\bar{x}$.

Proof. Under the assumptions, there exist positive numbers ε, r, Θ such that for any $(x, x^*) \in \operatorname{gph} \partial_p f$ with $\|x - \bar{x}\| < \varepsilon$, $\|x^*\| \le rt$ and $t \ge \Theta$,

$$f(x') \ge f(x) + \langle x^*, x' - x\rangle - \frac{t}{2}\|x' - x\|^2 \text{ for all } x' \in B(\bar{x}, \varepsilon). \tag{1.83}$$

Further, for any $x^* \in X^*, h \in X, t' > 0$ we may write

$$\begin{aligned}\langle x^*, h\rangle &= \langle J[J^*(x^*)] - J[J^*(x^*) - t'h], h\rangle + \langle J[J^*(x^*) - t'h], h\rangle \\ &\ge \frac{K}{t'}\|t'h\|^2 + \langle J[J^*(x^*) - t'h], h\rangle,\end{aligned}$$

the inequality being due to (1.52) with $q = 2$. Hence taking $h = x' - x$, we have for any $t' > 0$,

$$\langle x^*, x' - x\rangle - Kt'\|x' - x\|^2 \ge \langle J[J^*(x^*) - t'(x' - x)], x' - x\rangle.$$

Combining the latter inequality and (1.83), we see that for any $t' \ge \Theta/(2K)$, $(x, x^*) \in \operatorname{gph} \partial_p f$ with $\|x - \bar{x}\| < \varepsilon$ and $\|x^*\| \le (2Kr)t'$, we have

$$f(x') \ge f(x) + \langle J[J^*(x^*) - t'(x' - x)], x' - x\rangle \quad \text{for all } x' \in B(\bar{x}, \varepsilon),$$

that is, f is J-plr with parameters $\varepsilon, 2rK, \Theta/(2K)$. □

Corollary 1.4.15. Assume that the moduli of uniform convexity and smoothness of the norm $\|\cdot\|$ of X are of power type. Under the inclusion $\partial f(\cdot) \subset \partial_C f(\cdot)$ and the assumption of Proposition 1.4.14, the lower semicontinuous function f is ∂-pln at $\bar{x}$ if and only if f is ∂_p-pln at $\bar{x}$ and then for $\beta = 2$ one has

$$\partial_\beta f(x) = \partial_p f(x) = \partial f(x) = \partial_F f(x) = \partial_L f(x) = \partial_C f(x)$$

for all x in a neighbourhood of $\bar{x}$.

Proof. Suppose that f is ∂-pln at $\bar{x}$. On the one hand, by (1.82), it is ∂_p-pln at $\bar{x}$. On the other hand, by definition of ∂-pln property the inclusion $\partial f(x) \subset \partial_\beta f(x)$ holds for $\beta = 2$ and for x in a neighbourhood of $\bar{x}$, and by Proposition 1.4.14, the function f is also J-plr at $\bar{x}$. Therefore using (1.82) and Corollary 1.4.11 we obtain the equalities of the corollary.

Conversely, if f is ∂_p-pln at $\bar{x}$, then Proposition 1.4.14 again entails that f is J-plr at $\bar{x}$ and hence by Proposition 1.4.9 and Theorem 1.4.6 we have $\partial_p f(x) = \partial_C f(x)$ for all x in some neighbourhood of $\bar{x}$. Combining the latter equality with (1.82) and the inclusion assumption $\partial f(\cdot) \subset \partial_C f(\cdot)$ we obtain that $\partial f(x) = \partial_p f(x)$ for all x near $\bar{x}$. So f is ∂-pln at $\bar{x}$. □

We have a symmetrical result when the modulus of smoothness is $s = 2$.

Proposition 1.4.16. Assume that the moduli of uniform convexity and uniform smoothness of the norm $\|\cdot\|$ are of power type and that the power type of smoothness is $s = 2$. If f is J-plr at $\bar{x}$, then f is ∂-pln at $\bar{x}$ for any subdifferential ∂ satisfying the inclusions $\partial_P f(\cdot) \subset \partial f(\cdot) \subset \partial_C f(\cdot)$.

Proof. By Corollary 1.4.11, for all x near $\bar{x}$ we have $\partial_C f(x) = \partial_P f(x)$ and hence $\partial_P f(x) = \partial f(x)$. The J-plr assumption then entails that there exist positive numbers ε, r, Θ such that for any $t \geq \Theta$, $(x, x^*) \in \operatorname{gph} \partial_P f$ with $\|x - \bar{x}\| < \varepsilon$, $\|x^*\| \leq rt$,

$$f(x') \geq f(x) + \langle J[J^*(x^*) - t(x' - x)], x' - x\rangle \quad \text{for all } x' \in B(\bar{x}, \varepsilon). \tag{1.84}$$

Fix now any $t \geq \Theta$, $x', x \in B(\bar{x}, \varepsilon)$ and $x^* \in \partial f(x)$ with $\|x^*\| \leq rt$. By (1.53) since $s = 2$ (the power type of smoothness of the norm), we have

$$\langle J[J^*(x^*) - t(x' - x)] - J[J^*(x^*)], -t(x' - x)\rangle \leq L\|t(x' - x)\|^2,$$

which is equivalent to

$$\langle J[J^*(x^*) - t(x' - x)], x' - x\rangle \geq \langle x^*, x' - x\rangle - tL\|x' - x\|^2.$$

Then according to (1.84) we obtain

$$f(x') - f(x) \geq \langle x^*, x' - x\rangle - tL\|x' - x\|^2.$$

So f is ∂-pln at $\bar{x}$. □

Corollary 1.4.17. Assume that the moduli of uniform convexity and smoothness of the norm $\|\cdot\|$ of X are of power type and that the power type of smoothness is $s = 2$. Under the assumption of Proposition 1.4.16, if the lower semicontinuous function f is J-plr at $\bar{x}$ then for $\beta = 2$

$$\partial_\beta f(x) = \partial_P f(x) = \partial f(x) = \partial_F f(x) = \partial_L f(x) = \partial_C f(x)$$

for all x in a neighbourhood of $\bar{x}$.

Proof. This a consequence of Corollary 1.4.11 and Proposition 1.4.16. □

Uniformly ∂ primal lower-nice functions are defined analogously to uniformly J-plr functions, i.e., f is *uniformly ∂-pln* if there are $c, \Theta > 0$ such that whenever $t \geq \Theta$, $x^* \in \partial f(x)$ with $\|x^*\| \leq ct$,

$$f(x') \geq f(x) + \langle x^*, x' - x\rangle - \frac{t}{2}\|x' - x\|^2 \text{ for all } x' \in X.$$

The two previous results above proved for ∂-pln functions at a point $\bar{x}$ of X also hold for uniformly ∂-pln functions: just replace "∂-pln (resp. J-plr) at $\bar{x}$" with "uniformly ∂-pln (resp. uniformly J-plr)".

1.4.3 Prox-regularity and N-hyporegularity

Let $N(\cdot)$ be a given normal cone concept (e.g., $N^F(\cdot)$, $N^P(\cdot)$, etc.) associated with a subdifferential operator ∂; i.e., for any closed set C with indicator function ψ_C, one has $N_C = \partial\psi_C$ and $\partial f(x) = \{x^* \in X^* : (x^*, -1) \in N_{\mathrm{epi}\, f}(x, f(x))\}$.

The normal cone $N(\cdot)$ is assumed to satisfy, for any closed set C, the inclusion $N^P_C(\cdot) \subset N_C(\cdot)$.

Following Poliquin and Rockafellar [137] and the adaptation in Bernard and Thibault [14], the set $C \subset X$ is *N-hyporegular* at $\overline{x} \in C$ if there exist $\varepsilon, r > 0$ such that for any $x^* \in N_C(x)$ with $\|x - \overline{x}\| < \varepsilon$, $\|x^*\| \le 1$,

$$0 \ge \langle x^*, x' - x\rangle - \frac{1}{2r}\|x' - x\|^2 \text{ for all } x' \in B(\overline{x}, \varepsilon) \cap C. \tag{1.85}$$

Of course, (1.85) holds if and only if the truncated normal cone $N_C(\cdot) \cap B^*$ is *hypomonotone* near $\overline{x}$ in the usual sense, that is, there exist some $\varepsilon, r > 0$ such that for all $x_i \in C \cap B(\overline{x}, \varepsilon)$ and $x_i^* \in N_C(x_i) \cap B^*$, $i = 1, 2$ one has

$$\langle x_1^* - x_2^*, x_1 - x_2\rangle \ge -\frac{1}{r}\|x_1 - x_2\|^2. \tag{1.86}$$

In Bernard and Thibault [14] a set C satisfying (1.85) with $N(\cdot)$ has been called prox-regular with respect to the normal cone $N(\cdot)$. Here we prefer to use the name of *N-hyporegular set* because of the characterization (1.86) as the hypomonotonicity of $N_C(\cdot) \cap B^*$ (which has nothing to do with the metric projection mapping) and to reserve the name of prox-regularity set merely to translate the regularity property of the metric projection mapping in Definition 1.4.2 or in (c) of Theorem 1.3.25.

We observe by (1.85) that C is N-hyporegular at $\overline{x}$ if and only if its indicator function ψ_C is ∂-pln at $\overline{x}$. Remember (see Theorem 1.3.25 (l)) that C is prox-regular at $\overline{x}$ if and only if ψ_C is J-plr at $\overline{x}$. Similarly, the uniform N-hyporegularity means that the inequality (1.85) holds for all $x, x' \in C$ and $x^* \in N_C(x) \cap B^*$ and this corresponds to the uniform ∂-pln property for the indicator function ψ_C. In the same way, the uniform prox-regularity of C is equivalent to the uniform J-plr property of the function ψ_C. Remember that q and s are the moduli of convexity and of smoothness respectively, and that they satisfy $1 < s \le 2 \le q$. The case $q = 2$ corresponds for instance to L^p spaces with $p \in]1, 2]$ and $s = 2$ to L^p spaces with $p \in [2, +\infty[$.

In the case $q = 2$, we have from results in the previous subsection the following result.

Corollary 1.4.18. Assume that the moduli of uniform convexity and smoothness of the norm $\|\cdot\|$ of X are of power type. When the power type of the modulus of convexity is $q = 2$, a closed set $C \subset X$ is N-hyporegular at $\overline{x} \in C$ (resp. uniformly N-hyporegular) for a normal cone N which satisfies $N^P(\cdot) \subset N(\cdot) \subset N^{Cl}(\cdot)$ if and only if it is N^P-hyporegular at $\bar{x}$ (resp. uniformly N^P-hyporegular), and then it is also prox-regular at $\bar{x}$ (resp. uniformly prox-regular) and for $\beta = 2$ one has $N_C^\beta(x) = N_C^P(x) = N_C(x) = N_C^{Cl}(x)$ for any x in some neighbourhood of $\bar{x}$ (resp. any $x \in C$).

Proof. It is a direct consequence of Proposition 1.4.14 and Corollary 1.4.15. □

In the case $s = 2$ we have the reverse implication. If the moduli of uniform convexity and smoothness are of power type with the power type of smoothness $s = 2$, then its proof directly follows from Proposition 1.4.16 and Corollary 1.4.17. In fact the proof below shows that the result holds without requiring that the modulus of uniform convexity be of power type.

Proposition 1.4.19. Assume that the modulus of smoothness of the norm $\|\cdot\|$ of X is of power type $s = 2$. If a closed set $C \subset X$ is prox-regular at $\bar{x} \in C$ (resp. uniformly prox-regular), then it is N-hyporegular at $\bar{x}$ (resp. uniformly N-hyporegular) for any normal cone N which satisfies $N^P(\cdot) \subset N(\cdot) \subset N^{Cl}(\cdot)$, and further for $\beta = 2$ one has

$$N_C^\beta(x) = N_C^P(x) = N_C(x) = N_C^{Cl}(x)$$

for any x in some neighbourhood of $\bar{x}$ (resp. any $x \in C$).

Proof. We prove the result when we have local prox-regularity at a point. The prox-regularity assumption entails by Proposition 1.3.4 that there exist positive numbers ε, r such that for any $x \in B(\bar{x}, \varepsilon) \cap C$ and for any $x^* \in N_C^P(x)$ with $\|x^*\| \le 1$,

$$0 \ge \langle J[J^*(x^*) - r^{-1}(x' - x)], x' - x\rangle, \quad \forall x' \in C \text{ with } \|x' - \bar{x}\| < \varepsilon. \tag{1.87}$$

Fix now any $x', x \in C \cap B(\bar{x}, \varepsilon)$ and $x^* \in N_C^P(x)$ with $\|x^*\| \le 1$. By (1.53) since $s = 2$ (the power type of smoothness of the norm), we have

$$\langle J[J^*(x^*) - r^{-1}(x' - x)] - J[J^*(x^*)], -r^{-1}(x' - x)\rangle \le L\|r^{-1}(x' - x)\|^2,$$

which is equivalent to

$$\langle J[J^*(x^*) - r^{-1}(x' - x)], x' - x\rangle \ge \langle x^*, x' - x\rangle - Lr^{-1}\|x' - x\|^2.$$

Then according to (1.87) we obtain

$$0 \ge \langle x^*, x' - x\rangle - Lr^{-1}\|x' - x\|^2.$$

So the set C is also N^P-hyporegular at $\bar{x}$.

Further, putting $\beta = 2$, on the one hand the N^P-hyporegularity ensures us that for an appropriate neighbourhood U of $\bar{x}$, for any $x \in U \cap C$, the inclusion $N_C^P(x) \subset N_C^\beta(x)$ holds, and on the other hand by Theorem 1.4.6, we know that $N_C^\beta(x) \subset N_C^P(x) = N_C^{Cl}(x)$. Since by assumption we have the inclusions $N_C^P(\cdot) \subset N_C(\cdot) \subset N_C^{Cl}(\cdot)$, we obtain that for $\beta = 2$

$$N_C^\beta(x) = N_C^P(x) = N_C(x) = N_C^{Cl}(x) \quad \text{for all } x \in U \cap C$$

and consequently, C is also N-hyporegular. □

For a family of norms $(\|\cdot\|_i)_{i\in I}$, let us denote by $N_i^?$ the normal cone $N^?$ (e.g., N^F, N^{Cl}, N^P, etc.) obtained by using the norm $\|\cdot\|_i$ in its definition and computation. The $N_i^P(\cdot)$ cones for a closed set C will be denoted by $N_i^P(C;.)$. When the norms $\|\cdot\|_i$ and $\|\cdot\|_j$ are equivalent, some normal cone concepts $N^?$ yield $N_i^? = N_j^?$, as in the cases $N^? = N^F, N^\beta, N^{L,s}, N^L, N^{Cl}$ for instance. We will speak of $(\|\cdot\|_i, N)$-hyporegularity of a set C when the property (1.85) is satisfied with the norm $\|\cdot\| = \|\cdot\|_i$ and with the cone $N = N_i^?$. When C is $(\|\cdot\|_i, N^?)$-hyporegular for any $i \in I$, we will say that C is $N^?$-hyporegular relatively to the family $(\|\cdot\|_i)_{i\in I}$. Similarly, the concept of prox-regularity in Definition 1.4.2 is a priori norm dependent, so for a given norm $\|\cdot\|$ it has been said in that definition that the set C is $\|\cdot\|$-prox-regular.

From now on in the remaining of this subsection we address the problem of comparing the $(\|\cdot\|_i, N^?)$-hyporegularities and $\|\cdot\|_i$-prox-regularities of a set C for norms in a given family $(\|\cdot\|_i)_{i\in I}$, and of comparing accordingly the normal cones $N_i^P(C;.)$ for such a set C.

Note first that by passing to the limit in (1.85) and by using the comments preceding and following Proposition 1.4.5, it is easily seen, for a norm $\|\cdot\|$ which is both uniformly convex and uniformly smooth, that whenever C is $(\|\cdot\|, N^?)$-hyporegular at a point $\overline{x}$ with $N^? = N^P$ or $N^? = N^F$, then it is $(\|\cdot\|, N^{L,s})$-hyporegular at $\overline{x}$ with the same parameters. Further for some neighbourhood U of $\overline{x}$ and for $\beta = 2$ the inclusion $N_C^L(x) \subset N_C^\beta(x)$ holds for all $x \in U$ and hence, $N_C^\beta(x) = N_C^F(x) = N_C^L(x) = N_C^{Cl}(x)$ and C is $(\|\cdot\|, N^?)$-hyporegular at $\overline{x}$ for any $N^?$ with $N^?(\cdot) \subset N^{Cl}(\cdot)$.

Suppose that $(\|\cdot\|_i)_{i\in I}$ is a family of equivalent norms which are Fréchet differentiable outside zero and such that for some given $i_0 \in I$, the norm $\|\cdot\|_{i_0}$ is uniformly smooth and uniformly convex. From the previous remark we know that if C is $(\|\cdot\|_{i_0}, N^P)$-hyporegular at $\overline{x} \in C$ there exists some $\gamma > 0$ and some open neighbourhood U of $\overline{x}$ such that for all $x, x' \in U \cap C$ and $x^* \in N_{i_0}^{L,s}(C;x)$ with $\|x^*\|_{i_0} \le 1$ (denoting the dual norm of $\|\cdot\|_{i_0}$ in the same way)

$$0 \ge \langle x^*, x' - x\rangle - \gamma\|x' - x_0\|_{i_0}^2.$$

Then for the same open neighbourhood U of $\overline{x}$, for each $i \in I$ there exists some $\gamma_i > 0$ such that for all $x, x' \in U \cap C$ and $x^* \in N^F(C;x)$ with $\|x^*\|_i \le 1$ one has

$$0 \ge \langle x^*, x' - x\rangle - \gamma_i\|x' - x\|_i^2.$$

Therefore for any $i \in I$ the set C is $(\|\cdot\|_i, N^P)$-hyporegular at any $x \in U \cap C$ and for $\beta = 2$ one has

$$N_i^P(C;x) \subset N_C^F(x) = N_C^\beta(x) = N_C^{Cl}(x) \quad \text{for all } x \in U \cap C. \tag{1.88}$$

By specializing to certain families $(\|\cdot\|_i)_{i\in I}$ we have the next results of the subsection. Here is the first one.

Proposition 1.4.20. Suppose that $(\|\cdot\|_i)_{i\in I}$ is a family of equivalent norms which are uniformly smooth and uniformly convex with the moduli of convexity $\delta_{\|\cdot\|_i}$ of power type $q = 2$ for all $i \in I$.

If the closed set C is $(\|\cdot\|_{i_0}, N^P)$-hyporegular at $\overline{x} \in C$ for some $i_0 \in I$, then the following hold:
(a) the set C is $(\|\cdot\|_i, N)$-hyporegular at $\overline{x}$ for any $i \in I$, and for any normal cone N such that $N(\cdot) \subset N^{Cl}(\cdot)$;
(b) the set C is $\|\cdot\|_i$-prox-regular at $\overline{x}$ for any $i \in I$;
(c) there exists some neighbourhood U of $\overline{x}$ such that for any $x \in U \cap C$ and any $i \in I$ and for $\beta = 2$ the equalities $N_C^{\beta}(x) = N_i^P(C; x) = N_C^{Cl}(x)$ hold, and C is $(\|\cdot\|_i, N)$-hyporegular at x for any normal cone $N_C(\cdot)$ with $N_C^{\beta}(\cdot) \subset N_C(\cdot) \subset N_C^{Cl}(\cdot)$.

Proof. The first point has been seen above. The second and third ones come from the first one, from (1.88), and from Corollary 1.4.18. □

We specialize further, to the case of $\mathbb{R}^n$ endowed with the family $(\|\cdot\|_p)_{p>1}$ of classical l_p-norms, say $\|x\|_p = (\sum_{k=1}^{p} |x_k|^p)^{1/p}$.

Proposition 1.4.21. Suppose that $X = \mathbb{R}^n$ and $(\|\cdot\|_p)_{p>1}$ is the family of l_p-norms with $p > 1$.
1) If for some $p_0 > 1$ the set C is $(\|\cdot\|_{p_0}, N^P)$-hyporegular at $\overline{x} \in C$, then
(a) the set C is $(\|\cdot\|_p, N)$-hyporegular at $\overline{x}$ for any $p > 1$ and any normal cone N such that $N(\cdot) \subset N^{Cl}(\cdot)$;
(b) the set C is $\|\cdot\|_p$-prox-regular at $\overline{x}$ for any $p \in]1, 2]$;
(c) there is some neighbourhood U of $\overline{x}$ such that for all $p \in]1, 2]$ and $p' > 2$ and for $\beta = 2$ one has

$$N_{p'}^P(C; x) \subset N_p^P(C; x) = N_C^{\beta}(x) = N_C^{Cl}(x) \quad \text{for any } x \in U.$$

2) If for some $p_0 \geq 2$ the set C is $\|\cdot\|_{p_0}$-prox-regular at $\overline{x} \in C$, then
(a') the set C is $(\|\cdot\|_p, N)$-hyporegular at $\overline{x}$ for any $p > 1$ and any normal cone N such that $N(\cdot) \subset N^{Cl}(\cdot)$;
(b') the set C is $\|\cdot\|_p$-prox-regular at $\overline{x}$ for any $p \in]1, 2]$;
(c') there is some neighbourhood U of $\overline{x}$ such that for all $p \in]1, 2]$ and $p' > 2$ and for $\beta = 2$ one has

$$N_{p'}^P(C; x) \subset N_p^P(C; x) = N_{p_0}^P(C; x) = N_C^{\beta}(x) = N_C^{Cl}(x) \quad \text{for any } x \in U.$$

Proof. Under the assumption of the first case, the property (a) like (a) in Proposition 1.4.20 has been seen ahead this latter. The second property (b) is a consequence of (b) in Proposition 1.4.20 since for any $p \in]1, 2]$, the norm $\|\cdot\|_p$ has modulus of convexity of power type $q = 2$. Taking the open neighbourhood U given by (c) of

Proposition 1.4.20 for the family of norms $(\|\cdot\|_p)_{p\in]1,2]}$, only the first inclusion of the third property (c) remains to be argued. It is a consequence of C being $(\|\cdot\|_{p'}, N^P)$-hyporegular at any $x \in U \cap C$ for any $p' > 2$ according to the property (a), which entails for $\beta = 2$ the inclusion $N^P_{p'}(C; x) \subset N^\beta_C(x)$ for $\beta = 2$ and for any $x \in U \cap C$.

Now concerning the second case, suppose that for some $p_0 \geq 2$ the set C is $\|\cdot\|_{p_0}$-prox-regular at $\bar{x} \in C$. We first observe that Definition 1.4.2 furnishes some open neighbourhood U of $\bar{x}$ such that C is $\|\cdot\|_{p_0}$-prox-regular at any point $x \in C \cap U$. Since the norm $\|\cdot\|_{p_0}$ is uniformly smooth with modulus of smoothness $s = 2$, Proposition 1.4.19 yields that C is N^F-hyporegular at any $x \in C \cap U$ relatively to the family $(\|\cdot\|_p)_{p>1}$ and that $N^P_{p_0}(C; x) = N^F_C(x)$ for any $x \in U$. Then for any $p > 1$ the set C is $(\|\cdot\|_p, N^P)$-hyporegular at any $x \in C \cap U$ and then from the first case and what precedes, we have the second and third properties (b′) and (c′). □

1.4.4 Comparison of normal cones

The proposition of this subsection compares, for any closed subset C of X, the cone $N^P_C(\cdot)$ of proximal normal functionals with the normal cone $N^\beta_C(\cdot)$ when β is the power type of the modulus of uniform convexity or of smoothness of the norm $\|\cdot\|$ of X.

Proposition 1.4.22. Let C be a closed subset of X and let $x \in C$.
(a) If the modulus of convexity of the norm $\|\cdot\|$ of X is of power type q, then

$$N^q_C(x) \subset N^P_C(x).$$

(b) Similarly, if the modulus of smoothness of the norm $\|\cdot\|$ of X is of power type s, then

$$N^P_C(x) \subset N^s_C(x).$$

Proof. Suppose that the modulus of convexity of $\|\cdot\|$ is of power type q. From (1.50), there exists some constant $K > 0$ such that for all $x, y \in X$,

$$\|x + y\|^q \geq \|x\|^q + q\langle J_q(x), y\rangle + K\|y\|^q.$$

Take $x^* \in N^q_C(x)$, $x^* \neq 0$. There is some $\gamma, \varepsilon > 0$ such that for any $x' \in B(x, \varepsilon) \cap C$, $\langle x^*, x' - x\rangle \leq \gamma\|x' - x\|^q$. For any $x' \in B(x, \varepsilon) \cap C$ we then have for each $t > 0$,

$$\|x + tJ_q^{-1}(x^*) - x'\|^q \geq \|tJ_q^{-1}(x^*)\|^q + q\langle J_q(tJ_q^{-1}(x^*)), x - x'\rangle + K\|x - x'\|^q,$$

that is, by the equality in (1.47)

$$\|x + tJ_q^{-1}(x^*) - x'\|^q \geq \|tJ_q^{-1}(x^*)\|^q + t^{q-1}q\langle J_q(J_q^{-1}(x^*)), x - x'\rangle + K\|x - x'\|^q,$$

hence

$$\|x + tJ_q^{-1}(x^*) - x'\|^q \geq \|tJ_q^{-1}(x^*)\|^q - t^{q-1}q\gamma\|x' - x\|^q + K\|x - x'\|^q$$

so, whenever $t \leq (K/(\rho\gamma))^{1/(q-1)}$, we obtain

$$\|x + tJ_q^{-1}(x^*) - x'\|^q \geq \|tJ_q^{-1}(x^*)\|^q.$$

Because of the local character (1.57) of primal normal vector, the latter equality entails that $J_q^{-1}(x^*) \in PN_C(x)$, that is, $J(J_q^{-1}(x^*)) \in N_C^P(x)$. As $J(J_q^{-1}(x^*)) \in \mathbb{R}_+ x^*$, the first inclusion follows.

To prove now (b) fix any $x^* \in N_C^P(x)$ with $x^* \neq 0$. By definition of $N_C^P(\cdot)$, there exists $\delta > 0$ such that for any $t \in]0, \delta]$ and any $x' \in C$

$$\|x + tJ^{-1}(x^*) - x'\|^s \geq \|tJ^{-1}(x^*)\|^s.$$

From (1.51), there exists some constant $L > 0$ such that, for every $x, y \in X$,

$$\|x + y\|^s \leq \|x\|^s + s\langle J_s(x), y\rangle + L\|y\|^s.$$

From the two previous estimations we derive for any $x' \in C$

$$s\langle J_s(\delta J^{-1}(x^*)), x - x'\rangle + L\|x - x'\|^s \geq 0,$$

that is, according to the equality in (1.47)

$$s\delta^{s-1}\langle J_s(J^{-1}(x^*)), x - x'\rangle + L\|x - x'\|^s \geq 0.$$

As $J_s(J^{-1}(x^*)) \in \mathbb{R}_+ x^*$, we deduce that there exists some $\sigma > 0$ such that

$$\langle x^*, x' - x\rangle \leq \sigma\|x' - x\|^s \quad \text{for all } x' \in C.$$

This entails that $x^* \in N_C^s(x)$ and the inclusion of (b) is proved. □

1.4.5 Preservation of hyporegularity and prox-regularity

This subsection is devoted to the study of the preservation of prox-regularity for the intersection of finitely many sets and for the inverse image. The characterizations of prox-regularity in Theorem 1.3.25 by the hypomonotonicity of the proximal normal cone as well as by local single valuedness and by continuity of the metric projection mapping have been crucial in the study of Edmond and Thibault [68] of differential inclusions of sweeping process type governed by nonconvex prox-regular sets. The regularization of such differential inclusions in Hilbert space in Thibault [159] uses the property (f) of the same theorem. Below we will take advantage of the property (1.58)

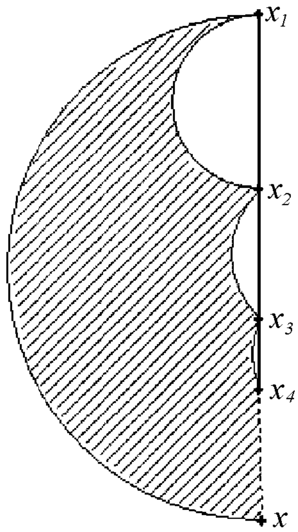

Fig. 1. Intersection of prox-regular sets

to investigate the stability of metrical prox-regularity under the above set-operations. In fact we will start with the stability of N^P-hyporegularity through the property (1.85) which is in the line of property (1.58).

To see first that the intersection of finitely many prox-regular sets may not be prox-regular, consider the example in $\mathbb{R}^2$ with its euclidian norm illustrated in Fig. 1, where the set C_1 is the whole line (x_1x_2), $\{x_i\}_i$ being a sequence of points on C_1 that converges to some $x \in C_1$, and the set C_2 is the hatched surface delimited on one side by the closed arc $\widehat{x_1x}$ of the half circle with radius $R = \|x_1 - x\|/2$ and on its other side by the arcs of the circles (*all with the same radius* $r < R$) $\widehat{x_ix_{i+1}}$, $i \in \mathbb{N}$, so that x belongs also to C_2. It is easily seen that while C_1 and C_2 are both prox-regular at x (they are even uniformly prox-regular), their intersection $C_1 \cap C_2$ is not prox-regular at the point x.

To provide general sufficient conditions under which the prox-regularity of intersection or inverse image is preserved, we need first to recall the concept of calmness. Translating the concept of calmness of a multivalued mapping in Rockafellar and Wets [152] we say that the intersection of a finite family of sets $\{C_k\}_{k=1}^m$ is *metrically calm* at a point $\overline{x} \in \bigcap_{k=1}^m C_k$ provided there exist a constant $\gamma > 0$ and a neighbourhood U of $\overline{x}$ such that

$$(1.89)\qquad d\left(x, \bigcap_{k=1}^m C_k\right) \le \gamma\,(d(x,C_1) + \cdots + d(x,C_m)) \quad \text{for all } x \in U.$$

Let now $F : X \to Y$ be a mapping from X into another uniformly convex Banach space Y and let D be a subset of Y and $\overline{x} \in F^{-1}(D)$. As above we say that the mapping

F is *metrically calm* at $\bar{x}$ relatively to the set D when there exist a constant $\gamma > 0$ and a neighbourhood U of $\bar{x}$ such that

$$d(x, F^{-1}(D)) \leq \gamma d(F(x), D) \quad \text{for all } x \in U, \tag{1.90}$$

where $d(x, A)$ is the distance from x to the set A.

Proposition 1.4.23. Let $\{C_k\}_{k=1}^m$ be a finite family of closed sets of X and let D be a closed set of Y.
(a) If all sets C_k are N^P-hyporegular at a point $\bar{x}$ of all the sets C_k and if the intersection is metrically calm at $\bar{x}$, then this intersection set $\bigcap_{k=1}^{m} C_k$ is N^P-hyporegular at $\bar{x}$ too.
(b) If a mapping $F : X \to Y$ is of class $C^{1,1}$ around a point $\bar{x} \in F^{-1}(D)$ and metrically calm at $\bar{x}$ relatively to D and if D is N^P-hyporegular at $F(\bar{x})$, then the set $F^{-1}(D)$ is N^P-hyporegular at $\bar{x}$.

Proof. Assume that each set C_k is N^P-hyporegular at $\bar{x}$ and put $C := \bigcap_{k=1}^{m} C_k$. By definition of N^P-hyporegularity (see 1.85) it is easily seen that there exist some real number $r > 0$ and some open neighbourhood U of $\bar{x}$ such that for each $k = 1, \dots, m$ we have for all $x \in C_k \cap U$ and $u^* \in N^P_{C_k}(x) \cap B^*$

$$\langle u^*, x' - x \rangle \leq (2r)^{-1} \|x' - x\|^2 \quad \text{for all } x' \in C_k \cap U. \tag{1.91}$$

Restricting the neighbourhood U if necessary, we may suppose that the inequality (1.89) holds upon U. Fix any $x \in C \cap U$ and take any $x^* \in N^P_C(x)$ with $\|x^*\| \leq 1$. We have that $x^* \in N^F_C(x)$ and hence $x^* \in \partial_F d_C(x)$ since one knows that $\partial_F d_C(x) = N^F_C(x) \cap B^*$ (see, e.g., Mordukhovich [124, Corollary 1.96]). This inequality (1.89) and the definition of Fréchet subgradient easily yields that $x^* \in \gamma \partial_F (d_{C_1} + \dots + d_{C_m})(x)$ and hence in particular $x^* \in \gamma \partial_L (d_{C_1} + \dots + d_{C_m})(x)$. The functions d_{C_k} being Lipschitzian, the formula of the Mordukhovich subdifferential of a finite sum of locally Lipschitz functions (see, e.g., Mordukhovich [124, Theorem 3.36]) ensures us that there exist $u_k^* \in \partial_L d_{C_k}(x)$ such that $\gamma^{-1} x^* = u_1^* + \dots + u_m^*$. Restricting again the neighbouhood U, by Theorem 1.4.6 we have $u_k^* \in \partial_p d_{C_k}(x) = N^P_C(x) \cap B^*$ which by (1.91) gives $\langle u_k^*, x' - x \rangle \leq (2r)^{-1} \|x' - x\|^2$ for every $x' \in C_k \cap U$. Consequently for any $x' \in C \cap U$ we obtain

$$\langle \gamma^{-1} x^*, x' - x \rangle \leq m(2r)^{-1} \|x' - x\|^2$$

and this obviously implies that the intersection set C is N^P-hyporegular at $\bar{x}$, that is, (a) is proven.

Let us now establish (b). Fix some open convex neighbourhood U of $\bar{x}$ over which (1.90) holds and over which the mapping F as well as its derivative $DF(\cdot)$ are Lipschitz with Lipschitz constants K and K_1 respectively. For $S := F^{-1}(D)$ fix any $x \in U \cap S$ and $x^* \in N_S^P(x)$ with $\|x^*\| \le 1$. As above we have $x^* \in \partial_F d_S(x)$ and this entails by (1.90) that $x^* \in \gamma\partial_F(d_D \circ F)(x)$ and hence $x^* \in \gamma\partial_L(d_D \circ F)(x)$. The subdifferential chain rule (see, e.g., Mordukhovich [124, Corollary 3.43]) gives some $y^* \in \partial_L d_D(F(x))$ such that $\gamma^{-1}x^* = y^* \circ DF(x)$. Fix by (1.85) some open neighbourhood V of $F(\overline{x})$ and some constant $r > 0$ such that

$$(1.92)\quad \langle v^*, y' - y\rangle \le (2r)^{-1}\|y' - y\|^2 \quad \text{for all } y \in V \cap D, y' \in V \cap D \text{ and } v^* \in N_D^P(y) \cap B^*,$$

and such that the equality concerning the subdifferentials of the distance function in Theorem 1.4.6 holds with d_D at all points of $V \cap D$. Restricting the open convex neighbourhood U of $\overline{x}$ if necessary, we may suppose that $F(U) \subset V$ and that $\partial_L d_D(F(u)) = \partial_p d_D(F(u))$ for all $u \in U \cap S$ according to Theorem 1.4.6. Then (1.92) yields

$$(1.93)\qquad \langle y^*, y' - F(x)\rangle \le (2r)^{-1}\|y' - F(x)\|^2 \quad \text{for all } y' \in V \cap D.$$

Fix any $x' \in S \cap U$ and write

$$F(x') - F(x) = DF(x)(x' - x) + \int_0^1 \left(DF(x + t(x' - x)) - DF(x)\right)(x' - x)\, dt$$

and then

$$\langle \gamma^{-1}x^*, x' - x\rangle = \langle y^*, F(x') - F(x)\rangle - \int_0^1 \langle y^* \circ \left(DF(x + t(x' - x)) - DF(x)\right), x' - x\rangle\, dt.$$

Using (1.93) we obtain

$$\langle x^*, x' - x\rangle \le \gamma(2r)^{-1}\|F(x') - F(x)\|^2 + \gamma K_1\|x' - x\|^2 \le \gamma(K^2(2r)^{-1} + K_1)\|x' - x\|^2.$$

The latter being true for all $x, x' \in S \cap U$ and $x^* \in N_S^P(x) \cap B^*$, we conclude that the set S is N^P-hyporegular at $\overline{x}$. □

For the convenience of the reader we made the choice to prove first the N^P-hyporegularity of the intersection and then to give the additional arguments yielding to the N^P-hyporegularity of the inverse image. However the case of the intersection can be derived from that of the inverse image. Indeed putting $Y = X^m$, $F(x) = (x, \dots, x)$ for all $x \in X$, and $D = C_1 \times \cdots \times C_m$ we see that $F^{-1}(D) = \bigcap_{k=1}^{m} C_k$ and then some direct arguments and computation allow us to obtain through (b) the result of (a).

We state now the following corollary which is a direct consequence of Corollary 1.4.18.

Corollary 1.4.24. Assume that the moduli of convexity and smoothness of the norm of X are of power type and the power type of convexity is $q = 2$. Then under the assumptions of Proposition 1.4.23 the sets $\bigcap_{k=1}^{m} C_k$ and $F^{-1}(D)$ are prox-regular at $\overline{x}$.

When the spaces $(X, \|\cdot\|)$ and $(Y, \|\cdot\|)$ are Hilbert spaces, Proposition 1.4.23 translates, according to Corollary 1.4.18 and Proposition 1.4.19, the preservation of prox-regularity.

Corollary 1.4.25. Assume that X and Y are Hilbert spaces for their respective norms. Let $(C_k)_{k=1}^{m}$ be a finite family of closed sets of X and let D be a closed set of Y.
(a) If all sets C_k are prox-regular at a point $\overline{x}$ of all the sets C_k and if the intersection is metrically calm at $\overline{x}$, then this intersection set $\bigcap_{k=1}^{m} C_k$ is prox-regular at the point $\overline{x}$.
(b) If a mapping $F : X \to Y$ is of class $C^{1,1}$ around a point $\overline{x} \in F^{-1}(D)$ and metrically calm at $\overline{x}$ relatively to D and if D is prox-regular at $F(\overline{x})$, then the set $F^{-1}(D)$ is prox-regular at $\overline{x}$.

An analysis of the proof of Proposition 1.4.23 reveals that a uniform version also holds. We state it only in the case of an intersection and let the case of the inverse image to the reader.

Proposition 1.4.26. Assume that all the closed sets C_k, $k = 1, \dots, m$, are r-uniformly N^P-hyporegular (resp. r-uniformly prox-regular and $(X, \|\cdot\|)$ is a Hilbert space) and that their intersection is calm at anyone of its points with the same modulus γ of metric calmness in (1.89). Then the set $\bigcap_{k=1}^{m} C_k$ is r'-uniformly N^P-hyporegular (resp. r'-uniformly metrically prox-regular) with $r' := r/(m\gamma)$.

In the literature there are conditions with normal cones easy to handle ensuring the metric calmness inequalities (1.89) and (1.90). For example such a condition for (1.89) is known (see, e.g., Ioffe [90, p. 548-549]) under the name of general position condition for intersection of finitely many sets. In the context of our uniformly convex Banach space, the concept can be translated by saying that the closed sets $C_1, \dots, C_m$ are (*relative to the Fréchet normal cone*) in *sequential general position* at $\overline{x} \in \bigcap_{k=1}^{m} C_k$ whenever for any sequence of tuples $(x_{1,n}, \cdots, x_{m,n}, x^*_{1,n}, \cdots, x^*_{m,n})_n$ such that $x_{k,n} \in C_k$, $x_{k,n} \to \overline{x}$, $x^*_{k,n} \in N^F_{C_k}(x_{k,n}) \cap \mathbb{B}^*$ and such that $\left\| \sum_{k=1}^{m} x^*_{k,n} \right\| \to 0$, one has $\|x^*_{k,n}\| \to 0$ for any $k = 1, \dots, m$.

Another condition which is much easier to handle involves as above a property with normal cones but at the fixed point $\overline{x}$. It probably appears for the first time as one of the assumptions of Theorem 4.10 of Federer's seminal paper [73] for sets of $\mathbb{R}^n$ which are not submanifolds. The sets C_k, $k = 1, \dots, m$, are (*relative to the limiting* *normal cone*) in *pointbased general position* at $\bar{x}$ provided the equality $x_1^* + \cdots + x_m^* = 0$ with $x_k^* \in N_{C_k}^L(x_k)$ entails $x_1^* = \cdots = x_m^* = 0$.

Recall now that a closed set C of X is *compactly epi-Lipschitzian* at $\bar{x} \in C$ (a concept due to Borwein and Strójwas [30]) if there exist a compact set Q, a neighbourhood V of zero, a neighbourhood U of $\bar{x}$ in X, and a positive real number ε such that

$$C \cap U + tV \subset C + tQ \quad \text{for all } t \in]0, \varepsilon].$$

Of course any closed subset of X is compactly epi-Lipschitzian at any of its points whenever the space X is finite dimensional.

Corollary 1.4.27. Let $\{C_k\}_{k=1}^m$ be a finite family of closed sets of X and $\bar{x} \in \bigcap_{k=1}^m C_k$. Then the following hold:
(a) If the sets C_k are in sequential general position at $\bar{x}$ and if each set C_k is N^P-hyporegular at $\bar{x}$ (resp. prox-regular at $\bar{x}$ and $(X, \|\cdot\|)$ is a Hilbert space), then the intersection $\bigcap_{k=1}^m C_k$ is N^p-hyporegular (resp. prox-regular) at $\bar{x}$.
(b) If all the sets C_k except at most one are compactly epi-Lipschitzian at $\bar{x}$, then the sequential general position in (a) may be replaced by the corresponding pointbased general position at $\bar{x}$.
(c) If the space X is finite dimensional, again the sequential general position in (a) may be replaced by the corresponding pointbased general position at $\bar{x}$.

Proof. (a) Since the uniformly convex Banach space X is an Asplund space, the sequential general position is known to entail the metric calmness property (1.89) (and even more) according to Proposition 6 in p. 548 and Theorem 2 in p. 545 of Ioffe [89] (for example). Then the conclusion of (a) follows from Proposition 1.4.23 (resp. Corollary 1.4.25).
(b) The result comes from the fact that the assumption of compactly epi-Lipschitzian property combined with the pointbased general position at $\bar{x}$ ensures (see, e.g., Jourani and Thibault [98, Theorem 3.4]) that the intersection is metrically regular and hence metrically calm at $\bar{x}$.
(c) is a direct consequence of (b). □

The result in (c) was previously established in (5) of Theorem 4.10 in Federer [73].

1.4.6 Conical derivative of the mapping P_C

The following result is well-known for convex sets of Hilbert space (see Zarantonello [171, p. 300]). The term conical derivative has been coined in page 301 of Zarantonello [171]. The result has been independently extended by Canino [37] to closed p-convex sets of Hilbert space and by Shapiro [153] to closed sets with 2-order tangential property of Hilbert space too. It is actually known (see Poliquin, Rockafellar and Thibault [138]) that, in the context of Hilbert space, the concept of p-convex set is equivalent to uniform prox-regularity and the 2-order tangential property is equivalent to the local prox-regularity. The related result in Shapiro [153] may then be translated for prox-regular sets. Note also that the proof in Canino [37] still holds for sets which are prox-regular at the considered point of the Hilbert space. Our proposition below deals with the context of uniformly convex space.

Proposition 1.4.28. Assume that the moduli of uniform convexity and smoothness of the norm $\|\cdot\|$ of X are of power type. Assume also that the closed set C of X is prox-regular at $\bar{x} \in C$. Then there exists an open neighbourhood U of $\bar{x}$ such that P_C is single valued and continuous on U and such that for any $x \in U \cap C$ and $y \in X$ for which the mapping $t \mapsto \frac{1}{t}[P_C(x+ty) - x]$ is bounded on some interval $]0, t_0]$, one has

$$\lim_{t\downarrow 0} \frac{1}{t}[P_C(x + ty) - x] = P_{K_C(x)}(y)$$

and the directional derivative $d'_C(x; y) := \lim_{t\downarrow 0} \frac{1}{t}[d_C(x + ty) - d_C(x)]$ of d_C at x in the direction y exists and

$$d'_C(x; y) = d(y, K_C(x)),$$

which translates in the terminology of Zarantonello [171] that $d'_C(x; \cdot)$ is a conical derivative.

Proof. By the inequality (1.74) in the proof of Theorem 1.4.6, there exist $\varepsilon, r > 0$ with $\varepsilon < 1/2$ such that P_C is a single-valued continuous mapping on $B(\bar{x}, \varepsilon)$ and

$$\langle x^* - u^*, u - x\rangle \le \frac{2}{r}\|x - u\|\,\omega_{r+1}(\|x - u\|) \tag{1.94}$$

for all $x, u \in B(\bar{x}, \varepsilon)$ and all $x^* \in N^P_C(x)$, $u^* \in N^P_C(u)$ with $\|x^*\| \le 1$ and $\|u^*\| \le 1$, where ω_{r+1} is the modulus of uniform continuity of the duality mapping J over the bounded set $(r + 1)B$. Let $x \in B(\bar{x}, \varepsilon) \cap C$, $y \in X$, and $t_0 > 0$ such that the mapping $t \mapsto \frac{1}{t}[P_C(x+ty) - x]$ is bounded on $]0, t_0]$, say by $\beta > 0$. We may suppose that for all $t \in]0, t_0]$ we have $x + ty \in B(\bar{x}, \varepsilon)$ and $P_C(x + ty) \in B(\bar{x}, \varepsilon)$. For any $t \in]0, t_0]$ putting $h'_t := \frac{1}{t}[P_C(x + ty) - x]$, we see that

$$y - h'_t = \frac{1}{t}[x + ty - P_C(x + ty)]$$

and hence $y - h'_t \in PN_C(P_C(x + ty))$, i.e., $J(y - h'_t) \in N^P_C(P_C(x + ty))$.

Now take any sequence $(t_n)_n$ of $]0, t_0]$ converging to 0. The sequence $(J(y - h'_{t_n}))_n$ being bounded, there exists an increasing function $\sigma : \mathbb{N} \to \mathbb{N}$ such that the sequence $(J(y - h'_{t_{\sigma(n)}}))_n$ converges weakly star to $J(y - h)$ for some vector $h \in X$. As $N^P_C(x) = N^L_C(x)$ (see Theorem 1.4.6), we have $J(y-h) \in N^P_C(x)$. Put $h_n := h'_{t_{\sigma(n)}}$. By (1.94) there exists some constant $\gamma > 0$ independent of n such that for $\lambda_n := t_{\sigma(n)}$

$$\langle J(y - h_n) - J(y - h), x - P_C(x + \lambda_n y)\rangle \leq \frac{\gamma}{r}\|P_C(x + \lambda_n y) - x\|\, \omega_{r+1}(\|P_C(x + \lambda_n y) - x\|).$$

The latter inequality is equivalent to

$$\langle J(y - h_n) - J(y - h), -\lambda_n h_n\rangle \leq \frac{\gamma}{r}\|P_C(x + \lambda_n y) - x\|\, \omega_{r+1}(\|P_C(x + \lambda_n y) - x\|),$$

that is,

$$\langle J(y - h_n) - J(y - h), (y - h_n) - (y - h)\rangle - \langle J(y - h_n) - J(y - h), h\rangle \leq$$

$$\frac{\gamma}{r}\left\|\frac{1}{\lambda_n}[P_C(x + \lambda_n y) - x]\right\|\, \omega_{r+1}(\|P_C(x + \lambda_n y) - x\|).$$

If $\max\{\|y - h_n\|, \|y - h\|\} = 0$ for infinitely many n, then for all these integers n we have $y = h_n = h$ and hence by definition of h_n we obtain $h = \frac{1}{\lambda_n}[P_C(x + \lambda_n y) - x]$, which yields $h \in K_C(x)$. Suppose now that $\max\{\|y - h_n\|, \|y - h\|\} > 0$ for all n not less than some n_0. According to (1.48) for some constant $K_2 > 0$ we have for all $n \geq n_0$

$$K_2(\max\{\|y - h_n\|, \|y - h\|\})^2 \delta_{\|\cdot\|}\left(\frac{\|h_n - h\|}{2(\max\{\|y - h_n\|, \|y - h\|\})}\right) \leq$$

$$\langle J(y - h_n) - J(y - h), h\rangle + \frac{\gamma}{r}\left\|\frac{1}{\lambda_n}[P_C(x + \lambda_n y) - x]\right\|\, \omega_{r+1}(\|P_C(x + \lambda_n y) - x\|),$$

and since the first expression of the second member tends to 0 and the modulus of convexity $\delta_{\|\cdot\|}$ is an increasing function with $\delta_{\|\cdot\|}(t) \underset{t\downarrow 0}{\longrightarrow} 0$, it is not difficult to see that $\|h_n - h\| \to 0$ and by definition of h_n we obtain $h \in K_C(x)$. So in any case, we have $h \in K_C(x)$ and the sequence $(h_n)_n$ strongly converges to h.

Now take any $v \in K_C(x)$. Since $K_C(x) = T_C(x)$ according to Theorem 1.4.6, there is a sequence $(z_n)_n$ converging to v such that $x + \lambda_n z_n \in C$ for all n. This entails that for all n

$$\|x + \lambda_n y - (x + \lambda_n z_n)\| \geq \|x + \lambda_n y - P_C(x + \lambda_n y)\|,$$

which is equivalent to

$$\|y - z_n\| \geq \|y + \frac{1}{\lambda_n}[x - P_C(x + \lambda_n y)]\|.$$

Then passing to the limit with $n \to \infty$, we obtain $\|y - v\| \geq \|y - h\|$ for all $v \in K_C(x)$, which means that $h = P_{K_C(x)}(y)$. So for any sequence of positive numbers $(t_n)_n$ converging to 0 there exists a subsequence $(t_{\sigma(n)})_n$ such that

$$\lim_{n\to\infty} \frac{1}{t_{\sigma(n)}}[P_C(x + t_{\sigma(n)}y) - x] = P_{K_C(x)}(y).$$

Consequently

$$\lim_{t\downarrow 0} \frac{1}{t}[P_C(x + ty) - x] = P_{K_C(x)}(y). \tag{1.95}$$

Concerning the equality of the directional derivative of the distance function to the set C, observe first that for any positive number t small enough

$$\begin{aligned}\frac{1}{t}[d_C(x + ty) - d_C(x)] = \frac{1}{t}d_C(x + ty) &= \tfrac{1}{t}\|P_C(x + ty) - (x + ty)\| \\ &= \left\|\tfrac{1}{t}[P_C(x + ty) - x] - y\right\|.\end{aligned}$$

Then taking (1.95) into account and passing to the limit when $t \downarrow 0$ give that the directional derivative $d'_C(x; y)$ exists and

$$d'_C(x; y) = \|P_{K_C(x)}(y) - y\| = d(y, K_C(x)). \quad \square$$

In the case of a Hilbert space X the mapping P_C is Lipschitz near $\bar{x}$ (see Poliquin, Rockafellar and Thibault [138]) and hence the boundedness assumption of the proposition above is fulfilled and one recovers in Corollary 1.4.29 below the result of Proposition 2.12 in Canino [37] and Theorem 3.1 in Shapiro [153]. The result is proved in Shapiro [153] through some properties of solutions of perturbed optimization problems satisfying some appropriate conditions. The proof in Canino [37] is based on the duality in Hilbert space between the tangent cone $K_C(x)$ and the primal normal cone $PN_C(x)$ for prox-regular sets C. Such a duality property is not available in the non Hilbert setting because of the nonlinearity of the duality mapping J. Instead, our proof of Proposition 1.4.28 above is related to the tangential regularity of C established in Theorem 1.4.6 and to the sequential characterization of the Clarke tangent cone (recalled in the beginning of Subsection 1.4.1).

Corollary 1.4.29 (Canino [37, Proposition 2.12] and Shapiro [153, Theorem 3.1]). Assume that X is a Hilbert space and that C is prox-regular at $\bar{x} \in C$. Then for some neighbourhood U of $\bar{x}$ one has that for all $x \in U \cap C$ and $y \in X$ the directional derivatives of P_C and d_C at x in the direction y exist and

$$\lim_{t\downarrow 0} \frac{1}{t}[P_C(x + ty) - x] = P_{K_C(x)}(y) \text{ and } d'_C(x; y) = d(y, K_C(x)).$$

1.4.7 Convergence

In this subsection we are interested in the behaviour of the projection mapping under convergence of prox-regular sets.

Let us begin by recalling some properties of the projection mapping of a prox-regular set and by providing some additional facts. Consider a uniformly r-prox-regular closed set C. We see from the proof of Theorem 1.3.27 that the function d_C^2 is Fréchet differentiable on the r-open neighborhood $O_C(r) := \{x \in X : d_C(x) < r\}$ of C and that P_C is a single-valued and norm-to-norm continuous mapping from $O_C(r)$ into C. Moreover, from the proof of Proposition 1.3.30, for any $x \in O_C(r)$

$$\nabla^F \left(\frac{1}{2} d_C^2\right)(x) = J(x - P_C(x)). \tag{1.96}$$

Concerning the continuity of P_C, the proof of Theorem 1.3.27 says more, in the sense that P_C is locally Hölder continuous on $O_C(r)$. The following two propositions furnish some further points that we will need in Theorem 1.4.33 below. Recall that q (resp. s) denotes the power type of the modulus of convexity (resp. smoothness) of the norm of X.

Proposition 1.4.30. Assume that the moduli of uniform convexity and smoothness of the norm $\|\cdot\|$ of X are of power type q and s respectively. Let r, r' be two positive numbers with $r' < \frac{r}{2}$ and let a number $\rho > 0$. Then there exists some constant $\gamma \geq 0$ depending only on r, r' and ρ such that, for any uniformly r-prox-regular closed subset C of X, one has

$$\|P_C(x_1) - P_C(x_2)\| \leq \gamma \|x_1 - x_2\|^{\frac{1}{q}} \quad \text{and} \quad \|\nabla^F d_C^2(x_1) - \nabla^F d_C^2(x_2)\| \leq \gamma \|x_1 - x_2\|^{\frac{s-1}{q}} \tag{1.97}$$

for all $x_1, x_2 \in O_C(r') \cap \rho B$.

Proof. Take $\alpha \in]0, 1/2[$ such that $r' = \alpha r$. By Step 1 of the proof of Theorem 1.3.27 the truncated normal cone set-valued mapping $N_C^{P\,r}(\cdot)$ is J-hypomonotone of degree 1 and hence so is the set-valued mapping $N_C^{P\,\alpha r}(\cdot)$. Referring to the proof of Lemma 1.3.8 with $T(\cdot) := N_C^{P\,\alpha r}(\cdot)$ and $\overline{r} := 1$, since $\alpha^{-1} > 2\overline{r}$ the development there reveals that, fixing λ satisfying $2\alpha < \lambda < 1$, there exists some constant $\gamma' > 0$ depending only on α, ρ, and r such that $Q := (I + \alpha J^* \circ N_C^{P\,r})^{-1}$ is on $\rho B \cap \operatorname{Dom} Q$ a single valued mapping for which

$$\|Q(x_1) - Q(x_2)\| \leq \gamma' \|x_1 - x_2\|^{\frac{1}{q}} \quad \text{for all } x_1, x_2 \in \rho B \cap \operatorname{Dom} Q. \tag{1.98}$$

Observe now that for $x \in O_C(\alpha r)$ we have $J(x - P_C(x)) \in N_C^P(P_C(x))$ (by definition of proximal normal functional) and $\|J(x - P_C(x))\| < \alpha r$ and that those two facts imply $J(x - P_C(x)) \in \alpha N_C^{P\,r}(P_C(x))$, that is, $P_C(x) = (I + \alpha J^* \circ N_C^{P\,r})^{-1}(x)$. We may then reformulate (1.98) in the form

$$\|P_C(x_1) - P_C(x_2)\| \le \gamma' \|x_1 - x_2\|^{\frac{1}{q}} \quad \text{for all } x_1, x_2 \in O_C(r') \cap \rho B.$$

Concerning the square distance function d_C^2, we know that it is (see 1.96) Fréchet differentiable on $O_C(r)$ with $\nabla^F d_C^2(x) = 2J(x - P_C(x))$ for all $x \in O_C(r)$. So using (1.55) and the latter inequality above, we have for all $x_1, x_2 \in O_C(r') \cap \rho B$

$$\begin{aligned} \|\nabla^F d_C^2(x_1) - \nabla^F d_C^2(x_2)\| &\le 2L_\rho \|x_1 - P_C(x_1) - x_2 + P_C(x_2)\|^{s-1} \\ &\le 2L_\rho (\|x_1 - x_2\| + \gamma' \|x_1 - x_2\|^{\frac{1}{q}})^{s-1} \\ &\le 2L_\rho (\gamma' + \|x_1 - x_2\|^{1-\frac{1}{q}})^{s-1} \|x_1 - x_2\|^{\frac{s-1}{q}} \\ &\le 2L_\rho (\gamma' + (2\rho)^{1-\frac{1}{q}})^{s-1} \|x_1 - x_2\|^{\frac{s-1}{q}}. \end{aligned}$$

It suffices to take $\gamma = \max\left\{\gamma', 2L_\rho(\gamma' + (2\rho)^{1-\frac{1}{q}})^{s-1}\right\}$ to obtain both inequalities in the statement of the proposition. □

In the proposition above, under the restriction $r' < r/2$, the power of Hölder continuity of P_C over $O_C(r') \cap \rho B$ is the constant $1/q$ (the inverse of the power type of uniform convexity of $\|\cdot\|$), i.e., only the modulus γ of Hölder continuity varies with r' and ρ. Relaxing the restriction $r' < r/2$ into $r' < r$ and letting the power of Hölder continuity depending also on r' and ρ, we can prove the following result of Hölder continuity of P_C on $O_C(r') \cap \rho B$ but this time with any $r' < r$.

Before stating the result, recall that for any $r > 0$ the (closed) r-enlargement $C(r)$ of C is given by $C(r) := \{x \in X : d(x, C) \le r\}$.

Proposition 1.4.31. Assume that the moduli of uniform convexity and smoothness of the norm $\|\cdot\|$ of X are of power type q and s respectively. Let ρ, r, r' be positive real numbers with $r' < r$. Then there exist some positive constants γ and $\theta \le 1$ both depending only on r', r and ρ, such that for any uniformly r-prox-regular closed set C of X one has

$$(1.99)\quad \|P_C(x_1) - P_C(x_2)\| \le \gamma \|x_1 - x_2\|^\theta \text{ and } \|\nabla^F d_C^2(x_1) - \nabla^F d_C^2(x_2)\| \le \gamma \|x_1 - x_2\|^{\theta(s-1)}$$

for all $x_1, x_2 \in O_C(r') \cap \rho B$.

More precisely, defining $(\alpha_n)_n$ by $\alpha_1 = \frac{1}{2}$ and $\alpha_{n+1} = \frac{1+\alpha_n}{2}$, then for any $n \in \mathbb{N}$, $\rho > 0$, $\alpha' \in]0, \alpha_n[$, there exists some positive real number k (depending only on ρ, α' and n) such that, for any uniformly r-prox-regular closed set C of X, the metric projection mapping P_C is Hölder continuous on $O_C(\alpha' r) \cap \rho B$ with power $\theta := \frac{1}{q^n}$ and with modulus k.

Proof. Let us call $\mathcal{P}(n)$ the property above in the second part of the proposition for P_C of any uniformly r-prox-regular closed set C. Note that $\mathcal{P}(1)$ is Proposition 1.4.30. Suppose that $\mathcal{P}(n)$ is fulfilled, and fix any uniformly r-prox regular closed set C and any $\rho > 0, \alpha' \in]0, \alpha_n[$. We may suppose that $O_C(r) \cap \rho B \neq \emptyset$. By the proofs of Theorem 1.3.27 and Lemma 1.3.29, the closed set $C(\alpha' r)$ is uniformly $(1-\alpha')r$-prox-regular. Then applying now Proposition 1.4.30 to the set $C(\alpha' r)$ yields that for any $\lambda \in]0, 1[$,

$$P_{C(\alpha' r)} \text{ is Hölder continuous on } O_{C(\alpha' r)}\left(\lambda \frac{1-\alpha'}{2} r\right) \cap \rho B \text{ with power } 1/q \tag{1.100}$$

and with modulus depending only on ρ, λ, α', and r. On the other hand, we have that for any $t > 0$, $O_C(\alpha' r + t) = O_{C(\alpha' r)}(t)$. Indeed, from Lemma 1.3.28 (a) in (see also Bounkhel and Thibault [33]), for any $\tau > 0$ and any $u \notin C(\tau)$ one has $d_C(u) = \tau + d_{C(\tau)}(u)$. Hence, for any $u \in X$, one has the equivalence $d_C(u) < \alpha' r + t \Leftrightarrow d_{C(\alpha' r)}(u) < t$, that is, the desired equality holds. This entails that

$$O_{C(\alpha' r)}\left(\lambda \frac{1-\alpha'}{2} r\right) = O_C\left(\alpha' r + \lambda \frac{1-\alpha'}{2} r\right). \tag{1.101}$$

By (1.100) and (1.101) there exists some constant $K_1 > 0$ (depending only on ρ, λ, α', and r) such that

$$\|P_{C(\alpha' r)}(u_1) - P_{C(\alpha' r)}(u_2)\| \leq K_1 \|u_1 - u_2\|^{1/q} \text{ for all } u_i \in O_C\left(\alpha' r + \lambda \frac{1-\alpha'}{2} r\right). \tag{1.102}$$

Fixing $b_\rho \in O_C(r) \cap \rho B \neq \emptyset$, we find some $a_\rho \in C$ with $\|a_\rho\| \leq r + \rho$. Combining this with (1.102), we obtain that there is some $\rho' > 0$ (depending only on ρ, λ, α', and r) such that

$$\|P_{C(\alpha' r)}(u)\| \leq \rho' \quad \text{for all } u \in (\rho B) \cap O_C\left(\alpha' r + \lambda \frac{1-\alpha'}{2} r\right) \tag{1.103}$$

(take $u_1 = u$ and $u_2 = a_\rho$ in (1.102)). The equality (1.101) also implies, according to Theorem 1.3.27 that for any $u \in O_C\left(\alpha' r + \lambda \frac{1-\alpha'}{2} r\right) \setminus C(\alpha' r)$, the points $y := P_{C(\alpha' r)}(u)$ and $z := P_C(y)$ exist. By (c) of Lemma 1.3.28, for any such u we have consequently $z = P_C(u)$ and so, $P_C(u) = P_C \circ P_{C(\alpha' r)}(u)$. From this, if $u_1, u_2 \in (\rho B) \cap \left(O_C\left(\alpha' r + \lambda \frac{1-\alpha'}{2} r\right) \setminus C(\alpha' r)\right)$, then for any $\alpha'' \in]\alpha', \alpha_n[$ we have

$$\begin{aligned} \|P_C(u_1) - P_C(u_2)\| &= \|P_C(P_{C(\alpha' r)}(u_1)) - P_C(P_{C(\alpha' r)}(u_2))\| \\ &\leq K_2 \|P_{C(\alpha' r)}(u_1) - P_{C(\alpha' r)}(u_2)\|^{1/q^n} \end{aligned}$$

for some constant $K_2 > 0$ (depending only on ρ', α'', and n), according to $\mathcal{P}(n)$ and to the fact that, by (1.103), for $i = 1, 2$, $P_{C(\alpha' r)}(u_i) \in \rho' B \cap O_C(\alpha'' r)$. Taking (1.102) into account we obtain that there exists some constant K_3 (depending only on ρ, λ, α', r and n) such that for all $u_1, u_2 \in (\rho B) \cap \left(O_C\left(\alpha' r + \lambda \frac{1-\alpha'}{2} r\right) \setminus C(\alpha' r)\right)$

$$\|P_C(u_1) - P_C(u_2)\| \le K_3 \|u_1 - u_2\|^{1/q^{n+1}} \tag{1.104}$$

and this inequality is still true for all $u_i \in \rho B$ satisfying $\alpha' r \le d(u_i, C) < \alpha' r + \lambda \frac{1-\alpha'}{2} r$. On $\rho B \cap C(\alpha' r)$, $\mathcal{P}(n)$ ensures that P_C also has the Hölder continuity property with power $1/q^n$ and hence also with power $1/q^{n+1}$ (because of the boundedness of $\rho B \cap C(\alpha' r)$). Then there exists some positive constant K_4 (depending only on ρ, λ, α', r and n) such that

$$\|P_C(u_1) - P_C(u_2)\| \le (K_4/2) \|u_1 - u_2\|^{1/q^{n+1}}$$

for all $u_1, u_2 \in (\rho B) \cap O_C\left(\alpha' r + \lambda \frac{1-\alpha'}{2} r\right)$ satisfying either $d(u_i, C) \ge \alpha' r$ for $i = 1$ and $i = 2$ or $d(u_i, C) \le \alpha' r$ for $i = 1$ and $i = 2$. In the remaining case where $u_1, u_2 \in (\rho B) \cap O_C\left(\alpha' r + \lambda \frac{1-\alpha'}{2} r\right)$ with $d(u_1, C) < \alpha' r$ and $d(u_2, C) > \alpha' r$, the continuity of $d(\cdot, C)$ over the line segment $[u_1, u_2]$ yields some point $u_0 \in]u_1, u_2[$ such that $d(u_0, C) = \alpha' r$; then

$$\|P_C(u_1) - P_C(u_2)\| \le (K_4/2) \left(\|u_1 - u_0\|^{1/q^{n+1}} + \|u_0 - u_2\|^{1/q^{n+1}}\right) \le K_4 \|u_1 - u_2\|^{1/q^{n+1}},$$

the second inequality being due to the inequalities $\|u_i - u_0\| \le \|u_1 - u_2\|$. The mapping P_C is then Hölder continuous on $\rho B \cap O_C\left(\alpha' r + \lambda \frac{1-\alpha' r}{2} r\right)$ with modulus K_4 and power $1/q^{n+1}$.

We have shown that, for any $\lambda \in]0, 1[$, $\alpha' \in]0, \alpha_n[$, and $\rho > 0$, there exists some positive real number k (depending only on ρ, α', λ, and n) such that, for any uniformly r-prox-regular set C, the metric projection mapping P_C is Hölder continuous on $\rho B \cap O_C\left(\alpha' r + \lambda \frac{1-\alpha'}{2} r\right)$ with power $1/q^{n+1}$ and with modulus k. This means that for any $\alpha' \in]0, \alpha_n[$ there is some positive k (depending only on ρ, α', and n) such that, for any uniformly r-prox-regular set C of X, the mapping P_C is Hölder continuous on $\rho B \cap O_C\left(\alpha' r + \frac{1-\alpha'}{2} r\right)$ with power $1/q^{n+1}$ and modulus k. Consequently $\mathcal{P}(n+1)$ holds true since $\alpha' r + \frac{1-\alpha'}{2} r = \frac{1+\alpha'}{2} r$.

The Hölder continuity of the function $\nabla^F d_C^2(\cdot)$ is obtained like in the proof of Proposition 1.4.30. □

Remark 1.4.32. When $(X, \|\cdot\|)$ is a Hilbert space, where $\|\cdot\|$ is the norm associated with the inner product, the behaviour of $P_C(\cdot)$ and $\nabla^F d_C^2(\cdot)$ is distinctly better. Indeed, by Poliquin, Rockafellar and Thibault [138] for any real positive numbers r, r' with $r' < r$ and any uniformly r-prox-regular closed set C of X one has

$$\|P_C(x_1) - P_C(x_2)\| \le \frac{r}{r - r'} \|x_1 - x_2\|$$

for all $x_1, x_2 \in O_C(r')$, and hence

$$\|\nabla^F d_C^2(x_1) - \nabla^F d_C^2(x_2)\| \le 2\left(1 + \frac{r}{r - r'}\right) \|x_1 - x_2\|$$

according to (1.96) and to the Lipschitz property of J here with Lipschitz modulus 1. So, not only the Lipschitz continuity (instead of the Hölder one) is available for the metric projection but also no restriction to bounded subsets of $O_C(r')$ is required.

We can now study convergence properties of families of uniformly prox-regular sets. Let $T \cup \{t_0\}$ be a topological space with t_0 as a cluster point of T. Recall that a family $(C_t)_{t\in T}$ of non empty closed subsets of X converges in the sense of Attouch-Wets (see Attouch and Wets [2], Beer [13], Rockafellar and Wets [152]) to a closed subset C of X when t goes to t_0 provided for each positive real number ρ one has

$$\sup_{x\in\rho B} |d_{C_t}(x) - d_C(x)| \underset{t\to t_0}{\longrightarrow} 0. \tag{1.105}$$

Theorem 1.4.33. Assume that the moduli of uniform convexity and smoothness of the norm $\|\cdot\|$ of X are of power type. Let $(C_t)_{t\in T}$ be a family of closed uniformly r-prox-regular sets of X which converges in the sense of Attouch-Wets to a closed set C of X. Then C is uniformly r-prox-regular and for each $x_0 \in X$ satisfying $d(x_0, C) < r$ the mapping $t \mapsto P_{C_t}(x_0)$ and the function $t \mapsto \nabla^F d^2_{C_t}(x_0)$ are defined on a neighbourhood of t_0 and

$$P_{C_t}(x_0) \underset{t\to t_0}{\longrightarrow} P_C(x_0) \quad \text{and} \quad \nabla^F d^2_{C_t}(x_0) \underset{t\to t_0}{\longrightarrow} \nabla^F d^2_C(x_0). \tag{1.106}$$

Proof. Fix a positive number r' with $d(x_0, C) < r' < r$ and choose by (1.105) some neighbourhood T_0' of t_0 and some $\beta > 0$ such that $d(x, C_t) < r'$ for all $t \in T_0' \setminus \{t_0\}$ and $x \in x_0 + 2\beta B$. According to Proposition 1.4.31, there exist some positive constants γ and $\sigma \le 1$ (both depending only x_0, β, r' and r) such that for all $t \in T_0'$ with $t \neq t_0$ and $x_1, x_2 \in x_0 + 2\rho B$ the derivatives $\nabla^F d^2_{C_t}(x_i)$, $i = 1, 2$, exist and

$$\|\nabla^F d^2_{C_t}(x_1) - \nabla^F d^2_{C_t}(x_2)\| \le \gamma \|x_1 - x_2\|^\sigma. \tag{1.107}$$

Fix another neighbourhood T_0'' of t_0 with $T_0'' \subset T_0'$ and such that, according to (1.105), the function v given for $\rho := 1 + 2\beta + \|x_0\|$ by

$$v(t) := \sup_{x\in\rho B} |d_{C_t}(x) - d_C(x)|$$

is bounded on $T_0 := T_0'' \setminus \{t_0\}$. Writing

$$d_{C_t}(x) \le d_{C_t}(x_0) + \|x - x_0\| \le d_C(x_0) + \|x - x_0\| + v(t) \quad \text{for all } x \in X \text{ and } t \in T,$$

we see that there exists some constant $M > 0$ such that

$$d_{C_t}(x) \le M/2 \quad \text{for all } (t, x) \in T_0 \times \rho B. \tag{1.108}$$

Fix now any positive number $\varepsilon < \beta$. Choose by (1.105) a neighbourhood $T'_\varepsilon \subset T''_0$ of t_0 such that for $T_\varepsilon := T'_\varepsilon \setminus \{t_0\}$

$$|d_{C_t}(x) - d_C(x)| \le \frac{\varepsilon^{\sigma+1}}{2} \quad \text{for all } t \in T_\varepsilon \text{ and } x \in \rho B. \tag{1.109}$$

Fix any $x \in B(x_0, \beta)$, any $w \in X$ with $\|w\| = \varepsilon$, and any $t, \tau \in T_\varepsilon$. We have

$$d^2_{C_t}(x+w) - d^2_C(x) = \langle \nabla^F d^2_{C_t}(x), w\rangle + \int_0^1 \langle \nabla^F d^2_{C_t}(x+sw) - \nabla^F d^2_{C_t}(x), w\rangle\, ds$$

and a similar equality with $d^2_{C_\tau}$ in place of $d^2_{C_t}$, hence

$$\begin{aligned} &\langle \nabla^F d^2_{C_t}(x) - \nabla^F d^2_{C_\tau}(x), w\rangle \\ = \;& [d^2_{C_t}(x+w) - d^2_{C_\tau}(x+w)] - [d^2_{C_t}(x) - d^2_{C_\tau}(x)] \\ & - \int_0^1 \langle \nabla^F d^2_{C_t}(x+sw) - \nabla^F d^2_{C_t}(x), w\rangle + \int_0^1 \langle \nabla^F d^2_{C_\tau}(x+sw) - \nabla^F d^2_{C_\tau}(x), w\rangle. \end{aligned}$$

Further, on the one hand by (1.107) we have

$$\left| \int_0^1 \langle \nabla^F d^2_{C_t}(x+sw) - \nabla^F d^2_{C_t}(x), w\rangle\, ds \right| \le$$

$$\int_0^1 \gamma s^\sigma \|w\|^{\sigma+1}\, ds = \frac{1}{\sigma+1}\gamma \|w\|^{\sigma+1} \le \gamma \varepsilon^{\sigma+1}$$

and a similar inequality with $d^2_{C_\tau}$ in place of $d^2_{C_t}$. On the other hand by (1.108) and (1.109), with $y := x + w$ or $y := x$ we also have

$$|d^2_{C_t}(y) - d^2_{C_\tau}(y)| = (d_{C_t}(y) + d_{C_\tau}(y))|d_{C_t}(y) - d_{C_\tau}(y)| \le M\varepsilon^{\sigma+1}.$$

Consequently

$$\langle \nabla^F d^2_{C_t}(x) - \nabla^F d^2_{C_\tau}(x), w\rangle \le (2M + 2\gamma)\varepsilon^{\sigma+1}$$

and hence that for every u in X with $\|u\| = 1$

$$\langle \nabla^F d^2_{C_t}(x) - \nabla^F d^2_{C_\tau}(x), u\rangle \le (2M + 2\gamma)\varepsilon^{\sigma},$$

which ensures that

$$\|\nabla^F d^2_{C_t}(x) - \nabla^F d^2_{C_\tau}(x)\| \le (2M + 2\gamma)\varepsilon^{\sigma}.$$

This uniform Cauchy property says that the family $(\nabla^F d^2_{C_t})_t$ converges uniformly to some mapping from $B(x_0, \beta)$ into X^*. This is known to entail that the function d^2_C is

Fréchet differentiable on $B(x_0,\beta)$ and hence, in particular, Fréchet differentiable at x_0 with

$$\nabla^F d^2_{C_t}(x_0) \underset{t\to t_0}{\longrightarrow} \nabla^F d^2_C(x_0). \tag{1.110}$$

Then the function d_C is Fréchet differentiable on the open tube $U_C(r)$ and this differentiability property translates through Theorem 1.3.27 the uniform r-prox-regularity of C. Further (1.96) and (1.110) yield

$$P_{C_t}(x_0) \underset{t\to t_0}{\longrightarrow} P_C(x_0).$$

This completes the proof of the proposition. □

Chapter 2

Integrability of subdifferentials of functions

The purpose of studying integration of subdifferentials is to answer the question whether or not the condition that the subdifferential of g contains the subdifferential of f implies that f and g differ by a constant. It is accepted instead of integrability of subdifferentials of functions to speak of the integrability of functions, in both cases being understood one and the same.

The famous Moreau-Rockafellar's integration result (see Rockafellar [147]) gives a positive answer when f and g are lower semicontinuous convex functions defined on a Banach space. The answer is also affirmative for some classes of locally Lipschitz functions (see Borwein and Moors [27], Borwein, Moors and Wang [28], Correa and Jofre [50], Correa and Thibault [54]). The first significant extension outside the convex and locally Lipschitz case is due to Poliquin [135] who shows that the integration result holds in a finite dimensional setting for functions f and g that are primal lower nice. Later, Thibault and Zagrodny [160] extend the result of Poliquin to the class of convexly subdifferentially similar functions defined on a Banach space. This class includes primal lower nice functions defined on a Hilbert space, as well as the differences of convex functions. Ivanov and Zlateva [93] established the integration result for the class of semi-convex functions. New insight on this topic can be found in the work of Thibault and Zagrodny [161], where a more general inclusion of subdifferentials is investigated.

In Section 2.1 the famous Moreau-Rockafellar result is proved by using regularization (and approximation) techniques which was the initial idea of Moreau. The result is published by Zlateva in [172].

Using a quantitative version of subdifferential characterization of a class of directionally Lipschitz functions in Section 2.2 we establish the integrability of subdifferentials of such functions. The result is published by Thibault and Zlateva in [163].

In Section 2.3 we study integrability properties of bivariate functions defined on a product of Banach spaces. Progress in this direction was already made by Correa and Thibault [54], and by Wu and Ye [168]. The problem in this setting is interesting from the point of view that the product structure allows us to introduce different concepts of continuity and regularity of the function and study how they rely to the integrability. Special attention here will be given to directionally Lipschitz bivariate functions in order to develop the results for directionally Lipschitz functions of one variable from Section 2.2. The results are published by Thibault and Zlateva in [162].

In Section 2.4 we study on a product Banach space the properties of a class of saddle functions called partially ball weakly inf-compact. For such a function we prove that the domain of the subdifferential is non-empty, that the operator naturally associated with the subdifferential is maximal monotone, and that the subdifferential of the function is integrable. For a function in a large subclass of that class we prove the density of the domain of the subdifferential in the domain of the function. The results are published by Thibault and Zlateva in [164].

2.1 Integrability of the subdifferential of a convex function through infimal regularization

As usual, X is a Banach space with dual X^*. A function $f : X \to \mathbb{R} \cup \{+\infty\}$ is *convex*, whenever

$$f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda) f(y), \quad \forall x, y \in X, \text{ and } \forall \lambda \in [0,1],$$

or, which is equivalent, $\operatorname{epi} f$ is a convex set in $X \times \mathbb{R}$.

Subgradient of a convex function f at point x is any $x^* \in X^*$ such that

$$f(x') \ge f(x) + \langle x' - x, x^* \rangle, \quad \forall x' \in X.$$

The set of all subgradients of f at x (which could be empty) is denoted by

$$\partial f(x) = \{p \in X^*;\ \langle p, y - x\rangle \le f(y) - f(x),\ \forall y \in X\},$$

and the multivalued map $\partial f : X \to 2^{X^*}$ is called *convex subdifferential* (or *Fenchel subdifferential*) of f whose domain $\operatorname{dom} \partial f \subset \operatorname{dom} f$ if f is proper.

For a proper, convex and lower semicontinuous function $f : X \to \mathbb{R} \cup \{+\infty\}$ the subdifferential of f at $x \in \operatorname{dom} f$ *in the sense of convex analysis* (or *Fenchel subdifferential*) is denoted by $\partial f(x)$, that is

$$\partial f(x) = \{p \in X^*;\ \langle p, y - x\rangle \le f(y) - f(x),\ \forall y \in X\},$$

and $\partial f = \emptyset$ on $X \setminus \operatorname{dom} f$.

The famous integration theorem of Moreau and Rockafellar, see Moreau [128] and Rockafellar [144, 147] states that each proper, convex and lower semicontinuous function is determined up to an additive constant by its subdifferential.

This result was firstly stated and proved by Moreau in [128] on a Hilbert space H. The proof uses infimal regularizations and works also in reflexive Banach space as mentioned of Moreau at [129, p.87].

Below we present a brief sketch of Moreau proof (see Moreau [128]). When H is a Hilbert space, for all $z \in H$ the function

$$u \to f(u) + \frac{1}{2}\|u - z\|^2$$

possesses a strict minimum called $prox_f z$ and z can be represented as the sum $z = x+y$ of x such that $x = prox_f z$ and y such that $y \in \partial f(x)$. Moreover, the decomposition of z as a sum of x and y such that $y \in \partial f(x)$ is unique and holds only for $x = prox_f z$ (see Moreau [128, Proposition 4a]).

Consider the function $\varphi : H \to \mathbb{R}$

$$\varphi(z) = \inf_{u \in H}\left\{f(u) + \frac{1}{2}\|u - z\|^2\right\},$$

which is convex and Fréchet differentiable with Fréchet derivative $\varphi'(z) = prox_f z - z$.

Similarly, if $g : X \to \mathbb{R} \cup \{+\infty\}$ is a proper, convex and lower semicontinuous function, we consider

$$\psi(z) = \inf_{u \in H}\left\{g(u) + \frac{1}{2}\|u - z\|^2\right\},$$

which is convex and Fréchet differentiable with Fréchet derivative $\psi'(z) = prox_g z - z$ and for all $z \in H$ the unique decomposition of the form $z = x + y$ where $y \in \partial g(x)$ holds for $x = prox_g z$.

If the inclusion $\partial f(x) \subset \partial g(x)$ holds for all $x \in H$, then $prox_g z = prox_f z$ for all $z \in H$, which implies that $\varphi'(z) = \psi'(z)$ for all $z \in H$. This yields that for some constant c, $\varphi(z) = \psi(z) + c$ for all $z \in H$, which for their conjugate functions yields $\varphi^*(z) = \psi(z)^* - c$ for all $z \in H$. As it is well known that $\varphi^* = f^* + \frac{1}{2}\|\cdot\|^2$, where f^*

is the conjugate of f and that $\psi^* = g^* + \frac{1}{2}\|\cdot\|^2$, where g^* is the conjugate of g. So it holds that $f^* = g^* - c$, and then $f = g + c$ (see Moreau [128, Proposition 8a]).

It is clear that the idea of Moreau was to replace f and g with more regular functions φ and ψ generated by them. Rockafellar's proof in Banach space is completely different and uses duality arguments. Our purpose here is to prove the Moreau-Rockafellar result in Banach space by using regularization (and approximation) techniques which was the initial idea of Moreau.

Given a proper, convex and lower semicontinuous function $f : X \to \mathbb{R} \cup \{+\infty\}$ one defines the multi-valued operator $\partial f : X \rightrightarrows X^*$ which to any $x \in X$ assigns the (probably empty) set $\partial f(x)$ of the convex subdifferential of f at x.

Recall that an operator $T : X \rightrightarrows X^*$ is *monotone* if for all $x^* \in T(x)$ and $y^* \in T(y)$ it holds $\langle x^* - y^*, x - y\rangle \geq 0$.

We consider proper, convex and lower semicontinuous functions f, $g : X \to \mathbb{R} \cup \{+\infty\}$ such that

$$\partial f(x) \subset \partial g(x), \quad \forall x \in X. \tag{2.1}$$

At this place it is worth noting that the formally weaker assumption: it is given a proper lower semicontinuous function $f : X \to \mathbb{R} \cup \{+\infty\}$ and a proper, convex and lower semicontinuous function $g : X \to \mathbb{R} \cup \{+\infty\}$ such that (2.1) holds for ∂f, where $\partial : X \rightrightarrows X^*$ is some arbitrary presubdifferential in fact entails the convexity of f.

Following Thibault and Zagrodny [160, 161], Correa, Jofre and Thibault [53] we define

Definition 2.1.1. *Presubdifferential operator* is an operator ∂ which associates to any function $f : X \to \mathbb{R} \cup \{+\infty\}$ and any point $x \in X$ a (possibly empty) subset $\partial f(x)$ of X^* that we will call a *subdifferential* of f at x such that:

Property 1. $\partial f(x) = \emptyset$ if $x \notin \operatorname{dom} g$;

Property 2. $\partial f(x) = \partial g(x)$ whenever f and g coincide on a neighbourhood of x;

Property 3. $\partial f(x)$ is equal to the subdifferential in the sense of convex analysis whenever f is convex;

Property 4. For g lower semicontinuous near x and f convex and continuous on a neighbourhood of x, whenever $x \in \operatorname{dom}(f + g)$ is a local minimum point of $f + g$,

$$0 \in \partial f(x) + \limsup_{y \to_g x} \partial g(y),$$

where lim sup denotes the weak star sequential upper limit and $y \to_g x$ means that $y \to x$ and $g(y) \to g(x)$.

It is easy to see that any abstract subdifferential operator according to Definition 1.1.1 is a presubdifferential operator according to Definition 2.1.1. Further, most of the widely used subdifferentials over appropriate Banach spaces are subdifferentials in the sense of the above definition (see Thibault and Zagrodny [160]) – the Clarke-Rockafellar subdifferential [44], the Michel-Penôt subdifferential [120], the Mordukhovich subdifferential [121], the Kruger-Mordukhovich subdifferential [106, 107], the Ioffe A-subdifferential [87, 88]), the limiting Fréchet subdifferential, the limiting proximal subdifferential (see for example Ioffe [89], Kruger [106]), as well as the Fréchet subdifferential, the proximal subdifferential, etc.

It is well-known that if $g : X \to \mathbb{R} \cup \{+\infty\}$ is a proper, convex and lower semicontinuous function, then the convex subdifferential ∂g is a monotone operator. Inclusion (2.1) for a lower semicontinuous function f and ∂f where ∂ is arbitrary subdifferential entails that ∂f is a monotone operator too. But the monotonicity of a subdifferential ∂f yields the convexity of the function f (see Correa, Jofre and Thibault [53, Theorem 2.4] that generalizes Correa, Jofre and Thibault [52, Theorem 3]). So, any lower semicontinuous function f satisfying (2.1) with arbitrary subdifferential ∂ has to be convex and thanks to Property 3 in Definition 2.1.1 the inclusion (2.1) holds for its convex subdifferential ∂f.

Going back to the case of a Hilbert space H let us recall that the function

$$\varphi(z) := \min_{u \in H} \left\{ f(u) + \frac{1}{2}\|u - z\|^2 \right\} = f(prox_f z) + \frac{1}{2}\|prox_f z - z\|^2 \tag{2.2}$$

is convex with continuous Fréchet derivative. Moreover, $z = x + y$, where $x = prox_f z$ and it is easy to derive that

$$\varphi'(z) \in \partial f(x) \cap \left(\frac{1}{2}\|\cdot\|^2\right)'(y) \subset \partial g(x) \cap \left(\frac{1}{2}\|\cdot\|^2\right)'(y) = \psi'(z) \tag{2.3}$$

In the case when X is not a Hilbert space, the function

$$u \to f(u) + \frac{1}{2}\|u - z\|^2$$

may not attains its minimum and a weaker version of the equality (2.2) involving ε-minima has to be considered, as well as, a relevant generalization of (2.3).

We prefer to perturb the convex function $f(u)$ with functions of the type $n\|u - z\|$ instead of the square of the norm. In other words, we use Hausdorff regularization instead of Moreau-Yosida regularization. Hence, we will consider the functions

$$u \to f(u) + n\|u - z\|$$

and we will study the relations between the intersection of ε-subdifferentials of the convex functions $f(\cdot)$ and $n\|\cdot - z\|$ at the points of ε-minima of the function $u \to f(u) + n\|u - z\|$ and the ε-subdifferential of the approximating function.

Let us recall that for $\varepsilon \geq 0$, the *ε–subdifferential* of a proper, convex and lower semicontinuous function $f : X \to \mathbb{R} \cup \{+\infty\}$ at $x \in \operatorname{dom} f$ is the set

$$\partial_\varepsilon f(x) = \{p \in X^* \;:\; -\varepsilon + \langle p, y - x\rangle \leq f(y) - f(x), \quad \forall y \in X\}. \tag{2.4}$$

Of course, for $\varepsilon = 0$, $\partial_0 f(x) \equiv \partial f(x)$. But while $\partial f(x)$ could be empty, for any $x \in \operatorname{dom} f$ and $\varepsilon > 0$, the sets $\partial_\varepsilon f(x)$ are non-empty. Moreover, for any real numbers ε_1 and ε_2 such that $0 < \varepsilon_1 \leq \varepsilon_2$ one has $\partial_{\varepsilon_1} f(x) \subset \partial_{\varepsilon_2} f(x)$ and $\partial f(x) = \bigcap_{\varepsilon > 0} \partial_\varepsilon f(x)$.

The result of Brøndsted and Rockafellar saying that the graph of $\partial_\varepsilon f$ is close to the graph of ∂f is well known:

Brøndsted-Rockafellar Lemma ([36]). Let $f : X \to \mathbb{R} \cup \{+\infty\}$, be a proper, convex and lower semicontinuous function, let $\varepsilon > 0$ and $p \in \partial_\varepsilon f(x)$. Then there exists $q \in \partial f(z)$ such that

$$\|z - x\| \leq \sqrt{\varepsilon}, \text{ and } \|q - p\| \leq \sqrt{\varepsilon}.$$

Proof of the integration result for convex function

Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be a proper, lower semicontinuous and convex function.

For $n \in \mathbb{N}$ define the inf-convolutions $\{f_n\}$ by

$$f_n(x) := \inf_{y \in X} \{f(y) + n\|x - y\|\}. \tag{2.5}$$

(The approximating sequence $\{f_n\}$ was originally introduced by Hausdorff [82] for any lower bounded lower semicontinuous function f of a real variable.) It is clear that for sufficiently large n the function f_n is finite valued and we will always consider this case even if it is not stated explicitly. Some well-known properties of these inf-convolutions of f (see, for instance, Laurent [109], Hiriart-Urruty [84], Fitzpatrick and Phelps [76]) are listed in next

Lemma 2.1.2. For a proper, lower semicontinuous and convex function $f : X \to \mathbb{R} \cup \{+\infty\}$ and n large enough,

(i) f_n is convex and n-Lipschitzian;

(ii) $f_n(x) \leq f_{n+1}(x) \leq f(x)$ for all $x \in X$ and all $n \in \mathbb{N}$;

(iii) $f_n(x) \to f(x)$ as $n \to \infty$ for all $x \in X$.

Let us denote the set of all ε-minima of the function $f(\cdot) + n\|x - \cdot\|$ over X by

$$M_\varepsilon^n(f; x) := \{y \in X : f(y) + n\|x - y\| \le f_n(x) + \varepsilon\}.$$

It is clear that for $\varepsilon > 0$ the sets $M_\varepsilon^n(f; x)$ are non-empty.

Since f_n are convex continuous functions, then $\partial f_n(x)$ is a nonempty set for all $x \in X$. We will show that if p is a subgradient of f_n at x, and y is an arbitrary ε-minimum point of $f(\cdot) + n\|x - \cdot\|$, then actually p is an ε-subgradient of f at y, as well as, it is an ε-subgradient of the function $n\|\cdot\|$ at $x - y$:

Lemma 2.1.3. For any $\varepsilon \ge 0$, and any $y \in M_\varepsilon^n(f; x)$ it holds that

$$\partial f_n(x) \subset \partial_\varepsilon f(y) \cap \partial_\varepsilon n\|\cdot\|(x - y). \tag{2.6}$$

Proof. Take $p \in \partial f_n(x)$ and $y \in M_\varepsilon^n(f; x)$. By the definition of subgradient, for all $x' \in X$,

$$\langle p, x' - x\rangle \le f_n(x') - f_n(x). \tag{2.7}$$

We estimate from above the right hand side of (2.7) by using that $f_n(x') \le f(x')$ (see Lemma 2.1.2 (ii)) and $f(y) + n\|x - y\| \le f_n(x) + \varepsilon$ (because $y \in M_\varepsilon^n(f; x)$) and we get

$$f_n(x') - f_n(x) \le f(x') - f(y) - n\|x - y\| + \varepsilon. \tag{2.8}$$

We estimate from below the left hand side of (2.7) using that $\|p\| \le n$ (see Lemma 2.1.2 (i)), thus

$$\langle p, x' - x\rangle \ge \langle p, x' - y\rangle - \|p\|\|x - y\| \ge \langle p, x' - y\rangle - n\|x - y\|. \tag{2.9}$$

Combining (2.8) and (2.9) with (2.7) we get for all $x' \in X$ that

$$\langle p, x' - y\rangle - n\|x - y\| \le f(x') - f(y) - n\|x - y\| + \varepsilon,$$

or

$$\langle p, x' - y\rangle \le f(x') - f(y) + \varepsilon,$$

which means that $p \in \partial_\varepsilon f(y)$.

Other obvious estimation of the right hand side of (2.7) gives us for all $x' \in X$ that

$$\langle p, x' - x\rangle \le f_n(x') - f_n(x) \le$$
$$\le f(y) + n\|x' - y\| - f(y) - n\|x - y\| + \varepsilon =$$
$$n\|x' - y\| - n\|x - y\| + \varepsilon,$$

which means that $p \in \partial_\varepsilon n\|\cdot\|(x - y)$. □

We recall that the proper, convex and lower semicontinuous functions $f,\ g : X \to \mathbb{R} \cup \{+\infty\}$ are such that

$$\partial f(x) \subset \partial g(x), \quad \forall x \in X.$$

Since $\operatorname{dom} \partial f$ is dense in $\operatorname{dom} f$ and the latter is non-empty because f is assumed to be proper, then there exist $\overline{x} \in \operatorname{dom} \partial f$ and $\overline{p} \in \partial f(\overline{x})$.

Define the function $\overline{f}(x) := f(x + \overline{x}) - \langle \overline{p}, x \rangle - f(\overline{x})$.

The function $\overline{f} : X \to \mathbb{R} \cup \{+\infty\}$ is proper, convex and lower semicontinuous, $\overline{f}(0) = 0$ and $\partial \overline{f}(x) = \partial f(x + \overline{x}) - \overline{p}$. Hence, $0 \in \partial \overline{f}(0)$.

Analogously, define $\overline{g}(x) := g(x + \overline{x}) - \langle \overline{p}, x \rangle - g(\overline{x})$, and observe that $\overline{g} : X \to \mathbb{R} \cup \{+\infty\}$ is proper, convex and lower semicontinuous function, $\overline{g}(0) = 0$ and $\partial \overline{g}(x) = \partial g(x + \overline{x}) - \overline{p}$. Moreover, $0 \in \partial \overline{g}(0)$, since $\overline{p} \in \partial f(\overline{x}) \subset \partial g(\overline{x})$.

It is clear also that $\partial \overline{f}(x) \subset \partial \overline{g}(x)$ for all $x \in X$.

Lemma 2.1.4. Let $\overline{f} : X \to \mathbb{R} \cup \{+\infty\}$ be a proper, convex and lower semicontinuous function such that $\overline{f}(0) = 0$ and $0 \in \partial \overline{f}(0)$. Let $s > 0$. Then for $x \in B(0, s)$, $M_\varepsilon^n(\overline{f}, x) \subset B[0, 3s]$ for $n \geq 1/s$ and $\varepsilon \leq 1$.

Proof. From $\overline{f}(0) = 0$ and $0 \in \partial \overline{f}(0)$ it follows that $\overline{f}(y) \geq 0$ for any $y \in X$. Then

$$\overline{f}(y) + n\|x - y\| \geq n\|x - y\| \geq n\|y\| - n\|x\|,$$

and if $\|x\| < s$ and $\|y\| > 3s$ then

$$\overline{f}(y) + n\|x - y\| \geq n\|y\| - n\|x\| > 3ns - ns = 2ns = ns + ns >$$

$$n\|x\| + 1 \geq \inf_y \{\overline{f}(y) + n\|x - y\|\} + 1 = \overline{f}_n(x) + 1,$$

and no such y can be in $M_\varepsilon^n(\overline{f}, x)$ for $\varepsilon \leq 1$. □

Lemma 2.1.5. Let $\overline{f}, \overline{g} : X \to \mathbb{R} \cup \{+\infty\}$ be proper, convex and lower semicontinuous functions such that $\overline{f}(0) = \overline{g}(0) = 0$, $0 \in \partial \overline{f}(0) \cap \partial \overline{g}(0)$, and $\partial \overline{f}(x) \subset \partial \overline{g}(x)$ for all $x \in X$. Let $s > 0$. Then for $n \geq 1/s$, $\partial \overline{f}_n(x) \subset \partial \overline{g}_n(x)$ for all $x \in B(0, s)$.

Proof. Fix $\varepsilon \in (0, 1]$ and $n \geq \dfrac{1}{s}$.

Take arbitrary $x \in B(0, s)$ and let $p \in \partial \overline{f}_n(x)$. Obviously, $\|p\| \leq n$.

By Lemma 2.1.3, $p \in \partial_\varepsilon \overline{f}(y) \cap \partial_\varepsilon n\|x - y\|$ for all $y \in M_\varepsilon^n(\overline{f}, x)$. Fix $y \in M_\varepsilon^n(\overline{f}, x)$ and note that $y \in B[0, 3s]$ by Lemma 2.1.4.

Since $p \in \partial_\varepsilon \overline{f}(y)$ by Brøndsted-Rockafellar lemma there exists $q \in \partial \overline{f}(z)$ for some z such that $\|z - y\| \leq \sqrt{\varepsilon}$ and $\|p - q\| \leq \sqrt{\varepsilon}$.

Since $\partial\overline{f}(z) \subset \partial\overline{g}(z)$ it follows that $q \in \partial\overline{g}(z)$, and

$$\overline{g}(x') - \overline{g}(z) \geq \langle q, x' - z\rangle, \quad \forall x' \in X. \tag{2.10}$$

Since $p \in \partial_\varepsilon n\|x - y\|$ this means that

$$n\|w - x'\| - n\|x - y\| \geq \langle p, w - x' + y - x\rangle - \varepsilon, \quad \forall w, x' \in X.$$

Hence, for all $w, x' \in X$ we have

$$\begin{aligned} n\|w - x'\| - n\|x - z\| &\geq n\|x - y\| - n\|x - z\| + \langle p, w - x' + z - x\rangle + \\ &\quad \langle p, y - z\rangle - \varepsilon \\ &\geq \langle p, w - x' + z - x\rangle - 2n\|y - z\| - \varepsilon \\ &\geq \langle p, w - x' + z - x\rangle - 2n\sqrt{\varepsilon} - \varepsilon. \end{aligned} \tag{2.11}$$

By (2.10) and (2.11) for all $w, x' \in X$ we get

$$\begin{aligned} \overline{g}(x') + n\|w - x'\| &\geq \overline{g}(z) + n\|x - z\| + \langle p, w - x\rangle + \\ &\quad \langle q - p, x' - z\rangle - 2n\sqrt{\varepsilon} - \varepsilon \\ &\geq \overline{g}_n(x) + \langle p, w - x\rangle - \sqrt{\varepsilon}\|x' - z\| - 2n\sqrt{\varepsilon} - \varepsilon. \end{aligned}$$

Applying Lemma 2.1.4 for $w \in B(0, s)$ and the function $\overline{g}$ we get

$$\overline{g}_n(w) = \inf_{x' \in B[0,3s]} \{\overline{g}(x') + n\|w - x'\|\},$$

and

$$\begin{aligned} \overline{g}_n(w) &= \inf_{x' \in B[0,3s]} \{\overline{g}(x') + n\|w - x'\|\} \\ &\geq \overline{g}_n(x) + \langle p, w - x\rangle - \sqrt{\varepsilon} \sup_{x' \in B[0,3s]} \|x' - z\| - 2n\sqrt{\varepsilon} - \varepsilon \\ &\geq \overline{g}_n(x) + \langle p, w - x\rangle - \sqrt{\varepsilon}(6s + \sqrt{\varepsilon}) - 2n\sqrt{\varepsilon} - \varepsilon. \end{aligned}$$

Letting ε to zero we obtain that

$$\overline{g}_n(w) \geq \overline{g}_n(x) + \langle p, w - x\rangle, \quad \forall w \in B(0, s),$$

which means that $p \in \partial\overline{g}_n(x)$.

Finally, for any $n \geq 1/s$, $\partial\overline{f}_n(x) \subset \partial\overline{g}_n(x)$ for all $x \in B(0, s)$. □

For continuous convex functions – such as $\overline{f}_n$ and $\overline{g}_n$ – the inclusion of subdifferentials easily yields that they differ by a finite constant (see Rockafellar [147]).

As an immediate consequence of this and Lemma 2.1.5 one deduces

Corollary 2.1.6. Let $\overline{f}, \overline{g} : X \to \mathbb{R} \cup \{+\infty\}$ be proper, convex and lower semicontinuous functions such that $\overline{f}(0) = \overline{g}(0) = 0$, $0 \in \partial\overline{f}(0) \cap \partial\overline{g}(0)$, and $\partial\overline{f}(x) \subset \partial\overline{g}(x)$ for all $x \in X$. Let $s > 0$. For $n \geq 1/s$ there exists a finite constant c_n such that

$$\overline{f}_n(x) = \overline{g}_n(x) + c_n, \quad \text{for all} \quad x \in B(0, s). \tag{2.12}$$

Now it is easy to finish the reasoning.

Indeed, for $0 \in B(0,s)$ by Lemma 2.1.2 (iii) we have that $\overline{f}_n(0) \to \overline{f}(0) = 0$ and $\overline{g}_n(0) \to \overline{g}(0) = 0$ as $n \to \infty$. Passing to limit in (2.12) yields $\overline{f}(0) = \overline{g}(0) + \lim_{n\to\infty} c_n$, and then $\lim_{n\to\infty} c_n = 0$.

By Lemma 2.1.2 (iii) for all $x \in B(0,s)$, $\overline{f}(x) = \lim_{n\to\infty} \overline{f}_n(x)$ and $\overline{g}(x) = \lim_{n\to\infty} \overline{g}_n(x)$, hence passing to limit in (2.12) yields

$$\overline{f}(x) - \overline{g}(x) = \lim_{n\to\infty}(\overline{f}_n(x) - \overline{g}_n(x)) = \lim_{n\to\infty} c_n = 0,$$

and

$$\overline{f}(x) = \overline{g}(x), \quad \forall x \in B(0,s).$$

Having in mind that s was arbitrary, this means that

$$\overline{f}(x) = \overline{g}(x), \quad \forall x \in X.$$

The latter is equivalent to say that for all $x \in X$,

$$f(x+\overline{x}) - \langle \overline{p}, x\rangle - f(\overline{x}) = g(x+\overline{x}) - \langle \overline{p}, x\rangle - g(\overline{x}),$$

$$f(x+\overline{x}) - g(x+\overline{x}) = f(\overline{x}) - g(\overline{x}),$$

or $f(x) = g(x) + c$ for all $x \in X$ with $c = f(\overline{x}) - g(\overline{x})$.

This completes the proof.

The regularization technique presented in this section can be used for establishing integrability of convex composite function on Asplund spaces but this will be a subject of future work.

2.2 Integrability of subdifferentials of directionally Lipschitz functions

We will use a quantitative version of subdifferential characterization of directionally Lipschitz functions to study the integrability of subdifferentials of such functions over arbitrary Banach space.

We prove results concerning the subdifferential characterization of directionally Lipschitz property of a given function. We finish by establishing the local integrability of subdifferentials of strictly directionally Lipschitz regular functions, continuous on their domains (Theorem 2.2.6).

Let us begin by giving some necessary definitions and preliminaries.

We work in a real Banach space X with *open unit ball* B° and topological dual X^*.

We will work also with a general *presubdifferential operator* ∂ that associates with each function $g : X \to \mathbb{R} \cup \{+\infty\}$ and with each point $x \in X$ a subset $\partial g(x)$ of X^*, that we will call a *subdifferential of* g *at* x, and for which the properties listed in Definition 2.1.1 hold.

Recall that the lower semicontinuous function $g : X \to \mathbb{R} \cup \{+\infty\}$ is said to be *directionally Lipschitz* at $x_0 \in \operatorname{dom} g$ with respect to a vector $h_0 \in X$ (see Rockafellar [150]), if there exist constants $K \in \mathbb{R}$, $\varepsilon > 0$, $\delta > 0$ such that

$$\begin{aligned}
&t^{-1}[g(x+th) - g(x)] \le K, \\
&\forall t \in]0,\varepsilon],\ \forall h \in h_0 + \delta B^\circ,\ \forall x \in x_0 + \delta B^\circ \text{ with } |g(x) - g(x_0)| \le \varepsilon.
\end{aligned} \tag{2.13}$$

It is clear that g is Lipschitz around x_0 exactly when it is directionally Lipschitz at x_0 with respect to $h_0 = 0$. Also, it is easy to see that in the case when the lower semicontinuous function g considered in the definition above is convex, or continuous relative to its domain (i.e. for any $x_0 \in \operatorname{dom} g$ and any $\gamma > 0$ there exists $\eta > 0$ such that $|g(x) - g(x_0)| \le \gamma$ for all $x \in \operatorname{dom} g \cap (x_0 + \eta B^\circ)$), then (2.13) is equivalent to the existence of constants $K \in \mathbb{R}$, $\varepsilon > 0$, $\delta > 0$ such that

$$\begin{aligned}
&t^{-1}[g(x+th) - g(x)] \le K, \\
&\forall t \in]0,\varepsilon],\ \forall h \in h_0 + \delta B^\circ,\ \text{and } \forall x \in (x_0 + \delta B^\circ) \cap \operatorname{dom} g.
\end{aligned} \tag{2.14}$$

The lower semicontinuous function $g : X \to \mathbb{R} \cup \{+\infty\}$ is said to be (see Jofre and Thibault [97]) *strictly directionally Lipschitz* at $x_0 \in \operatorname{dom} g$ with respect to $h_0 \in X$ if it is directionally Lipschitz at x_0 with respect to h_0 with some constants K, ε, δ satisfying (2.13) and, moreover,

$$\begin{aligned}
&g \text{ is locally Lipschitz on any set } x+]0,\varepsilon](h_0 + \delta B^\circ), \\
&\text{where } x \in x_0 + \delta B^\circ \text{ and } |g(x) - g(x_0)| \le \varepsilon.
\end{aligned} \tag{2.15}$$

If the lower semicontinuous function g in the latter definition is also supposed to be continuous relative to its domain, or convex, then (2.15) can be replaced by:

$$\begin{aligned}
&g \text{ is locally Lipschitz on any set } x+]0,\varepsilon](h_0 + \delta B^\circ), \\
&\text{where } x \in (x_0 + \delta B^\circ) \cap \operatorname{dom} g.
\end{aligned} \tag{2.16}$$

Let us also recall that the lower semicontinuous function $g : X \to \mathbb{R} \cup \{+\infty\}$ is said to be *regular at* $x_0 \in \operatorname{dom} g$ if $\liminf\limits_{\substack{h' \to h \\ t \downarrow 0}} t^{-1}[g(x_0 + th') - g(x_0)] = g^\uparrow(x_0; h)$ for any

$h \in X$, where

$$d^{\uparrow}\varphi(x_0; h) = \sup_{\eta>0} \limsup_{\substack{x \to_{\varphi} x_0 \\ t \downarrow 0}} \left(\inf_{h' \in h+\eta B_X^{\circ}} t^{-1}[\varphi(x+th') - \varphi(x)] \right).$$

is the Clarke generalized derivative and it is said to be *regular* if it is regular at any point of its domain.

We wish to remind the Mean Value Theorem, established by Zagrodny in [170] (see also Thibault and Zagrodny [161]), as it will be essentially used in what follows.

Mean Value Theorem of Zagrodny [170, 161]. Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous function, let $a, b \in \operatorname{dom} f$ with $a \neq b$ and let ∂ be any presubdifferential operator. Then there exist $x_n \to_f c \in [a, b[:= \{(1-t)a + tb : t \in [0, 1[\}$, and $x_n^* \in \partial f(x_n)$ such that:

(i) $f(b) - f(a) \leq \lim_{n\to\infty} \langle x_n^*, b - a \rangle$;

(ii) $\dfrac{\|b - c\|}{\|b - a\|}(f(b) - f(a)) \leq \lim_{n\to\infty} \langle x_n^*, b - x_n \rangle$;

(iii) $\|b - a\|(f(c) - f(a)) \leq \|c - a\|(f(b) - f(a))$.

2.2.1 Subdifferential properties of directionally Lipschitz functions

In this subsection we study how the directionally Lipschitz property of the function refers to the properties of its subdifferential. Work in this direction with the Clarke subdifferential is that of Treiman [165], which is strongly based on the technique of Bishop and Phelps and on the result (see Treiman [165]) stating that $\liminf_{S \ni y \to x} K(S; y) \subset T(S; x)$ (here $K(S; .)$ and $T(S; .)$ denote respectively the Bouligand contingent cone and the Clarke tangent cone of a closed subset $S \subset X$). Our results are given in terms of arbitrary subdifferential and their proofs are merely based on the Mean Value Theorem. They also may be considered as a quantitative version of a result of Treiman [165, Theorem 6]. This quantitative version will be needed further to establish Lemma 2.2.4 that is a key step in our development.

In all the sequel ∂ stands for any presubdifferential operator such that the corresponding subdifferential is included in the Clarke subdifferential, i.e., $\partial g(x) \subset \partial_C g(x)$ for any function $g : X \to \mathbb{R} \cup \{+\infty\}$ and any $x \in X$.

Lemma 2.2.1. Assume that the lower semicontinuous function $g : X \to \mathbb{R} \cup \{+\infty\}$ is directionally Lipschitz at x_0 with respect to h_0 with constants K, ε, δ satisfying (2.13) (resp. (2.14)). Then

$$\langle x^*, h_0 \rangle + \delta \|x^*\| \le K, \ \forall x^* \in \partial g(x), \ \forall x \in x_0 + \delta B^\circ \text{ with } |g(x) - g(x_0)| \le \varepsilon, \tag{2.17}$$

(resp.

$$\langle x^*, h_0 \rangle + \delta \|x^*\| \le K, \ \forall x^* \in \partial g(x), \ \forall x \in x_0 + \delta B^\circ). \tag{2.18}$$

Proof. Let $g : X \to \mathbb{R} \cup \{+\infty\}$ be directionally Lipschitz at x_0 with respect to h_0 with constants K, ε, δ satisfying (2.14) (resp. (2.13)). Then the condition (2.14) (resp. (2.13)) obviously entails

$$g^\uparrow(x; h) \le K, \ \forall h \in h_0 + \delta B^\circ, \ \forall x \in x_0 + \delta B^\circ \text{ (with } |g(x) - g(x_0)| \le \varepsilon).$$

Hence, if $\partial g(x) \ne \emptyset$ for some $x \in x_0 + \delta B^\circ$ (with $|g(x) - g(x_0)| \le \varepsilon$), then as $\partial g(x) \subset \partial_C g(x)$ we have

$$\langle x^*, h \rangle \le g^\uparrow(x; h) \le K, \ \forall h \in h_0 + \delta B^\circ, \ \forall x^* \in \partial g(x) \text{ (with } |g(x) - g(x_0)| \le \varepsilon), \text{ i.e.,}$$
$$\langle x^*, h_0 \rangle + \delta \|x^*\| \le K, \ \forall x^* \in \partial g(x), \ x \in x_0 + \delta B^\circ \text{ (with } |g(x) - g(x_0)| \le \varepsilon). \quad \square$$

We proceed to show the reverse implication, i.e., that (2.18) yields the property (2.14). As it can be seen below, that case is more simple than the one establishing that (2.17) ensures the property (2.13). So, we made the choice of separating the two proofs.

Lemma 2.2.2. Let $g : X \to \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous function, let $x_0 \in \operatorname{dom} g$ and $h_0 \in X$. Assume that there exist constants $K \in \mathbb{R}$ and $\delta > 0$ such that (2.18) holds.

Then for any positive numbers ε_0 and δ_0 such that $\varepsilon_0(\|h_0\| + \delta_0) + \delta_0 < \delta$, the property (2.14) holds for the function g with K, ε_0, δ_0.

Proof. Fix $\varepsilon_0 > 0$ and $\delta_0 > 0$ such that $\varepsilon_0(\|h_0\| + \delta_0) + \delta_0 < \delta$ and take any $x \in (x_0 + \delta_0 B^\circ) \cap \operatorname{dom} g$, any $h \in h_0 + \delta_0 B^\circ$ with $h \ne 0$, and any $t \in]0, \varepsilon_0]$. Consider $g(x + th)$ and fix arbitrary real number $r \in \mathbb{R}$ with $r < g(x + th)$. Define the lower semicontinuous function

$$g_r(y) := \begin{cases} g(y), & \text{if } y \ne x + th, \\ r, & \text{if } y = x + th. \end{cases}$$

Apply the Mean Value Theorem of Zagrodny to estimate

$$g_r(x+th) - g_r(x) = r - g(x) \leq \lim_{n\to\infty} \langle x_n^*, th\rangle \tag{2.19}$$

with $x_n^* \in \partial g(x_n)$, $x_n \underset{n\to\infty}{\longrightarrow} z \in [x, x+th[$, $g(x_n) \underset{n\to\infty}{\longrightarrow} g(z)$. Then for n large enough,

$$\begin{aligned}\|x_n - x_0\| &\leq \|x_n - z\| + \|z - x\| + \|x - x_0\| \leq \|x_n - z\| + t\|h\| + \delta_0 \\ &\leq \|x_n - z\| + \varepsilon_0(\|h_0\| + \delta_0) + \delta_0 < \delta.\end{aligned}$$

Hence, $x_n \in x_0 + \delta B^\circ$ and we can use (2.18) to get

$$\begin{aligned}\langle x_n^*, th\rangle &= \langle x_n^*, th_0\rangle + t\langle x_n^*, h - h_0\rangle \leq t[\langle x_n^*, h_0\rangle + \langle x_n^*, h - h_0\rangle] \\ &\leq t[\langle x_n^*, h_0\rangle + \delta_0\|x_n^*\|] \leq tK.\end{aligned}$$

This implies because of (2.19)

$$r - g(x) \leq tK,$$

which yields, on the one hand, that $g(x+th)$ is finite (i.e. $x + th \in \operatorname{dom} g$) and, on the other hand, that

$$g(x+th) - g(x) \leq tK.$$

Observe that the inequality also holds for $h = 0$, in the case when $0 \in h_0 + \delta_0 B^\circ$ since $K \geq 0$ in that case because of (2.18). Therefore, the property (2.14) holds for g at x_0 with respect to h_0 with K, ε_0, δ_0 as above. □

Theorem 2.2.3 (subdifferential characterization of directionally Lipschitz property). Let $g : X \to \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous function, let $x_0 \in \operatorname{dom} g$ and $h_0 \in X$. Then g is directionally Lipschitz at x_0 with respect to h_0 with constants K, ε', δ' satisfying (2.13), if and only if, there exist constants $\varepsilon > 0$ and $\delta > 0$ such that (2.17) holds with K, ε, δ.

Proof. The implication $\Longrightarrow$ follows from Lemma 2.2.1.

Let us prove the converse.

Without loss of generality we may suppose that $\varepsilon \in]0, 1[$ and $\delta \in]0, \varepsilon[$, and that they are such that

$$g(x_0) \leq g(x) + \varepsilon, \ \forall x \in x_0 + \delta B^\circ. \tag{2.20}$$

Now take $\varepsilon' > 0$ such that $\varepsilon'(\|h_0\| + |K| + 2) < \delta$, and $\delta' \in]0, \varepsilon'[$. Fix any $x \in x_0 + \delta' B^\circ$ with $|g(x) - g(x_0)| \leq \varepsilon'$, any $h \in h_0 + \delta' B^\circ$ with $h \neq 0$, and any $t \in]0, \varepsilon']$. Consider $g(x + th)$ and suppose $g(x + th) > g(x) + tK$. We may choose and fix a real number

$\mu \in]0, 1[$ such that $g(x + th) > g(x) + t(K + \mu)$. Set $r := g(x) + t(K + \mu)$ and define the lower semicontinuous function

$$g_r(y) := \begin{cases} g(y), & \text{if } y \neq x + th, \\ r, & \text{if } y = x + th. \end{cases}$$

Apply the Mean Value Theorem of Zagrodny to estimate

$$g_r(x + th) - g_r(x) = r - g(x) \leq \lim_{n\to\infty} \langle x_n^*, th\rangle \tag{2.21}$$

with some sequences $x_n^* \in \partial g(x_n)$, $x_n \underset{n\to\infty}{\longrightarrow} z \in [x, x + th[$, $g(x_n) \underset{n\to\infty}{\longrightarrow} g(z)$. Then, for n large enough, $\|x_n - x_0\| \leq \|x_n - z\| + \varepsilon'(\|h_0\| + \delta') + \delta' < \delta$.

Case I. $z = x$.
Then $g(x_n) \underset{n\to\infty}{\longrightarrow} g(x)$ and for n large enough

$$|g(x_n) - g(x_0)| \leq |g(x_n) - g(x)| + |g(x) - g(x_0)| \leq |g(x_n) - g(x)| + \varepsilon' < \varepsilon.$$

Case II. $z \neq x$.
Then Mean Value Theorem of Zagrodny and in particular (iii) ensures that

$$t\|h\|(g(z) - g(x)) \leq \|z - x\|(r - g(x)).$$

We rewrite the left hand side

$$\begin{aligned} t\|h\|(g(z) - g(x_0)) &\leq \|z - x\|(r - g(x)) + t\|h\|(g(x) - g(x_0)) \\ &\leq \|z - x\|(r - g(x)) + t\|h\|\varepsilon' \\ &\leq t\|h\|(r - g(x) + \varepsilon'), \end{aligned}$$

from where

$$\begin{aligned} g(z) - g(x_0) &\leq r - g(x) + \varepsilon' = t(K + \mu) + \varepsilon' \\ &\leq t(|K| + \mu) + \varepsilon' \leq \varepsilon'(|K| + \mu + 1) \leq \varepsilon'(|K| + 2). \end{aligned}$$

Hence, for sufficiently large n we have

$$g(x_n) - g(x_0) = g(x_n) - g(z) + g(z) - g(x_0) \leq |g(x_n) - g(z)| + \varepsilon'(|K| + 2) < \varepsilon.$$

Obviously, for n large enough, the points $x_n \in x_0 + \delta B^\circ$ and hence by (2.20) we also have that $g(x_0) - g(x_n) \leq \varepsilon$. Therefore, we obtain $|g(x_n) - g(x_0)| \leq \varepsilon$.

In both cases one has for n large enough, $x_n \in x_0 + \delta B^\circ$ and $|g(x_n) - g(x_0)| \leq \varepsilon$, then by (2.17) one has $\langle x_n^*, h\rangle \leq K$. So, by (2.21)

$$r - g(x) \leq tK, \text{ i.e., } t(K + \mu) \leq tK,$$

and the latter yields $\mu \leq 0$, which is a contradiction, since $\mu > 0$. We conclude that $g(x + th) \leq g(x) + tK$, and further, if $0 \in h_0 + \delta' B^\circ$, the inequality still holds for $h = 0$ because, according to (2.17), $K \geq 0$ in that case. The proof is then complete. □

2.2.2 Local integrability

In this subsection we use the results proved in the previous one to establish the local integrability of any subdifferential of a class of directionally Lipschitz functions.

We begin by showing how the property (2.14) is entailed by inclusion of subdifferentials for lower semicontinuous functions.

Lemma 2.2.4. Let $g : X \to \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous function which satisfies property (2.14) at $x_0 \in \operatorname{dom} g$ with respect to $h_0 \in X$ with some constants K, ε, δ. Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous function such that $\partial f(x) \subset \partial g(x)$ for any $x \in x_0 + \delta B^\circ$.

Then for any constants $\delta_0 > 0$ and $\varepsilon_0 > 0$ such that $\delta_0 + \varepsilon_0(\|h_0\| + \delta_0) < \delta$ with $\operatorname{dom} f \cap (x_0 + \delta_0 B^\circ) \neq \emptyset$, one has that the property (2.14) holds for f with the constants K, ε_0, δ_0, i.e.,

$$t^{-1}[f(x + th) - f(x)] \leq K, \ \forall t \in]0, \varepsilon_0], \ \forall h \in h_0 + \delta_0 B^\circ, \text{ and } \forall x \in (x_0 + \delta_0 B^\circ) \cap \operatorname{dom} f.$$

In particular, $x + [0, \varepsilon_0](h_0 + \delta_0 B^\circ) \subset \operatorname{dom} f$, for all $x \in \operatorname{dom} f \cap (x_0 + \delta_0 B^\circ)$. Further, f is directionally Lipschitz at x_0 with respect to h_0, whenever $x_0 \in \operatorname{dom} f$.

Proof. Suppose that g satisfies (2.14) with constants K, ε, δ. By Lemma 2.2.1, for any $x \in \operatorname{dom} g \cap (x_0 + \delta B^\circ)$ and any $x^* \in \partial g(x)$ we have $\langle x^*, h_0\rangle + \delta\|x^*\| \leq K$. Because of the assumption $\partial f(x) \subset \partial g(x)$ for any $x \in x_0 + \delta B^\circ$, that inequality holds in particular for all $x \in \operatorname{dom} \partial f \cap (x_0 + \delta B^\circ)$ and $x^* \in \partial f(x)$.

Fix any $\delta_0 > 0$ and $\varepsilon_0 > 0$ such that $\delta_0 + \varepsilon_0(\|h_0\| + \delta_0) < \delta$. Lemma 2.2.2 ensures that for all $x \in \operatorname{dom} f \cap (x_0 + \delta_0 B^\circ)$, $t \in]0, \varepsilon_0]$, and $h \in h_0 + \delta_0 B^\circ$ one has $t^{-1}[f(x + th) - f(x)] \leq K$. The proof is then complete. □

We will also need the following second lemma of general interest. It concerns the graphical density of the domain of any presubdifferential.

Lemma 2.2.5. Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous function with $\operatorname{dom} f \neq \emptyset$. Then $\operatorname{dom} \partial f$ is f-graphically dense in $\operatorname{dom} f$, i.e., for any $a \in \operatorname{dom} f$ there exists a sequence $x_n \in \operatorname{dom} \partial f$ with $x_n \underset{n\to\infty}{\longrightarrow} a$ and $f(x_n) \underset{n\to\infty}{\longrightarrow} f(a)$.

Proof. Fix $a \in \operatorname{dom} f$ and $\varepsilon > 0$. Choose by the lower semicontinuity of f some positive number $r < \varepsilon$ such that $f(x) > f(a) - \varepsilon$ for all $x \in a + rB^\circ$. If for any $b \in a + rB^\circ$ one has $f(b) \geq f(a)$, then a is a local minimum point of f and by Property 4 we have that

$$0 \in \limsup_{x_n \to_f a} \partial f(x_n),$$

i.e., there exist $x_n \in \operatorname{dom} \partial f$ such that $x_n \underset{n\to\infty}{\longrightarrow} a$ and $f(x_n) \underset{n\to\infty}{\longrightarrow} f(a)$. Otherwise, there exists some $b \in a + rB^\circ$ with $f(b) < f(a)$ and the Mean Value Theorem of Zagrodny yields $x_n \underset{n\to\infty}{\longrightarrow} c \in [a, b[$ with $f(x_n) \underset{n\to\infty}{\longrightarrow} f(c)$ and $\partial f(x_n) \neq \emptyset$, and such that the conclusion (iii) of that theorem holds. The latter gives $\lim_{n\to\infty} f(x_n) = f(c) \leq f(a)$. We deduce the existence of some N such that $\|x_n - a\| < r < \varepsilon$ and $|f(x_n) - f(a)| < \varepsilon$ for all $n \geq N$, and hence the proof is complete. □

We establish now the integrability result.

Theorem 2.2.6 (integrability of regular directionally Lipschitz functions). Let $g : X \to \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous regular function, continuous relative to its domain, and strictly directionally Lipschitz at $x_0 \in \operatorname{dom} g$.

Then there exist constants $\alpha > 0$ and $\beta \in]0, \alpha[$ such that for any lower semicontinuous function $f : X \to \mathbb{R} \cup \{+\infty\}$ with $\operatorname{dom} f \cap (x_0 + \beta B^\circ) \neq \emptyset$ the inclusion $\partial f(x) \subset \partial g(x)$ for all $x \in x_0 + \alpha B^\circ$ implies

$$f = g + \text{const} \quad \text{on} \quad x_0 + \beta B^\circ.$$

If the strict directionally Lipschitz property for g at x_0 with respect to h_0 holds with constants K, ε, δ satisfying (2.14), one may take $\alpha = \delta$ and $\beta = \min\left\{\frac{\delta^2}{4(\|h_0\|+2\delta)}, \frac{\varepsilon\delta}{2}\right\}$.

Proof. Let K, ε, and δ be given by the strictly directionally Lipschitz property of g at x_0 with respect to h_0, i.e., properties (2.14) and (2.16) hold for them.

Set $\delta_0 := \frac{\delta}{2}$ and $\varepsilon_0 := \min\left\{\frac{\delta}{2(\|h_0\|+2\delta)}, \varepsilon\right\}$, and observe that $\varepsilon_0 < 1$. It is easy to see that the conclusion of Lemma 2.2.4 holds for such δ_0 and ε_0.

Put now $\alpha := \delta$ and $\beta := \varepsilon_0\delta_0 = \min\left\{\frac{\delta^2}{4(\|h_0\|+2\delta)}, \frac{\varepsilon\delta}{2}\right\}$.

We claim that for arbitrary $x \in x_0 + \beta B^\circ$ there exists $h_x \in h_0 + \frac{\delta}{2}B^\circ$ such that

$$x + \varepsilon_0 h_x = x_0 + \varepsilon_0 h_0. \tag{2.22}$$

Indeed, for $x \in x_0 + \beta B^\circ$ we may set $h_x := \frac{x_0 - x}{\varepsilon_0} + h_0$ and since $\frac{\|x - x_0\|}{\varepsilon_0} < \frac{\beta}{\varepsilon_0} = \frac{\varepsilon_0 \delta_0}{\varepsilon_0} = \frac{\delta}{2}$, we obtain that $h_x \in h_0 + \frac{\delta}{2} B^\circ$.

Moreover, the inclusion $h_x \in h_0 + \frac{\delta}{2} B^\circ$ ensures that

$$x +]0, \varepsilon_0] \left(h_x + \frac{\delta}{2} B^\circ \right) \subset x +]0, \varepsilon](h_0 + \delta B^\circ), \quad \forall x \in x_0 + \beta B^\circ. \tag{2.23}$$

Now, fix arbitrary $v \in \operatorname{dom} f \cap (x_0 + \beta B^\circ)$ which is a non-empty set by assumption. As the set $\operatorname{dom} \partial f$ is f-graphically dense in $\operatorname{dom} f$ by Lemma 2.2.5, we obtain a sequence $\{x_n\} \subset \operatorname{dom} \partial f \subset \operatorname{dom} \partial g \subset \operatorname{dom} g$, such that $x_n \underset{n \to \infty}{\longrightarrow} v$ and $f(x_n) \underset{n \to \infty}{\longrightarrow} f(v)$. Writing

$$\|x_n - x_0\| \le \|x_n - v\| + \|v - x_0\| < \|x_n - v\| + \beta,$$

we see that for sufficiently large n (for example $n \ge N_1$) we have

$$\|x_n - x_0\| < \beta. \tag{2.24}$$

For $n \ge N_1$ denote by C_n the open convex set $C_n := x_n +]0, \varepsilon_0] \left(h_{x_n} + \frac{\delta}{2} B^\circ \right)$. Recall that for $n \ge N_1$ we have $x_n \in (x_0 + \beta B^\circ) \cap \operatorname{dom} g$ and observe also, by what precedes, that $h_{x_n} \in \left(h_0 + \frac{\delta}{2} B^\circ \right)$. So, by the definition of strictly directionally Lipschitz property of g and by (2.23) it follows that g is locally Lipschitz and regular on C_n. From Lemma 2.2.4 it is clear that $\operatorname{dom} f \cap C_n \ne \emptyset$ for $n \ge N_1$, since $x_n +]0, \varepsilon_0] h_{x_n} \subset \operatorname{dom} f$.

It remains to observe that the inclusion of subdifferentials holds on C_n for $n \ge N_1$, since $C_n \subset x_0 + \alpha B^\circ$. Indeed, for any $x \in C_n$ we may write $x := x_n + t(h_{x_n} + p)$ for some $t \in]0, \varepsilon_0]$ and some $p \in \frac{\delta}{2} B^\circ$ which ensures by (2.24) that

$$\begin{aligned} \|x - x_0\| &= \|x_n + t(h_{x_n} + p) - x_0\| \le \|x_n - x_0\| + t\|h_{x_n} + p\| \\ &< \beta + \varepsilon_0 \left(\|h_0\| + \frac{\delta}{2} + \frac{\delta}{2} \right) = \varepsilon_0 \frac{\delta}{2} + \varepsilon_0 (\|h_0\| + \delta) \\ &= \varepsilon_0 \left(\|h_0\| + \frac{3\delta}{2} \right) \le \frac{\delta}{2(\|h_0\| + 2\delta)} \left(\|h_0\| + \frac{3\delta}{2} \right) < \delta = \alpha. \end{aligned}$$

Hence, we can apply the integrability result for locally Lipschitz regular functions, see Correa and Jofre [50] and Theorem 4.1 in Thibault and Zagrodny [161], to obtain that there exists some real constant c_n such that

$$f(x) = g(x) + c_n, \quad \forall x \in C_n, \quad \forall n \ge N_1.$$

From (2.22) we have that $x_n + \varepsilon_0 h_{x_n} = x_0 + \varepsilon_0 h_0 \in C_n$, for any $n \ge N_1$, hence

$$f(x_0 + \varepsilon_0 h_0) = g(x_0 + \varepsilon_0 h_0) + c_n, \quad \forall n \ge N_1,$$

and, in particular because $g(x_0 + \varepsilon_0 h_0)$ is finite according to (2.14), the value of c_n does not depend on n for $n \geq N_1$, say $c_n = c = f(x_0 + \varepsilon_0 h_0) - g(x_0 + \varepsilon_0 h_0)$, for all $n \geq N_1$, i.e.,

$$f(x) = g(x) + c, \quad \forall x \in C_n, \quad \forall n \geq N_1. \tag{2.25}$$

Now we proceed to prove the equality $f(v) = g(v) + c$.

First, let us show that $f(x_n) = g(x_n) + c$ for $n \geq N_1$. Observing by the definition of C_n that for some vectors h_0 one may have $x_n \notin C_n$, we begin by verifying that $f(x_n + th_{x_n}) \underset{t\downarrow 0}{\longrightarrow} f(x_n)$.

As f is lower semicontinuous, we always have that $f(x_n) \leq \liminf_{t\downarrow 0} f(x_n + th_{x_n})$. Further, for $t \in]0, \varepsilon_0]$, Lemma 2.2.4 gives that $x_n + th_{x_n} \in \operatorname{dom} f$ and $f(x_n + th_{x_n}) - f(x_n) \leq Kt$. Hence, $\limsup_{t\downarrow 0} f(x_n + th_{x_n}) \leq f(x_n)$ and the claim is proved.

Analogously, using the strictly directionally Lipschitz property of g at x_0 with respect to h_0 along with its lower semicontinuity, it is not difficult to see that $g(x_n + th_{x_n}) \underset{t\downarrow 0}{\longrightarrow} g(x_n)$, and then, reminding that $f(x_n)$ and $g(x_n)$ are finite because of $x_n \in \operatorname{dom} \partial f \subset \operatorname{dom} \partial g$, we may conclude that

$$(f - g)(x_n) = \lim_{t\downarrow 0}(f - g)(x_n + th_{x_n}) = c, \quad \forall n \geq N_1. \tag{2.26}$$

As $x_n \underset{n\to\infty}{\longrightarrow} v$, the lower semicontinuity of g implies that

$$g(v) \leq \liminf_{n\to\infty} g(x_n) = \lim_{n\to\infty} f(x_n) - c = f(v) - c,$$

in particular, $v \in \operatorname{dom} g$, hence $v \in \operatorname{dom} g \cap (x_0 + \beta B^\circ)$. The continuity of g relative to its domain ensures that $g(v) = \lim_{n\to\infty} g(x_n)$. By (2.26) and the fact that $f(x_n) \underset{n\to\infty}{\longrightarrow} f(v)$ we obtain

$$g(v) = \lim_{n\to\infty} g(x_n) = \lim_{n\to\infty} [f(x_n) - c] = f(v) - c,$$

i.e. $f(v) = g(v) + c$.

Hence we have proved that

$$f(v) = g(v) + c, \quad \forall v \in \operatorname{dom} f \cap (x_0 + \beta B^\circ),$$

and in the same time we obtain via that equality

$$\operatorname{dom} f \cap (x_0 + \beta B^\circ) \subset \operatorname{dom} g \cap (x_0 + \beta B^\circ).$$

To finish the proof, it remains only to establish the opposite inclusion of the domains. Take arbitrary $u \in \operatorname{dom} g \cap (x_0 + \beta B^\circ)$ and set $C := u +]0, \varepsilon_0] \left(h_u + \frac{\delta}{2} B^\circ\right)$. Note that for any $x \in \operatorname{dom} f \cap (x_0 + \beta B^\circ) \neq \emptyset$, the point $x + \varepsilon_0 h_x \in \operatorname{dom} f$ by Lemma 2.2.4 and, moreover

$$x + \varepsilon_0 h_x = x_0 + \varepsilon_0 h_0 = u + \varepsilon_0 h_u. \tag{2.27}$$

This ensures that $\operatorname{dom} f \cap C \neq \emptyset$. The assumptions of Theorem 4.1 in Thibault and Zagrodny [161] hold for f, g and C, and we apply it to conclude that

$$f(x) = g(x) + c, \quad \forall x \in C.$$

The constant is still $c = f(x_0 + \varepsilon_0 h_0) - g(x_0 + \varepsilon_0 h_0)$ because by (2.27) one has $x_0 + \varepsilon_0 h_0 \in C$. Observe that for any $t \in]0, \varepsilon_0]$ the points $u + th_u \in C \subset \operatorname{dom} g$, where the last inclusion holds because of (2.14). Using the lower semicontinuity of f at u, and the continuity of g with respect to its domain, we obtain that

$$f(u) \leq \liminf_{t \downarrow 0} f(u + th_u) = \liminf_{t \downarrow 0} g(u + th_u) + c = g(u) + c,$$

hence, $u \in \operatorname{dom} f$. The proof is then complete. □

Note that the continuity assumption of the restriction of g on $V \cap \operatorname{dom} g$ (where V is some neighbourhood of x_0) is crucial. It suffices to consider the function g from $\mathbb{R}$ into $\mathbb{R}$ with $g(x) = 0$ if $x \geq 0$ and $g(x) = 1$ if $x < 0$ and the function f from $\mathbb{R}$ into $\mathbb{R}$ with $f(x) = 0$ if $x \geq 0$ and $f(x) = 2$ if $x < 0$. We have $\partial f(x) \subset \partial g(x)$ for all $x \in \mathbb{R}$ but the functions f and g are not equal near 0 up to a constant.

In the same way, it is easily seen that the lower semicontinuity assumption of f is also essential.

2.3 Integrability of subdifferentials of certain bivariate functions

The aim of this section is to develop the subject in studying integrability properties of bivariate functions defined on a product of two Banach spaces. The problem in this setting is interesting from the point of view that the product structure allows us to introduce different concepts of continuity and regularity of the function and study how they rely to the integrability.

In Subsection 2.3.1 we introduce and discuss two concepts of regularity for a bivariate function.

The integrability of subdifferentials of locally Lipschitz bivariate functions for which some regularity concept holds is proved in Subsection 2.3.2.

In the following Subsection 2.3.3 we define and study two notions of directional Lipschitzness for a bivariate function.

These notions are used in Subsection 2.3.4 to establish the local integrability of certain directionally Lipschitz bivariate functions.

Throughout the section X, Y are real Banach spaces with *open unit balls* B_X° and B_Y° and topological duals X^* and Y^*, respectively. The product space $Z = X \times Y$ equipped with the norm $\|(x, y)\| = \max\{\|x\|, \|y\|\}$ is a Banach space whose *open unit ball* is denoted by $B^\circ := B_X^\circ \times B_Y^\circ$. For a function $\varphi : X \to \mathbb{R} \cup \{-\infty, +\infty\}$ we denote the effective domain by $\operatorname{dom} \varphi := \{x \in X : |\varphi(x)| < +\infty\}$.

Recall also that, for a function $\varphi : X \to \mathbb{R} \cup \{-\infty, +\infty\}$ with $x \in \operatorname{dom} \varphi$, the *lower Dini derivative* in the direction $h \in X$ is defined by

$$d^- \varphi(x; h) := \liminf_{\substack{h' \to h \\ t \downarrow 0}} t^{-1}[\varphi(x + th') - \varphi(x)].$$

Similarly one defines $d^+\varphi(x; h) := -d^-(-\varphi)(x; h)$ and $d^\downarrow\varphi(x; h) := -d^\uparrow(-\varphi)(x; h)$.

Recall that the function $\varphi : X \to \mathbb{R} \cup \{-\infty, +\infty\}$ is said to be *upper* (resp. *lower*) *regular* at x in the direction h when

$$d^- \varphi(x; h) = d^\uparrow \varphi(x; h) \quad \text{(resp. } d^+\varphi(x; h) = d^\downarrow \varphi(x; h)),$$

and it is said to be *upper* (resp. *lower*) *regular* at x when it is so in any direction.

When φ is locally Lipschitz near x, it is easy to see that the upper regularity in the direction h corresponds to the existence of $d\varphi(x; h) := \lim_{t \downarrow 0} t^{-1}[\varphi(x+th) - \varphi(x)]$ and to the equality $d\varphi(x; h) = d^\uparrow \varphi(x; h)$.

We will also use the notation

$$d^\circ \varphi(x; h) := \limsup_{\substack{(x',\alpha) \downarrow_\varphi (x,\varphi(x)) \\ h' \to h,\ t \downarrow 0}} t^{-1}[\varphi(x' + th') - \alpha].$$

Obviously, when φ is Lipschitz near x one has

$$d^\uparrow \varphi(x; h) = d^\circ \varphi(x; h) = \limsup_{\substack{x' \to x \\ t \downarrow 0}} t^{-1}[\varphi(x' + th) - \varphi(x')].$$

It is trivial that all above definitions hold for a bivariate function $g : X \times Y \to \mathbb{R} \cup \{-\infty, +\infty\}$ considered as a function of one variable defined on the Banach space $Z = X \times Y$. To simplify the notations, for $(x, y) \in \operatorname{dom} g$ we set

$$d^- g(x, y; h, k) := d^- g((x, y); (h, k)),$$

as well as,

$$d_1^- g(x,y;h) := d^- g(\cdot,y)(x;h) \text{ and } d_2^- g(x,y;k) := d^- g(x,\cdot)(y;k),$$

where $g(\cdot,y)$ denotes the function $x' \to g(x',y)$ and $g(x,\cdot)$ denotes the function $y' \to g(x,y')$. One defines similarly $d^\uparrow g(x,y;h,k)$, $d_1^\uparrow g(x,y;h)$, $d_2^\uparrow g(x,y;k)$, etc.

We work with a general *presubdifferential operator* ∂, that associates with each function $g : X \to \mathbb{R} \cup \{-\infty,+\infty\}$ and with each point $z = (x,y) \in Z$ a subset $\partial g(z)$ of Z^*, that we will call a *joint subdifferential* of g at z, and for which the four properties from Definition 2.1.1 hold, i.e., a subdifferential of g when it is considered as a function of one "joint" variable z. Also, we suppose that the joint subdifferential is included in the Clarke subdifferential, i.e., $\partial g(x,y) \subset \partial_C g(x,y)$ for any bivariate function $g : X \times Y \to \mathbb{R} \cup \{-\infty,+\infty\}$ and any $(x,y) \in X \times Y$. We will also consider the *partial subdifferentials* of the function g, namely, $\partial_1 g(x,y) := \partial_1 g(\cdot,y)(x)$ and $\partial_2 g(x,y) := \partial_2 g(x,\cdot)(y)$, where $\partial_1 \varphi$ (resp. $\partial_2 \varphi$) denotes some subdifferential in the sense of Definition 2.1.1 for functions φ defined on X (resp. on Y). The subdifferentials ∂_1 and ∂_2 are also assumed to be included in the Clarke subdifferential.

2.3.1 Concepts of regularity

For a bivariate function g defined on a product Banach space $Z = X \times Y$ we consider two regularity concepts.

Definition 2.3.1. (Correa and Thibault [54]) The function $g : X \times Y \to \mathbb{R} \cup \{-\infty,+\infty\}$ is said to be *upper-upper regular* (resp. *upper-lower regular*) at $(x,y) \in \operatorname{dom} g$ in the direction $(h,k) \in X \times Y$ if
(i) $d^- g(x,y;h,0) = d^\uparrow g(x,y;h,0)$ and
(ii) $d^- g(x,y;0,k) = d^\uparrow g(x,y;0,k)$ (resp. $d^+ g(x,y;0,k) = d^\downarrow g(x,y;0,k)$).
The function g is said to be *upper-upper regular* (resp. *upper-lower regular*) at $(x,y) \in \operatorname{dom} g$ if it is so in any direction $(h,k) \in X \times Y$.

It is a direct consequence from the definition that the class of upper-upper regular bivariate functions contains all upper regular bivariate functions (considered as functions of one variable (x,y)).

Definition 2.3.2. The function $g : X \times Y \to \mathbb{R} \cup \{-\infty,+\infty\}$ is said to be *separately upper-upper regular* (resp. *separately upper-lower regular*) at $(x,y) \in \operatorname{dom} g$ in the directions $h \in X$ and $k \in Y$ if

(i) $g(\cdot, y)$ is upper regular at x in the direction $h \in X$ (i.e. $d_1^- g(x, y; h) = d_1^{\uparrow} g(x, y; h)$) and
(ii) $g(x, \cdot)$ is upper regular at y in the direction $k \in Y$ (i.e. $d_2^- g(x, y; k) = d_2^{\uparrow} g(x, y; k)$) (resp. $g(x, \cdot)$ is lower regular at y in the direction $k \in Y$ (i.e. $d_2^+ g(x, y; k) = d_2^{\downarrow} g(x, y; k)$)).
The function g is said to be *separately upper-upper regular* (resp. *separately upper-lower regular*) at $(x, y) \in \operatorname{dom} g$ if it is so in any directions $h \in X$ and $k \in Y$.

Below we consider only the upper-upper case of regularity in order to simplify the presentation. As it is stated in Subsection 2.3.4, the upper-lower case can be dealt in a similar way.

It is not difficult to see that when the function g is locally Lipschitz, the upper-upper regularity implies the separate upper-upper regularity. Indeed, suppose that g is locally Lipschitz and upper-upper regular at (x, y) with respect to (h, k). Then $g(\cdot, y)$ is locally Lipschitz and

$$d_1^- g(x, y; h) = d^- g(x, y; h, 0) = d^{\circ} g(x, y; h, 0) \geq d_1^{\circ} g(x, y; h) \geq d_1^- g(x, y; h).$$

Hence, $d_1^- g(x, y; h) = d_1^{\circ} g(x, y; h) = d_1^{\uparrow} g(x, y; h)$. Analogously, one has $d_2^- g(x, y; k) = d_2^{\circ} g(x, y; k) = d_2^{\uparrow} g(x, y; k)$ and, hence, g is separately upper-upper regular at (x, y) with respect to h and k.

Any continuous convex-convex function $g : X \times Y \to \mathbb{R}$ is upper-upper regular (see Correa and Thibault [54, Proposition 2.2]) and then, the above observation implies that continuous convex-convex functions are also separately upper-upper regular.

A locally Lipschitz function may be separately upper-upper regular at some point without being upper-upper regular at that point as the following example demonstrates.

Example 2.3.3. The locally Lipschitz function $g : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ defined by

$$g(x, y) := \min\{|x|, |y|\}$$

is separately upper-upper regular at $(0, 0)$ in any directions h and k, but it is not upper-upper regular at $(0, 0)$, since it is not upper-upper regular at $(0, 0)$ in the direction $(h, k) = (1, 1)$.

Outside the case of a locally Lipschitz function, the upper-upper regularity may not imply the separate upper-upper regularity, as the following example shows.

Example 2.3.4. The function $g : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ defined by

$$g(x, y) := \begin{cases} x^2 \sin \frac{1}{x}, & \text{if } y = 0,\ x > 0; \\ -x^2, & \text{otherwise} \end{cases}$$

is continuous and upper-upper regular at $(0,0)$, but it is not separately upper-upper regular at $(0,0)$, since it is not separately upper-upper regular at $(0,0)$ in directions $h=1$ and $k=1$.

2.3.2 Integrability of subdifferentials of certain locally Lipschitz bivariate functions

In this section we establish the integrability of subdifferentials of locally Lipschitz bivariate functions $g : X \times Y \to \mathbb{R}$ for which the introduced concepts of regularity in the previous subsection hold.

First, we prove the integrability of the partial subdifferentials of a separate upper-upper regular locally Lipschitz bivariate function.

Theorem 2.3.5. Let $g : X \times Y \to \mathbb{R}$ be a locally Lipschitz function which is separately upper-upper regular and let $C \subset X \times Y$ be an open convex set.

Then for any function $f : X \times Y \to \mathbb{R} \cup \{+\infty\}$ which is separately lower semicontinuous with $\operatorname{dom} f \cap C \neq \emptyset$ the condition on partial subdifferentials

$$\text{(PS)} \qquad \partial_1 f(x,y) \subset \partial_1 g(x,y) \text{ and } \partial_2 f(x,y) \subset \partial_2 g(x,y), \quad \forall (x,y) \in C$$

implies

$$f(x,y) = g(x,y) + \text{const}, \quad \forall (x,y) \in C.$$

Proof. Take any $(x_0, y_0) \in C \cap \operatorname{dom} f$ which is a non-empty set by assumption, and consider the functions $f(\cdot, y_0)$ and $g(\cdot, y_0)$ on the set $U(y_0) := \{x \in X : (x, y_0) \in C\}$.

The function $g(\cdot, y_0) : X \to \mathbb{R}$ is locally Lipschitz and regular and the set $U(y_0)$ is open and convex in X. The function $f(\cdot, y_0) : X \to \mathbb{R} \cup \{+\infty\}$ is lower semicontinuous, such that $\operatorname{dom} f(\cdot, y_0) \cap U(y_0) \neq \emptyset$ (since $x_0 \in U(y_0)$) and by (PS) we have $\partial_1 f(x, y_0) \subset \partial_1 g(x, y_0)$, for all $x \in U(y_0)$. Hence, the conditions of Theorem 4.1 in Thibault and Zagrodny [161] are satisfied and we apply it to $f(\cdot, y_0)$, $g(\cdot, y_0)$ and $U(y_0)$ to deduce that there exists a finite constant $c(y_0)$ such that

$$\text{(2.28)} \qquad \begin{aligned} &f(x, y_0) = g(x, y_0) + c(y_0), \ \forall x \in U(y_0) \text{ and} \\ &\operatorname{dom} f(\cdot, y_0) \cap U(y_0) \equiv \operatorname{dom} g(\cdot, y_0) \cap U(y_0) \equiv U(y_0). \end{aligned}$$

Now, consider $f(x_0, \cdot)$ and $g(x_0, \cdot)$ on $U(x_0) := \{y \in Y : (x_0, y) \in C\}$. The function $g(x_0, \cdot) : Y \to \mathbb{R}$ is regular and locally Lipschitz and the set $U(x_0)$ is open and convex in Y. The function $f(x_0, \cdot) : Y \to \mathbb{R} \cup \{+\infty\}$ is lower semicontinuous with $\operatorname{dom} f(x_0, \cdot) \cap U(x_0) \neq \emptyset$ (since $y_0 \in U(x_0)$). Moreover, from (PS) we have

$\partial_2 f(x_0, y) \subset \partial_2 g(x_0, y)$, for all $y \in U(x_0)$. We apply again Theorem 4.1 in Thibault and Zagrodny [161] to $f(x_0, \cdot)$, $g(x_0, \cdot)$ and $U(x_0)$ to obtain a finite constant $d(x_0)$ such that

$$\begin{aligned} &f(x_0, y) = g(x_0, y) + d(x_0), \ \forall y \in U(x_0) \text{ and} \\ &\operatorname{dom} f(x_0, \cdot) \cap U(x_0) \equiv \operatorname{dom} g(x_0, \cdot) \cap U(x_0) \equiv U(x_0). \end{aligned} \tag{2.29}$$

We claim that for arbitrary $(x, y) \in C$ we have

$$(f - g)(x, y) = (f - g)(x_0, y_0), \tag{2.30}$$

and succeeding in proving the claim the proof will be complete.

To establish the claim, let us fix arbitrary $(x, y) \in C$ with $(x, y) \neq (x_0, y_0)$.

There exists $\delta > 0$ such that the inclusion

$$\bigcup \{(x' + \delta B_X^\circ) \times (y' + \delta B_Y^\circ) : (x', y') \in \text{ the segment } [(x_0, y_0), (x, y)]\} \subset C$$

holds, which means in particular that

$$x' + \delta B_X^\circ \subset U(y') \text{ and } y' + \delta B_Y^\circ \subset U(x'), \text{ whenever } (x', y') \in [(x_0, y_0), (x, y)]. \tag{2.31}$$

Put $u := x - x_0$, $v := y - y_0$, and $w := (u, v) \in X \times Y$. Choose $0 < t < \min\left\{\frac{\delta}{\|w\|}, 1\right\}$ such that $M := t^{-1} \in \mathbb{N}$, recalling that $\|w\| := \max\{\|u\|, \|v\|\}$.

First, we will show that

$$(f - g)(x_0, y_0) = (f - g)(x_0 + tu, y_0 + tv) = (f - g)((x_0, y_0) + t(u, v)). \tag{2.32}$$

To this end, observe that $x_0 \in U(y_0)$ and $x_0 + tu \in U(y_0)$. The latter is a consequence of (2.31), because (x_0, y_0) is in the segment $[(x, y), (x_0, y_0)]$ and $t\|u\| < \delta$. Using (2.28) we obtain that

$$c(y_0) = (f - g)(x_0, y_0) = (f - g)(x_0 + tu, y_0). \tag{2.33}$$

Similarly, we have that $y_0 + tv \in U(x_0 + tu)$ and $y_0 \in U(x_0 + tu)$ as a consequence of (2.31), because $(x_0 + tu, y_0 + tv)$ is in the segment $[(x, y), (x_0, y_0)]$ and $t\|v\| < \delta$. Note also by (2.33) that $(x_0 + tu, y_0) \in C \cap \operatorname{dom} f$. Then we may apply (2.29) to obtain that

$$d(x_0 + tu) = (f - g)(x_0 + tu, y_0) = (f - g)(x_0 + tu, y_0 + tv),$$

and combining with (2.33) we get (2.32).

Repeating the considerations, after M steps we obtain (2.30). The proof is then complete. □

Theorem 2.3.5 extends to the product of arbitrary Banach spaces a result of Wu and Ye (see [168, Corollary 3.16] where some restrictions on the Banach spaces are required).

Investigating the integrability of the joint subdifferential of a locally Lipschitz upper-upper regular bivariate function, one can prove as in Correa and Thibault [54] the following result.

Theorem 2.3.6 (Correa and Thibault [54, Proposition 3.7]). Let $g : X \times Y \to \mathbb{R}$ be a locally Lipschitz function which is upper-upper regular and let $C \subset X \times Y$ be an open convex set.

Then for any lower semicontinuous function $f : X \times Y \to \mathbb{R} \cup \{+\infty\}$ with $\operatorname{dom} f \cap C \neq \emptyset$ the condition on the joint subdifferentials

$$\text{(JS)} \qquad \partial f(x,y) \subset \partial g(x,y), \quad \forall (x,y) \in C$$

implies

$$f(x,y) = g(x,y) + \text{const}, \quad \forall (x,y) \in C.$$

It is easy to see that, if in Theorem 2.3.6 one replaces the condition (JS) on joint subdifferentials by the condition (PS) on partial subdifferentials, then its conclusion still holds. This is because the upper-upper regular locally Lipschitz function g is also separately upper-upper regular and one may invoke Theorem 2.3.5 to conclude. Hence, for a locally Lipschitz upper-upper regular bivariate function the integrability result holds when the inclusion assumption is made on the partial subdifferentials, as well as when it is made on the joint subdifferentials.

Let us note that Theorem 2.3.5 and Theorem 2.3.6 hold for convex-convex continuous functions.

2.3.3 Concepts of directional Lipschitzness

In this section we will consider two notions of directional Lipschitzness for a bivariate function. First is the classical definition stated for a bivariate function, namely:

Definition 2.3.7 (Rockafellar [150]). The lower semicontinuous function $g : X \times Y \to \mathbb{R} \cup \{+\infty\}$ is said to be *directionally Lipschitz* at $(x_0, y_0) \in \operatorname{dom} g$ with respect to a vector $(h_0, k_0) \in X \times Y$ if there exist constants $K \in \mathbb{R}$, $\varepsilon > 0$, $\delta > 0$ such that:

$$\begin{aligned} &t^{-1}[g(x+th, y+tk) - g(x,y)] \leq K, \\ &\text{for all } t \in]0,\varepsilon],\ (h,k) \in (h_0,k_0) + \delta B^\circ, \\ &\text{and } (x,y) \in (x_0,y_0) + \delta B^\circ \text{ with } |g(x,y) - g(x_0,y_0)| \leq \varepsilon. \end{aligned} \tag{2.34}$$

It is not difficult to see that g is Lipschitz around (x_0, y_0) exactly when it is directionally Lipschitz at (x_0, y_0) with respect to $(h_0, k_0) = (0, 0)$. Also, it is easy to see that in the case when the lower semicontinuous function g considered in the above definition is convex, or continuous relative to its domain (i.e. for any $(x_0, y_0) \in \operatorname{dom} g$ and any $\eta > 0$ there exists $\gamma > 0$ such that $|(g(x, y) - g(x_0, y_0)| < \eta$ for any $(x, y) \in ((x_0, y_0) + \gamma B^\circ) \cap \operatorname{dom} g$), then (2.34) is equivalent to the existence of constants $K \in \mathbb{R}$, $\varepsilon > 0$, $\delta > 0$ such that

$$\begin{aligned}&t^{-1}[g(x + th, y + tk) - g(x, y)] \leq K,\\&\text{for all } t \in]0, \varepsilon],\ (h, k) \in (h_0, k_0) + \delta B^\circ,\\&\text{and}\ \ (x, y) \in ((x_0, y_0) + \delta B^\circ) \cap \operatorname{dom} g.\end{aligned} \tag{2.35}$$

The lower semicontinuous function $g : X \times Y \to \mathbb{R} \cup \{+\infty\}$ is said to be *strictly directionally Lipschitz* at $(x_0, y_0) \in \operatorname{dom} g$ with respect to $(h_0, k_0) \in X \times Y$ if it is directionally Lipschitz at (x_0, y_0) with respect to (h_0, k_0) with some constants K, ε, δ satisfying (2.34) and, moreover (see Jofre and Thibault [97])

$$\begin{aligned}&g \text{ is locally Lipschitz on any drop } (x, y) +]0, \varepsilon]((h_0, k_0) + \delta B^\circ),\\&\text{where } (x, y) \in (x_0, y_0) + \delta B \text{ with } |g(x, y) - g(x_0, y_0)| \leq \varepsilon.\end{aligned} \tag{2.36}$$

If the lower semicontinuous function g in the latter definition is also supposed to be continuous relative to its domain, or convex, then (2.36) can be replaced by:

$$\begin{aligned}&g \text{ is locally Lipschitz on any drop } (x, y) +]0, \varepsilon]((h_0, k_0) + \delta B^\circ),\\&\text{where } (x, y) \in ((x_0, y_0) + \delta B^\circ) \cap \operatorname{dom} g.\end{aligned} \tag{2.37}$$

Now we will introduce another concept of directional Lipschitzness for a bivariate function which is inspired by the product structure of the space.

Definition 2.3.8. The separately lower semicontinuous function $g : X \times Y \to \mathbb{R} \cup \{+\infty\}$ will be said to be *separately directionally Lipschitz* at $(x_0, y_0) \in \operatorname{dom} g$ with respect to vectors $h_0 \in X$ and $k_0 \in Y$, if there exist constants $K \in \mathbb{R}$, $\varepsilon > 0$, $\delta > 0$ such that:

$$\begin{aligned}&t^{-1}[g(x + th, y) - g(x, y)] \leq K, \text{ and } t^{-1}[g(x, y + tk) - g(x, y)] \leq K,\\&\text{for all } t \in]0, \varepsilon],\ (h, k) \in (h_0, k_0) + \delta B^\circ, \text{ and}\\&(x, y) \in (x_0, y_0) + \delta B^\circ \text{ with } |g(x, y) - g(x_0, y_0)| \leq \varepsilon.\end{aligned} \tag{2.38}$$

Note that Definition 2.3.8 may be easily adapted to the function that is upper semicontinuous on some of the variables in assuming that the corresponding inequality in (2.38) holds for $-g$ instead of g.

It is clear that when the separately lower semicontinuous function g, considered in the above definition is supposed to be also continuous relative to its domain, then the separate directionally Lipschitz property (2.38) is equivalent to the existence of constants $K \in \mathbb{R}$, $\varepsilon > 0$, $\delta > 0$ such that

$$\begin{aligned} &t^{-1}[g(x+th,y)-g(x,y)] \le K, \text{ and } t^{-1}[g(x,y+tk)-g(x,y)] \le K, \\ &\text{for all } t \in]0,\varepsilon], \ (h,k) \in (h_0,k_0)+\delta B^\circ, \ \text{ and} \\ &(x,y) \in ((x_0,y_0)+\delta B^\circ) \cap \operatorname{dom} g. \end{aligned} \tag{2.39}$$

To compare so introduced definitions of directional Lipschitzness, let us observe that for a lower semicontinuous function $g : X \times Y \to \mathbb{R} \cup \{+\infty\}$ the directionally Lipschitz property at (x_0, y_0) with respect to (h_0, k_0), i.e. (2.34) is equivalent to

$$\begin{aligned} &d^\circ g(x,y;h,k) \le K, \\ &\forall (h,k) \in (h_0,k_0)+\delta B^\circ, \ \forall (x,y) \in (x_0,y_0)+\delta B^\circ \text{ with } |g(x,y)-g(x_0,y_0)| \le \varepsilon, \end{aligned}$$

while the separate directionally Lipschitz property at (x_0, y_0) with respect to h_0 and k_0, i.e., (2.38) is equivalent to

$$\begin{aligned} &d_1^\circ g(x,y;h) \le K, \text{ and } d_2^\circ g(x,y;k) \le K, \\ &\forall (h,k) \in (h_0,k_0)+\delta B^\circ, \ \forall (x,y) \in (x_0,y_0)+\delta B^\circ \text{ with } |g(x,y)-g(x_0,y_0)| \le \varepsilon. \end{aligned}$$

The following example (due to Rockafellar [148] and considered there in another context) provides in $\mathbb{R} \times \mathbb{R}$ a lower semicontinuous function which is directionally Lipschitz at the origin with respect to (h_0, k_0) but it is not separately directionally Lipschitz with respect to h_0 and k_0.

Example 2.3.9. The lower semicontinuous function

$$f(x,y) := \begin{cases} \frac{x^2}{y} - y, & \text{if } y > 0, \\ 0, & \text{if } x = y = 0, \\ +\infty, & \text{otherwise} \end{cases}$$

is directionally Lipschitz at $(0, 0)$ with respect to $(h_0, k_0) = (1, 1)$ but it is not separately directionally Lipschitz with respect to $h_0 = 1$ and $k_0 = 1$, as it is infinite on the positive x-ray.

Further, the separate directionally Lipschitz property is stronger than the non-separate one in such a way that, for a lower semicontinuous function, it implies its directional Lipschitzness. Here we give the proof for a lower semicontinuous function, which is continuous relative to its domain, as this is the case that we will need later.

Lemma 2.3.10. Let $g : X \times Y \to \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous function, which is continuous relative to its domain. Assume that g is separately directionally

Lipschitz at $(x_0, y_0) \in \operatorname{dom} g$ with respect to $h_0 \in X$ and $k_0 \in Y$ with constants K, ε, δ satisfying (2.39) and fix any $m \geq \max\{\|h_0\|, \|k_0\|\}$. Then g is directionally Lipschitz at (x_0, y_0) with respect to (h_0, k_0) with constants $2K, \tilde{\varepsilon}, \tilde{\delta}$ satisfying (2.35) for any positive constants $\tilde{\varepsilon}$ and $\tilde{\delta}$ such that $\tilde{\delta} + \tilde{\varepsilon}(m + \tilde{\delta}) < \delta$ and $\tilde{\varepsilon} \leq \varepsilon$.

Proof. Take any K, ε, δ satisfying (2.39) and any positive $\tilde{\varepsilon}$ and $\tilde{\delta}$ such that $\tilde{\delta} + \tilde{\varepsilon}(m + \tilde{\delta}) < \delta$ and $\tilde{\varepsilon} \leq \varepsilon$. Then, for any $(x, y) \in ((x_0, y_0) + \tilde{\delta} B^\circ) \cap \operatorname{dom} g$, any $t \in]0, \tilde{\varepsilon}[$, and any $k \in k_0 + \tilde{\delta} B_Y^\circ$ one has from (2.39) that

$$t^{-1}[g(x, y + tk) - g(x, y)] \leq K. \tag{2.40}$$

Hence,

$$(x, y + tk) \in \operatorname{dom} g \cap ((x_0, y_0) + (\tilde{\delta} + \tilde{\varepsilon}(m + \tilde{\delta}))B^\circ) \subset \operatorname{dom} g \cap ((x_0, y_0) + \delta B^\circ) \tag{2.41}$$

by the choice of $\tilde{\delta}$ and $\tilde{\varepsilon}$.

Again from (2.39) we have for any $h \in h_0 + \tilde{\delta} B_X^\circ$ that

$$t^{-1}[g(x + th, y + tk) - g(x, y + tk)] \leq K. \tag{2.42}$$

Combining (2.40) and (2.42) one gets

$$\begin{aligned} &t^{-1}[g(x + th, y + tk) - g(x, y)] \leq 2K, \\ &\text{for all } t \in]0, \tilde{\varepsilon}],\ (h, k) \in (h_0, k_0) + \tilde{\delta} B^\circ, \quad \text{and} \quad (x, y) \in ((x_0, y_0) + \tilde{\delta} B^\circ) \cap \operatorname{dom} g. \end{aligned}$$

So, according to the definition, g is directionally Lipschitz at (x_0, y_0) with respect to (h_0, k_0) with constants $2K$, $\tilde{\varepsilon}$, and $\tilde{\delta}$. □

Further, we show that the inclusion of partial subdifferentials entails the separate directional Lipschitzness. The proof of this result follows the main steps of the proof of Lemma 2.2.4 and for this reason it is omitted.

Lemma 2.3.11. Let $g : X \times Y \to \mathbb{R} \cup \{+\infty\}$ be a separately lower semicontinuous function, which is continuous relative to its domain. Suppose that g is separately directionally Lipschitz at $(x_0, y_0) \in \operatorname{dom} g$ with respect to $h_0 \in X$ and $k_0 \in Y$ with constants K, ε, δ satisfying (2.39) and fix any $m \geq \max\{\|h_0\|, \|k_0\|\}$.

Let $f : X \times Y \to \mathbb{R} \cup \{+\infty\}$ be a separately lower semicontinuous function such that

$$\partial_1 f(x, y) \subset \partial_1 g(x, y) \text{ and } \partial_2 f(x, y) \subset \partial_2 g(x, y), \quad \forall (x, y) \in (x_0, y_0) + \delta B^\circ. \tag{PS}$$

Then for any positive constants $\bar{\delta}$ and $\bar{\varepsilon}$ such that $\bar{\delta} + \bar{\varepsilon}(m + \bar{\delta}) < \delta$ and $\bar{\varepsilon} \leq \varepsilon$ with $\operatorname{dom} f \cap ((x_0, y_0) + \bar{\delta} B^\circ) \neq \emptyset$ one has

$$\begin{aligned} &t^{-1}[f(x + th, y) - f(x, y)] \leq K, \quad \text{and} \quad t^{-1}[f(x, y + tk) - f(x, y)] \leq K, \\ &\text{for all } t \in]0, \bar{\varepsilon}],\ (h, k) \in (h_0, k_0) + \bar{\delta} B^\circ, \text{ and } (x, y) \in ((x_0, y_0) + \bar{\delta} B^\circ) \cap \operatorname{dom} f. \end{aligned} \tag{2.43}$$

Now we put together Lemma 2.3.10 and Lemma 2.3.11 to obtain

Proposition 2.3.12. Let $g : X \times Y \to \mathbb{R} \cup \{+\infty\}$ be a separately lower semicontinuous function, which is continuous relative to its domain. Suppose that g is separately directionally Lipschitz at $(x_0, y_0) \in \operatorname{dom} g$ with respect to $h_0 \in X$ and $k_0 \in Y$ with constants K, ε, δ satisfying (2.39) and fix any $m \geq \max\{\|h_0\|, \|k_0\|\}$. Let $f : X \times Y \to \mathbb{R} \cup \{+\infty\}$ be a separately lower semicontinuous function such that

$$\text{(PS)} \qquad \partial_1 f(x,y) \subset \partial_1 g(x,y) \text{ and } \partial_2 f(x,y) \subset \partial_2 g(x,y), \quad \forall (x,y) \in (x_0,y_0) + \delta B^\circ.$$

Then for any positive constants δ_0 and ε_0 such that $\delta_0 + \varepsilon_0(m + \delta_0) < \delta/2$ and $\varepsilon_0 \leq \min\left\{\varepsilon, \frac{\delta}{2(m+\delta)}\right\}$ with $\operatorname{dom} f \cap ((x_0,y_0) + \delta_0 B^\circ) \neq \emptyset$ one has

$$(2.44) \qquad \begin{array}{l} t^{-1}[f(x+th, y+tk) - f(x,y)] \leq 2K, \\ \text{for all } t \in]0, \varepsilon_0], \ (h,k) \in (h_0,k_0) + \delta_0 B^\circ, \text{ and} \\ (x,y) \in ((x_0,y_0) + \delta_0 B^\circ) \cap \operatorname{dom} f. \end{array}$$

In particular, $(x,y) + [0,\varepsilon_0]((h_0,k_0) + \delta_0 B^\circ) \subset \operatorname{dom} f$, whenever $(x,y) \in \operatorname{dom} f \cap ((x_0,y_0) + \delta_0 B^\circ)$.

Proof. We have from Lemma 2.3.11 that, for $\overline{\delta} = \delta/2$ and $\overline{\varepsilon} = \min\left\{\varepsilon, \frac{\overline{\delta}}{m+2\overline{\delta}}\right\}$, property (2.43) holds for f. Further, exactly as in the proof of Lemma 2.3.10, but now working with f and with $\overline{\delta}$ and $\overline{\varepsilon}$ in the place of δ and ε, respectively, we can take any positive ε_0 and δ_0 such that $\delta_0 + \varepsilon_0(m + \delta_0) < \overline{\delta} = \delta/2$ and $\varepsilon_0 \leq \overline{\varepsilon} = \min\left\{\varepsilon, \frac{\delta}{2(m+\delta)}\right\}$ (for example $\delta_0 = \delta/4$ and $\varepsilon_0 = \min\left\{\varepsilon, \frac{\delta}{4(m+\delta/2)}\right\}$) for which (2.44) holds. □

2.3.4 Local integrability of subdifferentials of certain directionally Lipschitz bivariate functions

The first result concerns the local integrability of the joint subdifferential of a bivariate function for which non-separate regularity and non-separate directional Lipschitzness hold.

Theorem 2.3.13. Let $g : X \times Y \to \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous function, which is continuous relative to its domain. Suppose that g is upper-upper regular and strictly directionally Lipschitz at $(x_0, y_0) \in \operatorname{dom} g$ with respect to $(h_0, k_0) \in X \times Y$.

Then there exist some constants $\alpha > 0$ and $\beta \in]0, \alpha[$ such that for any lower semicontinuous function $f : X \times Y \to \mathbb{R} \cup \{+\infty\}$ with $\operatorname{dom} f \cap ((x_0,y_0) + \beta B^\circ) \neq \emptyset$ the condition

$$\text{(JS)} \qquad \partial f(x,y) \subset \partial g(x,y), \quad \forall (x,y) \in (x_0,y_0) + \alpha B^\circ$$

implies

$$f(x,y) = g(x,y) + \text{const}, \quad \forall (x,y) \in (x_0,y_0) + \beta B^\circ.$$

Proof. The proof may be derived by following the steps of the proof of Theorem 2.2.6 concerning the case of a directionally Lipschitz function of one variable, so we made the choice to omit it. We need only to observe first, that Lemma 2.2.4 (that is essentially used in the above mentioned proof of Theorem 2.2.6) holds in the product Banach space $Z = X \times Y$, and second, that we can apply Theorem 2.3.6 working on any drop that appears in the proof of Theorem 2.2.6. □

The second result establishes the local integrability of the partial subdifferentials of a bivariate function for which the separate regularity and the separate directional Lipschitzness hold.

We need first to introduce the strict version of the separately directionally Lipschitz property as we have already done for the directionally Lipschitz property.

Definition 2.3.14. Let $g : X \times Y \to \mathbb{R} \cup \{+\infty\}$ be a separately lower semicontinuous function that is continuous relative to its domain. We say that g is *strictly separately directionally Lipschitz* at $(x_0, y_0) \in \operatorname{dom} g$ with respect to $h_0 \in X$ and $k_0 \in Y$ if it is separately directionally Lipschitz at (x_0, y_0) with respect to h_0 and k_0 with some constants K, ε, δ satisfying (2.39) and, moreover, for each $(x, y) \in ((x_0, y_0) + \delta B^\circ) \cap \operatorname{dom} g$ the function $g(\cdot, y)$ is locally Lipschitz on the drop $x+]0, \varepsilon](h_0 + \delta B_X^\circ)$, and the function $g(x, \cdot)$ is locally Lipschitz on the drop $y+]0, \varepsilon](k_0 + \delta B_Y^\circ)$.

We can now prove the result.

Theorem 2.3.15. Let $g : X \times Y \to \mathbb{R} \cup \{+\infty\}$ be a separately lower semicontinuous function which is continuous relative to its domain. Suppose also that g is separately upper-upper regular and strictly separately directionally Lipschitz at $(x_0, y_0) \in \operatorname{dom} g$ with respect to $(h_0, k_0) \in X \times Y$.

Then there exist some constants $\alpha > 0$ and $\beta \in]0, \alpha[$ such that for any function $f : X \times Y \to \mathbb{R} \cup \{+\infty\}$ that is separately lower semicontinuous with $\operatorname{dom} f \cap ((x_0, y_0) + \beta B^\circ) \neq \emptyset$ the condition

$$\text{(PS)} \qquad \partial_1 f(x,y) \subset \partial_1 g(x,y) \text{ and } \partial_2 f(x,y) \subset \partial_2 g(x,y), \quad \forall (x,y) \in (x_0,y_0) + \alpha B^\circ$$

implies

$$f(x,y) = g(x,y) + \text{const}, \quad \forall (x,y) \in (x_0,y_0) + \beta B^\circ.$$

Proof. Let g be strictly separately directionally Lipschitz at $(x_0, y_0) \in \operatorname{dom} g$ with respect to $h_0 \in X$ and $k_0 \in Y$ with constants K, ε and δ satisfying (2.39). Fix any $m > \max\{\|h_0\|, \|k_0\|\}$. Further, fix $\tilde{\delta} \in]0, \delta/4[$ such that

$$\frac{\tilde{\delta}}{m + 2\tilde{\delta}} < \min\left\{\varepsilon, \frac{\delta}{2(m+\delta)}\right\}$$

and

$$0 < 4\tilde{\delta} < \min\left\{\frac{(\delta - 2\tilde{\delta})^2}{4(m + 2(\delta - 2\tilde{\delta}))}, \frac{\varepsilon(\delta - 2\tilde{\delta})}{2}\right\}.$$

Set $\overline{\alpha} := \delta$, $\overline{\beta} := \frac{\tilde{\delta}^2}{m+2\tilde{\delta}}$ and $\beta := \min\left\{\frac{(\delta-2\tilde{\delta})^2}{4(m+2(\delta-2\tilde{\delta}))}, \frac{\varepsilon(\delta-2\tilde{\delta})}{2}\right\}$. Note that $\overline{\beta} < \tilde{\delta}$ and $4\tilde{\delta} < \beta$.

Let f be any function satisfying the assumptions of the theorem with $\overline{\alpha}$ and $\overline{\beta}$. We need several steps.

I. Set $\Delta := ((x_0, y_0) + 2\tilde{\delta}B^\circ) \cap \operatorname{dom} f \neq \emptyset$. Fix arbitrary $y' \in \operatorname{proj}_Y \Delta$. By density of $\operatorname{dom} \partial_1 f(., y')$ in $\operatorname{dom} f(., y')$ (see Lemma 2.2.5), one may find $\tilde{x}$ with $(\tilde{x}, y') \in ((x_0, y_0) + 2\tilde{\delta}B^\circ)$ such that $(\tilde{x}, y') \in \operatorname{dom} \partial_1 f \subset \operatorname{dom} \partial_1 g \subset \operatorname{dom} g$.

The function $g(\cdot, y')$ is lower semicontinuous, regular, continuous relative to its domain and according to (2.39) and Definition 2.3.14 it is also strictly directionally Lipschitz at $\tilde{x}$ with respect to h_0 with constants K, ε and $\delta - 2\tilde{\delta}$.

The function $f(\cdot, y')$ is lower semicontinuous and such that $\operatorname{dom} f(\cdot, y') \cap (\tilde{x} + \beta B_X^\circ) \neq \emptyset$ (because it contains $\tilde{x}$), and $\partial_1 f(x, y') \subset \partial_1 g(x, y')$, for all $x \in \tilde{x} + (\delta - 2\tilde{\delta})B_X^\circ \subset x_0 + \overline{\alpha}B_X^\circ$.

Hence, according to Theorem 2.2.6 applied with $\alpha = \overline{\alpha}$ and β, one has that there exists a finite constant $c(y')$ such that

$$\begin{gathered} f(x, y') - g(x, y') = c(y'), \quad \forall x \in (\tilde{x} + \beta B_X^\circ) \cap \operatorname{dom} f(\cdot, y'), \text{ and} \\ \operatorname{dom} f(\cdot, y') \cap (\tilde{x} + \beta B_X^\circ) \equiv \operatorname{dom} g(\cdot, y') \cap (\tilde{x} + \beta B_X^\circ). \end{gathered}$$

Further, since $\tilde{x} + \beta B_X^\circ \supset x_0 + 2\tilde{\delta}B_X^\circ$, it follows that

$$\begin{gathered} f(x, y') - g(x, y') = c(y'), \quad \forall x \in (x_0 + 2\tilde{\delta}B_X^\circ) \cap \operatorname{dom} f(\cdot, y'), \text{ and} \\ \operatorname{dom} f(\cdot, y') \cap (x_0 + 2\tilde{\delta}B_X^\circ) \equiv \operatorname{dom} g(\cdot, y') \cap (x_0 + 2\tilde{\delta}B_X^\circ). \end{gathered} \tag{2.45}$$

II. Consider now $x' \in \operatorname{proj}_X \Delta$. One may find $\tilde{y}$ with $(x', \tilde{y}) \in ((x_0, y_0) + 2\tilde{\delta}B^\circ)$ such that $(x', \tilde{y}) \in \operatorname{dom} \partial_2 f \subset \operatorname{dom} \partial_2 g \subset \operatorname{dom} g$.

The function $g(x', \cdot)$ is lower semicontinuous, regular, continuous relative to its domain and according to (2.39) and Definition 2.3.14 it is also strictly directionally Lipschitz at $\tilde{y}$ with respect to k_0 with constants K, ε and $\delta - 2\tilde{\delta}$.

The function $f(x', \cdot)$ is lower semicontinuous and such that $\operatorname{dom} f(x', \cdot) \cap (\tilde{y} + \beta B_Y^\circ) \neq \emptyset$ (because it contains $\tilde{y}$), and $\partial_2 f(x', y) \subset \partial_2 g(x', y)$, for all $y \in \tilde{y} + (\delta - 2\tilde{\delta})B_Y^\circ \subset y_0 + \overline{\alpha} B_Y^\circ$.

Again applying Theorem 2.2.6 with the same α and β, one obtains the existence of a finite constant $d(x')$ such that

$$f(x', y) - g(x', y) = d(x'), \quad \forall y \in (\tilde{y} + \beta B_X^\circ) \cap \operatorname{dom} f(x', \cdot), \text{ and}$$
$$\operatorname{dom} f(x', \cdot) \cap (\tilde{y} + \beta B_Y^\circ) \equiv \operatorname{dom} g(x', \cdot) \cap (\tilde{y} + \beta B_Y^\circ).$$

Since $\tilde{y} + \beta B_Y^\circ \supset y_0 + 2\tilde{\delta} B_Y^\circ$, it follows that

$$\begin{aligned} f(x', y) - g(x', y) = d(x'), \quad \forall y \in (y_0 + 2\tilde{\delta} B_Y^\circ) \cap \operatorname{dom} f(x', \cdot), \text{ and} \\ \operatorname{dom} f(x', \cdot) \cap (y_0 + 2\tilde{\delta} B_Y^\circ) \equiv \operatorname{dom} g(x', \cdot) \cap (y_0 + 2\tilde{\delta} B_Y^\circ). \end{aligned} \tag{2.46}$$

Combining (2.45) and (2.46) we obtain for any $(x', y') \in \operatorname{dom} f \cap ((x_0, y_0) + 2\tilde{\delta} B^\circ)$ that

$$c(y') = f(x', y') - g(x', y') = d(x'). \tag{2.47}$$

III. We proceed to showing that $c(y')$ does not depend on y', as well as $d(x')$ does not depend on x', and that they are equal. Let us mention that the latter is not a direct consequence of (2.47) because the set $\operatorname{dom} f \cap ((x_0, y_0) + 2\tilde{\delta} B^\circ)$ may not contain a subset that is a cartesian product.

Observe first that for any $(x, y) \in (x_0, y_0) + \overline{\beta} B^\circ$ one may find $(h_x, k_y) \in (h_0, k_0) + \tilde{\delta} B^\circ$ such that

$$(x, y) + \tilde{\varepsilon}(h_x, k_y) = (x_0, y_0) + \tilde{\varepsilon}(h_0, k_0), \text{ where } \tilde{\varepsilon} := \frac{\overline{\beta}}{\tilde{\delta}} = \frac{\tilde{\delta}}{m + 2\tilde{\delta}}.$$

Indeed, to obtain h_x we set $h_x := h_0 + \frac{x_0 - x}{\tilde{\varepsilon}}$. It is easy to see that $h_x \in h_0 + \tilde{\delta} B_X^\circ$, since

$$\|h_x - h_0\| = \frac{\|x_0 - x\|}{\tilde{\varepsilon}} < \frac{\overline{\beta}}{\tilde{\varepsilon}} = \tilde{\delta}.$$

To obtain k_y we proceed analogously.

Put $u_0 := x_0 + \tilde{\varepsilon} h_0$ and $v_0 := y_0 + \tilde{\varepsilon} k_0$. Further, observe that for $\tilde{\delta}$ and $\tilde{\varepsilon}$ the conclusion of Lemma 2.3.10 holds, since its conditions $\tilde{\delta} + \tilde{\varepsilon}(m + \tilde{\delta}) \leq \tilde{\delta} + \frac{\tilde{\delta}(m+\tilde{\delta})}{m+2\tilde{\delta}} < 2\tilde{\delta} < \delta$ and $\tilde{\varepsilon} \leq \varepsilon$ are satisfied. Hence, from the proof of Lemma 2.3.10 we have that for any $(x, y) \in \operatorname{dom} g \cap ((x_0, y_0) + \tilde{\delta} B)$ and any $(h, k) \in (h_0, k_0) + \tilde{\delta} B$ the points $(x, y + \tilde{\varepsilon} k)$ and $(x + \tilde{\varepsilon} h, y)$ are in $\operatorname{dom} g \cap ((x_0, y_0) + 2\tilde{\delta} B)$.

Having in mind that $\overline{\beta} < \tilde{\delta}$ we combine with the preceding to obtain

$$(x, y) \in \operatorname{dom} g \cap ((x_0, y_0) + \overline{\beta} B^\circ) \Rightarrow (x, y + \tilde{\varepsilon} k_y),\ (x + \tilde{\varepsilon} h_x, y) \in \operatorname{dom} g \cap ((x_0, y_0) + 2\tilde{\delta} B^\circ), \text{ i.e.,}$$

(2.48) $(x,y) \in \operatorname{dom} g \cap ((x_0,y_0)+\overline{\beta}B^\circ) \Rightarrow (x,v_0),\ (u_0,y) \in \operatorname{dom} g \cap ((x_0,y_0)+2\tilde{\delta}B^\circ)$.

Now, take any point $y' \in \operatorname{proj}_Y[\operatorname{dom} f \cap ((x_0,y_0)+\overline{\beta}B^\circ)] \subset \operatorname{proj}_Y \Delta$. Then there exists $x' \in x_0 + \overline{\beta}B_X^\circ$ such that $(x',y') \in \operatorname{dom} f \cap ((x_0,y_0)+\overline{\beta}B^\circ)$ and hence $(x',y') \in \operatorname{dom} g \cap ((x_0,y_0)+\overline{\beta}B^\circ)$ where the latter holds according to (2.45). From (2.48) the point $(u_0,y') \in \operatorname{dom} g \cap ((x_0,y_0)+2\tilde{\delta}B^\circ)$ and hence $(u_0,y') \in \operatorname{dom} f \cap ((x_0,y_0)+2\tilde{\delta}B^\circ)$, the latter again by (2.45).

Since we have established that $u_0 \in \operatorname{dom} f(\cdot,y')$, we use (2.45) to obtain

$$f(u_0,y') - g(u_0,y') = c(y'). \tag{2.49}$$

Further, from the fact that $u_0 \in \operatorname{proj}_X \Delta$ and from (2.46) because $y' \in \operatorname{dom} f(u_0,\cdot)$, we have that

$$f(u_0,y') - g(u_0,y') = d(u_0). \tag{2.50}$$

Combining (2.49) and (2.50) we obtain

$$c(y') = d(u_0) \text{ for any } y' \in \operatorname{proj}_Y[\operatorname{dom} f \cap ((x_0,y_0)+\overline{\beta}B^\circ)].$$

Analogously, one gets

$$d(x') = c(v_0) \text{ for any } x' \in \operatorname{proj}_X[\operatorname{dom} f \cap ((x_0,y_0)+\overline{\beta}B^\circ)].$$

Since for arbitrary $(x',y') \in \operatorname{dom} f \cap ((x_0,y_0)+\overline{\beta}B^\circ)$ from (2.47) we have $c(y') = d(x')$, those two constants are equal, i.e., $d(u_0) = c(v_0) =: c$. Finally, we obtain that there exists a finite constant c such that

$$\begin{gathered} f(x,y) - g(x,y) = c, \quad \forall (x,y) \in \operatorname{dom} f \cap ((x_0,y_0)+\overline{\beta}B^\circ), \text{ and} \\ \operatorname{dom} f \cap ((x_0,y_0)+\overline{\beta}B^\circ) \subset \operatorname{dom} g \cap ((x_0,y_0)+\overline{\beta}B^\circ). \end{gathered} \tag{2.51}$$

IV. To finish the proof it remains only to establish the opposite inclusion of the domains, i.e., $\operatorname{dom} g \cap ((x_0,y_0)+\overline{\beta}B^\circ) \subset \operatorname{dom} f$.

To this end, let us take arbitrary $(u,v) \in \operatorname{dom} g \cap ((x_0,y_0)+\overline{\beta}B^\circ)$. First we will show that $(u,v_0) \in \operatorname{dom} f$. Consider the drop $U := (u,v_0) +]0,\tilde{\varepsilon}]\left(\{h_0 + \tilde{\delta}B_X^\circ\} \times \{0\}\right) =: U_1 \times \{v_0\}$ and observe that $u_0 \in U_1$ because

$$u_0 = x_0 + \tilde{\varepsilon}h_0 = u + \tilde{\varepsilon}\left(h_0 + \frac{x_0-u}{\tilde{\varepsilon}}\right) \text{ and } \frac{\|x_0-u\|}{\tilde{\varepsilon}} < \frac{\overline{\beta}}{\tilde{\varepsilon}} = \tilde{\delta}.$$

We claim that $U \subset ((x_0,y_0)+\overline{\alpha}B^\circ)$. Indeed, one has

$$\overline{\beta} + \tilde{\varepsilon}(m+\tilde{\delta}) = \tilde{\varepsilon}\tilde{\delta} + \tilde{\varepsilon}(m+\tilde{\delta}) = \tilde{\varepsilon}(m+2\tilde{\delta}) = \frac{\tilde{\delta}}{m+2\tilde{\delta}}(m+2\tilde{\delta}) = \tilde{\delta} < \overline{\alpha},$$

and if we take any $(x, y) \in U$ we use the above estimation to obtain

$$\|x - x_0\| \leq \|x - u\| + \|u - x_0\| \leq \tilde{\varepsilon}(\|h_0\| + \tilde{\delta}) + \overline{\beta} < \tilde{\varepsilon}(m + \tilde{\delta}) + \overline{\beta} < \overline{\alpha},$$

$$\|y - y_0\| = \|v_0 - y_0\| = \tilde{\varepsilon}\|k_0\| < \tilde{\varepsilon}m < \overline{\alpha}$$

and the claim is proved. Both last inequalities allow us to invoke (PS) which implies

$$\partial_1 f(x, v_0) \subset \partial_1 g(x, v_0), \quad \forall x \in U_1. \tag{2.52}$$

We claim that $(u_0, v_0) \in \operatorname{dom} f$. This is a consequence of Proposition 2.3.12. Indeed, as $\tilde{\delta} + \tilde{\varepsilon}(m + \tilde{\delta}) < 2\tilde{\delta} < \delta/2$ and $\tilde{\varepsilon} \leq \min\left\{\varepsilon, \frac{\delta}{2(m+\delta)}\right\}$, we may take $\delta_0 = \tilde{\delta}$ and $\varepsilon_0 = \tilde{\varepsilon}$ in Proposition 2.3.12. Using the inequality $\overline{\beta} < \tilde{\delta}$ and fixing any $(x', y') \in \operatorname{dom} f \cap ((x_0, y_0) + \overline{\beta} B^\circ)$ we then get

$$(x' + \tilde{\varepsilon} h_{x'}, y' + \tilde{\varepsilon} k_{y'}) = (u_0, v_0) \in \operatorname{dom} f, \tag{2.53}$$

since $(h_{x'}, k_{y'}) \in (h_0, k_0) + \delta_0 B^\circ$.

We observe now that we have $(u, v_0) \in \operatorname{dom} g$ because of (2.48). So, the choice of $\tilde{\delta}$ and $\tilde{\varepsilon}$ together with (2.39) imply that $U_1 \subset \operatorname{dom} g(\cdot, v_0)$. Moreover, from the strict directionally Lipschitz property, $g(\cdot, v_0)$ is locally Lipschitz and regular on U_1. As $u_0 \in U_1$ (see the beginning of the present step), we have by (2.53) that $u_0 \in U_1 \cap \operatorname{dom} f(\cdot, v_0)$, hence the latter is non-empty. Taking all those facts and (2.52) into account, we can apply Theorem 4.1 in Thibault and Zagrodny [161] with $f(\cdot, v_0)$, $g(\cdot, v_0)$ and U_1 to obtain that there exists a finite constant c_1 such that

$$f(x, v_0) = g(x, v_0) + c_1, \quad \forall x \in U_1. \tag{2.54}$$

In fact one can show that $c_1 = c$. For $t \in]0, \tilde{\varepsilon}]$ we have that the point $(u + th_0, v_0) \in U \cap \operatorname{dom} g$. We use the lower semicontinuity of f at (u, v_0) with respect to the first variable and the continuity of g relative to its domain, as well as (2.54), to get

$$f(u, v_0) \leq \liminf_{t \downarrow 0} f(u + th_0, v_0) = \liminf_{t \downarrow 0} g(u + th_0, v_0) + c_1 = g(u, v_0) + c_1.$$

Hence, $(u, v_0) \in \operatorname{dom} f$.

Now, consider the drop $V := (u, v) +]0, \tilde{\varepsilon}] \left(\{0\} \times \{k_0 + \tilde{\delta} B_Y^\circ\}\right) =: \{u\} \times V_2$. We have that $V \subset ((x_0, y_0) + \overline{\alpha} B^\circ)$ exactly by the same reasons as above, so (PS) implies that $\partial_2 f(u, y) \subset \partial_2 g(u, y)$ for all $y \in V_2$. We have already showed that $v_0 \in \operatorname{dom} f(u, \cdot) \cap V_2$. Again the choice of $\tilde{\delta}$ and $\tilde{\varepsilon}$ and (2.39) ensure that $V_2 \subset \operatorname{dom} g(u, \cdot)$. Moreover, $g(u, \cdot)$ is locally Lipschitz and regular on V_2. We apply Theorem 4.1 in Thibault and

Zagrodny [161] with $f(u,.)$, $g(u,.)$ and V_2 to obtain that there exists a finite constant c_2 such that

$$f(u,y) = g(u,y) + c_2, \quad \forall y \in V_2.$$

For any $t \in]0,\tilde{\varepsilon}]$ the point $v + tk_0 \in V_2$, the function f is lower semicontinuous at (u,v) with respect to the second variable, and the function g is continuous relative to its domain, hence

$$f(u,v) \leq \liminf_{t\downarrow 0} f(u, v+tk_0) = \liminf_{t\downarrow 0} g(u, v+tk_0) + c_2 = g(u,v) + c_2.$$

This means that $(u,v) \in \operatorname{dom} f$ and the proof is then complete. □

Theorem 2.3.15 implies for example the local integrability of any extended real valued convex-convex function g, which is lower semicontinuous on each variable and continuous relative to its domain, in a neighbourhood of a point (x_0,y_0) such that $\operatorname{int}[\operatorname{dom} g(\cdot,y_0) \times \operatorname{dom} g(x_0,\cdot)] \neq \emptyset$.

In order to deal together with the separate regularity and the joint (non-separate) directionally Lipschitz property, we prove the following result.

Theorem 2.3.16. Let $g : X \times Y \to \mathbb{R} \cup \{+\infty\}$ be a separately lower semicontinuous function which is continuous relative to its domain. Suppose also that g is separately upper-upper regular and strictly directionally Lipschitz at $(x_0,y_0) \in \operatorname{dom} g$ with respect to $(h_0,0) \in X \times Y$.

Then there exist some constants $\alpha > 0$ and $\beta \in]0,\alpha[$ such that for any function $f : X \times Y \to \mathbb{R} \cup \{+\infty\}$ that is separately lower semicontinuous with $\operatorname{dom} f \cap ((x_0,y_0) + \beta B^\circ) \neq \emptyset$ the condition

(PS) $\quad \partial_1 f(x,y) \subset \partial_1 g(x,y)$, and $\partial_2 f(x,y) \subset \partial_2 g(x,y), \quad \forall (x,y) \in (x_0,y_0) + \alpha B^\circ$

implies

$$f(x,y) = g(x,y) + \text{const}, \quad \forall (x,y) \in (x_0,y_0) + \beta B^\circ.$$

Proof. We will give only a sketch of the proof because most of the techniques applied can be found in the proof of Theorem 2.2.6.

Fix constants K, ε, δ such that properties (2.35) and (2.37) hold for g with respect to $(h_0,0)$. Set $\delta_0 := \delta/2$, $\varepsilon_0 := \min\left\{\varepsilon, \frac{\delta}{2(\|h_0\|+2\delta)}\right\}$ and put $\alpha := \delta$, $\beta := \varepsilon_0\delta_0$.

Let f be any function satisfying the assumptions of the theorem with such α and β. According to Lemma 2.2.4, it is not difficult to see that the properties of g and the condition (PS) ensure that if $(x,y) \in \operatorname{dom} f \cap ((x_0,y_0) + \delta_0 B^\circ)$ then

$$(x,y) + [0,\varepsilon_0](h,0) \subset \operatorname{dom} f, \quad \forall h \in h_0 + \delta_0 B_X^\circ. \tag{2.55}$$

For any $(x, y) \in X \times Y$ we will put $h_x = h_0 + \frac{x_0 - x}{\varepsilon_0}$ and $k_y = \frac{y_0 - y}{\varepsilon_0}$. Fix any $(u, v) \in \operatorname{dom} f \cap ((x_0, y_0) + \beta B^\circ)$. Take (see Lemma 2.2.5) for each integer n some x_n in X such that

$$(x_n, v) \in ((x_0, y_0) + \beta B^\circ) \cap \operatorname{dom} \partial_1 f \subset \operatorname{dom} \partial_1 g \subset \operatorname{dom} g,$$

and such that $x_n \to u$ and $f(x_n, v) \to f(x, v)$. The drop

$$C_n := (x_n, v) +]0, \varepsilon_0] \left(\{h_{x_n} + \delta_0 B_X^\circ\} \times \{k_v + \delta_0 B_Y^\circ\}\right)$$

contains the point $(x_0 + \varepsilon_0 h_0, y_0) = (x_n + \varepsilon_0 h_{x_n}, v + \varepsilon_0 k_v)$ and, since $\|k_v\| < \delta_0$, it also contains the half open segment $(x_n, v) +]0, \varepsilon_0](h_{x_n}, 0)$. Further, since $\beta < \delta_0$ by (2.55) one has

$$(x_n, v) +]0, \varepsilon_0](h_{x_n}, 0) \subset \operatorname{dom} f.$$

On the open convex set $C_n \subset X \times Y$ the function g is locally Lipschitz and separately upper-upper regular. Thus, Theorem 2.3.5 entails that there exists a finite constant $c := (f - g)(x_0 + \varepsilon_0 h_0, y_0)$ such that

$$f(x, y) = g(x, y) + c, \quad \forall (x, y) \in C_n.$$

Now applying the techniques from the proof of Theorem 2.2.6 to the function $f(., v)$ we obtain that

$$f(u, v) = g(u, v) + c.$$

Therefore,

$$\begin{gathered} f(x, y) = g(x, y) + c, \quad \forall (x, y) \in \operatorname{dom} f \cap ((x_0, y_0) + \beta B^\circ), \text{ and} \\ \operatorname{dom} f \cap ((x_0, y_0) + \beta B^\circ) \subset \operatorname{dom} g \cap ((x_0, y_0) + \beta B^\circ). \end{gathered}$$

To establish the opposite inclusion of the domains it is enough to see that for any $(u, v) \in \operatorname{dom} g \cap ((x_0, y_0) + \beta B^\circ)$, the drop $C := (u, v) +]0, \varepsilon_0] \left((h_u + \delta_0 B_X^\circ) \times (k_v + \delta_0 B_Y^\circ)\right)$ contains the point $(x_0 + \varepsilon_0 h_0, y_0)$ and the latter belongs to $\operatorname{dom} f$, according to the definition of c. Hence, applying again Theorem 2.3.5 we obtain that $f(x, y) = g(x, y) + c$ for all $(x, y) \in C$, in particular for all $t \in]0, \varepsilon_0]$

$$f(u + t h_u, v) = g(u + t h_u, v) + c.$$

Exploring the lower semicontinuity of f with respect to the first variable and continuity of g on its domain, we obtain that $(u, v) \in \operatorname{dom} f$. The proof is then complete. □

Theorem 2.3.16 in particular implies the local integrability of any extended real valued convex-convex function g, which is lower semicontinuous with respect to each variable and continuous relative to its domain, around a point (x_0, y_0) such that there exists some $x \in X$ for which $(x, y_0) \in \operatorname{int} \operatorname{dom} g$.

To conclude, let us note that all integration results obtained in this section can be easily adapted for continuous bivariate functions, for which a type of upper-lower regularity assumption holds (for example, continuous convex-concave functions).

2.4 Partially ball weakly inf-compact saddle functions

In this section we study saddle functions $K : X \times Y \to [-\infty, +\infty]$ defined on a product of Banach spaces X and Y. Such functions are closely related to minimax problems. The main contribution in the study of saddle functions with values in $[-\infty, +\infty]$ is due to Rockafellar. In the case when X and Y are finite dimensional, their properties are investigated in detail in a number of works of Rockafellar (cf., e.g., [142, 145, 148]), McLinden [117, 118, 119], etc. Most of those properties are generalized to the case when X and Y are Banach spaces, one of which is reflexive, in subsequent papers of Rockafellar [146, 149], Gossez [79], etc. Our intention here is to extend their results to a class of saddle functions defined on a product of two arbitrary Banach spaces X and Y.

The properties of saddle functions on the Banach space $X \times Y$ are laid out in Subsection 2.4.1.

Subsection 2.4.2 is devoted to the study of subdifferentiability properties of saddle functions, and more precisely, the subdifferential properties of a proper closed saddle function $K : X \times Y \to [-\infty, +\infty]$ that we call *partially ball weakly inf-compact* (Definition 2.4.8) and write for short pbwc. We establish that the domain of the subdifferential ∂K is non-empty (Theorem 2.4.11). Moreover, in this setting the operator T_K associated with ∂K is maximal monotone (Theorem 2.4.14). For proper closed saddle functions K in a large subclass of pbwc saddle functions we show that the domain of ∂K is dense in the domain of K (Theorem 2.4.12).

In the final Subsection 2.4.3 we prove that for proper closed pbwc saddle functions K the subdifferential ∂K is integrable (Theorem 2.4.15 and Theorem 2.4.16).

We can conclude that there exist clear parallels between basic properties of proper closed convex functions defined on a Banach space X and those of proper closed pbwc saddle functions defined on a product Banach space $X \times Y$.

We work in a real Banach space $(X, \|\cdot\|)$ with topological *dual space* X^*. The dual of the Banach space X^* is called the *bidual* of X and it is denoted by X^{**}. Any element of X "gives" an element of X^{**} via the *canonical embedding* $\widehat{\ }: X \hookrightarrow X^{**}$ defined by

$$\langle x^*, \widehat{x} \rangle = \langle x, x^* \rangle \quad \text{for } x \in X \text{ and } x^* \in X^*.$$

Let us recall that two normed linear spaces are called *congruent* (relation denoted by $\cong$) if there exists a norm preserving isomorphism (called *a congruence*) between them. It is well-known (see Holmes [86]) that the canonical embedding is a congruence between X and its image $\widehat{X}$ in X^{**}, i.e., $X \cong \widehat{X}$. The image $\widehat{X}$ of X is a norm closed subspace of X^{**} and $\|\widehat{x}\| = \|x\|$. When $\widehat{X}$ coincides with X^{**}, the Banach space X is *reflexive*. Analogously, the Banach space X^* is congruent to $\widehat{X^*}$, i.e., $X^* \cong \widehat{X^*}$ via the canonical embedding $\widehat{\ }: X^* \hookrightarrow X^{***}$ defined by

$$\langle x^{**}, \widehat{x^*} \rangle = \langle x^*, x^{**} \rangle \quad \text{for } x^* \in X^* \text{ and } x^{**} \in X^{**}.$$

Moreover, $\left(\widehat{X}\right)^* \cong X^*$ (see Holmes [86, p. 123]).

A *convex function* f on X is an everywhere defined function with values in the extended real interval $[-\infty, +\infty]$ whose *epigraph* $\operatorname{epi} f := \{(x, r) \in X \times \mathbb{R} : f(x) \le r\}$ is a convex set in $X \times \mathbb{R}$. The *effective domain* of f is defined by $\operatorname{dom} f := \{x \in X : f(x) < +\infty\}$. If $f(x) > -\infty$ for all x and $f(x) < +\infty$ for at least one x, then f is said to be *proper*. Otherwise, f is said to be *improper*. The convex function f is said to be *closed* if it is proper and lower semicontinuous, or else, if it is identically $+\infty$ or $-\infty$. Through the subsection, unless otherwise be specified, the *closure* and *lower semicontinuity operations* will be taken with respect to the *norm topology*. Given any convex function f on X, there exists a greatest closed convex function majorized by f. This function is called the *closure* of f and is denoted by $\operatorname{cl} f$. It is clear that $\operatorname{cl} f \le f$ and f is closed exactly when $f = \operatorname{cl} f$. When f does not take the value $-\infty$, then

$$\operatorname{cl} f(x) = \liminf_{x' \to x} f(x'), \quad \text{for all } x \in X. \tag{2.56}$$

For any convex function f on X, the function $f^* : X^* \to [-\infty, +\infty]$ defined by

$$f^*(x^*) = \sup_{x \in X}\{\langle x, x^* \rangle - f(x)\}$$

is called the *conjugate* of f. The conjugate of a proper lower semicontinuous convex function f on X is a proper convex function on X^* which is lower semicontinuous with respect to the weak-star topology $w(X^*, X)$, as well as, to the norm topology of X^*. One defines the *biconjugate* of a convex function f on X as the conjugate of its conjugate function, i.e., it is the function $f^{**} : X^{**} \to [-\infty, +\infty]$ on the bidual space X^{**} defined by

$$f^{**}(x^{**}) = \sup_{x^* \in X^*} \{\langle x^*, x^{**}\rangle - f^*(x^*)\}.$$

The biconjugate of a proper convex lower semicontinuous function f is a proper convex function on X^{**} which is lower semicontinuous with respect to the $w(X^{**}, X^*)$ topology, as well as, to the norm topology of X^{**}. By Fenchel's duality result, for a convex function f one has

$$f^{**}(\widehat{x}) = \operatorname{cl} f(x) \text{ for all } x \in X. \tag{2.57}$$

The reader interested in the theory of conjugate convex functions could consult for example Brøndsted [35], Fenchel [74], Moreau [127, 129], Rockafellar [143, 148].

When f is convex and proper, and ∂f is its convex subdifferential, $\operatorname{dom} \partial f \subset \operatorname{dom} f$. We will often use the following well-known result of Brøndsted and Rockafellar (see [36]):

$$\begin{array}{l}\text{if the convex function } f : X \to \mathbb{R} \cup \{+\infty\} \text{ is proper} \\ \text{and lower semicontinuous, then } \operatorname{dom} \partial f \text{ is dense in } \operatorname{dom} f.\end{array} \tag{2.58}$$

Recall that the *range* of ∂f is the subset of X^* given by $\operatorname{Rge} \partial f := \bigcup_{x \in X} \partial f(x)$ while the *graph* of ∂f is the set $\operatorname{gph} \partial f := \{(x, x^*) \in X \times X^* : x^* \in \partial f(x)\}$.

It is well known (see for example Aubin and Ekeland [8], Moreau [129] and Rockafellar [148]) that for a proper convex function f the following are equivalent

$$x^* \in \partial f(x) \iff f(x) + f^*(x^*) = \langle x, x^*\rangle,$$

and any of those implies that $x \in \operatorname{dom} f$. If, in addition f is lower semicontinuous at x, then

$$x^* \in \partial f(x) \iff f(x) + f^*(x^*) = \langle x, x^*\rangle \iff \widehat{x} \in \partial f^*(x^*). \tag{2.59}$$

Following Rockafellar [148, Corollary 23.5.2], one derives that

$$\begin{array}{l}\text{if a proper convex function } f \text{ is subdifferentiable at } x \text{ then} \\ \operatorname{cl} f(x) = f(x) \text{ and } \partial(\operatorname{cl} f)(x) = \partial f(x).\end{array} \tag{2.60}$$

If X is reflexive and f is a proper lower semicontinuous convex function, from (2.59) it is clear that one can identify f^{**} with f and that ∂f^* is just the "inverse" of ∂f. In other words, $x \in \partial f^*(x^*)$, if and only if, $x^* \in \partial f(x)$. If X is not reflexive, the relationship between ∂f^* and ∂f is more complicated, but ∂f^* and ∂f still completely determine each other, according to the following result due to Rockafellar [147].

Theorem 2.4.1 (Rockafellar [147, Proposition 1]). Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be a proper lower semicontinuous convex function and let $x^* \in X^*$, $x^{**} \in X^{**}$. Then $x^{**} \in \partial f^*(x^*)$, if and only if, there exists a net $\{x^*_\alpha\}_{\alpha \in A}$ in X^* converging to x^* in the norm topology and a bounded net $\{\widehat{x_\alpha}\}_{\alpha \in A}$ in $\widehat{X}$ (with the same partially ordered index set A) converging to x^{**} in the $w(X^{**}, X^*)$ topology such that $x^*_\alpha \in \partial f(x_\alpha)$ for every $\alpha \in A$.

For any $r \in \mathbb{R}$, the *r-sublevel set* of the convex function f is the (possibly empty) set $\{f \leq r\} := \{x \in X : f(x) \leq r\}$. Obviously, it is a convex set and when f is lower semicontinuous, it is closed in the norm topology as well as in the weak topology of X. Let B_X denote the closed unit ball of the Banach space X.

We say that the function $f : X \to \mathbb{R} \cup \{+\infty\}$ is *ball weakly inf-compact* (*bwc* for short) if for any $r \in \mathbb{R}$ the sets $\mathrm{Lev}_{r,n}(f) := \{f \leq r\} \cap nB_X$ are $w(X, X^*)$ compact for any $n \in \mathbb{N}$. We make the convention that the empty set is weakly compact. Recall that the notion of inf-compactness was introduced by Moreau (see [129]).

From Theorem 2.4.1 we derive the following characterizations of a proper closed bwc convex function.

Theorem 2.4.2. Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be a proper lower semicontinuous convex function. Then the following are equivalent:
(a) f is bwc;
(b) the range of ∂f^* is a non-empty subset of $\widehat{X}$;
(c) $\mathrm{dom}\, f^{**} \subset \widehat{X}$.

Proof. (a) $\Rightarrow$ (b). Since f^* is a proper closed convex function, the range of ∂f^* is non-empty according to (2.58). Take any $x^{**} \in \partial f^*(x^*)$. Then by Theorem 2.4.1 we have that there exist a net $\{x^*_\alpha\}_{\alpha \in A}$ in X^* converging to x^* in the norm topology and a bounded net $\{\widehat{x_\alpha}\}_{\alpha \in A}$ in $\widehat{X}$ converging to x^{**} in the $w(X^{**}, X^*)$ topology such that $x^*_\alpha \in \partial f(x_\alpha)$ for every $\alpha \in A$. As f is proper we take any $x_0 \in \mathrm{dom}\, f$ and by the definition of the subdifferential we have

$$f(x_\alpha) \leq f(x_0) + \langle x_\alpha - x_0, x^*_\alpha \rangle, \quad \forall \alpha \in A.$$

If we set $r := 1 + f(x_0) + \langle x^*, x^{**} - \widehat{x_0} \rangle$, there exists some $\alpha_0 \in A$ such that $f(x_\alpha) \leq r$ for all $\alpha \geq \alpha_0$. Since the net $\{x_\alpha\}_{\alpha \in A}$ is norm bounded, for sufficiently large $n \in \mathbb{N}$, the points $x_\alpha \in \{f \leq r\} \cap nB_X$ for all $\alpha \geq \alpha_0$. The $w(X, X^*)$ compactness of the latter set ensures that the $w(X^{**}, X^*)$ closure of its embedding in X^{**} lies in $\widehat{X}$ (cf., e.g., Holmes [86, p. 149]). Hence, there exists some $x \in X$ such that $x^{**} = \widehat{x}$.

(b) $\Rightarrow$ (c). Suppose the contrary, i.e., there exists some $x^{**} \in \operatorname{dom} f^{**}$ such that $x^{**} \in X^{**} \setminus \widehat{X}$. Then there exists a norm open neighbourhood U of x^{**} in X^{**} such that $U \cap \widehat{X} = \emptyset$. By the norm density of the domain of the subdifferential of a proper closed convex function in its effective domain (see (2.58)) it follows the existence of some $x_0^{**} \in U \cap \operatorname{dom} \partial f^{**}$. Let $x_0^{***} \in \partial f^{**}(x_0^{**})$. By Theorem 2.4.1 we have that there exist a net $\{x_\alpha^{**}\}_{\alpha\in A}$ in X^{**} converging to x_0^{**} in the norm topology of X^{**} and a bounded net $\{\widehat{x_\alpha^*}\}_{\alpha\in A}$ in $\widehat{X^*}$ converging to x_0^{***} in the $w(X^{***}, X^{**})$ topology such that $x_\alpha^{**} \in \partial f^*(x_\alpha^*)$ for every $\alpha \in A$. From the norm convergence of $\{x_\alpha^{**}\}$ to x_0^{**} we have that x_α^{**} are eventually in U, in particular $x_\alpha^{**} \notin \widehat{X}$. But from (b) we have that $x_\alpha^{**} \in \widehat{X}$, which yields a contradiction. Hence, $\operatorname{dom} f^{**} \subset \widehat{X}$.

(c) $\Rightarrow$ (a). Let us fix $r \in \mathbb{R}$ and $n \in \mathbb{N}$ such that $\operatorname{Lev}_{r,n}(f)$ is a non-empty set, otherwise the claim is trivial. Take an arbitrary net $\{x_\alpha\}_{\alpha\in A} \subset \operatorname{Lev}_{r,n}(f)$. Since $\|x_\alpha\| = \|\widehat{x_\alpha}\|$ and $f(x_\alpha) = f^{**}(\widehat{x_\alpha})$ by (2.57), we have $\widehat{x_\alpha} \in S := \{x^{**} \in X^{**} : f^{**}(x^{**}) \leq r\} \cap nB_{X^{**}}$. The later set being obviously $w(X^{**}, X^*)$ compact one may extract from $\{\widehat{x_\alpha}\}_{\alpha\in A}$ a subnet $\{\widehat{x_{s(\gamma)}}\}_{\gamma\in\Gamma}$ that converges to some $x^{**} \in S$ in the $w(X^{**}, X^*)$ topology. The definition of S yields that $x^{**} \in \operatorname{dom} f^{**}$. By the assumption of (c) we have that $x^{**} = \widehat{x}$ for some $x \in X$, and $x^{**} \in nB_{X^{**}}$ implies that $x \in nB_X$. Obviously, $\{x_{s(\gamma)}\}$ tends to x in the $w(X, X^*)$ topology and $x \in \operatorname{Lev}_{r,n}(f)$ since $f(x) = f^{**}(\widehat{x})$ by (2.57). From the arbitrary net $\{x_\alpha\}_{\alpha\in A} \subset \operatorname{Lev}_{r,n}(f)$ we have shown how to obtain a subnet that is $w(X, X^*)$ convergent to an element of $\operatorname{Lev}_{r,n}(f)$. This means that the latter set is $w(X, X^*)$ compact. The proof is then complete. □

To finish, let us recall that a *concave function* g on X is an everywhere defined function with values in the extended real interval $[-\infty, +\infty]$ such that the function $(-g)$ is convex.

2.4.1 Saddle functions. Properties

We will work in a space Z, which is a product of two real Banach spaces X and Y, i.e., $Z = X \times Y$. Setting any reasonable norm on $X \times Y$ we have that Z is a Banach space (take for instance $\|(x, y)\| := \max\{\|x\|, \|y\|\}$) and its dual can be identified with $X^* \times Y^*$ using the pairing $\langle (x, y), (x^*, y^*)\rangle = \langle x, x^*\rangle + \langle y, y^*\rangle$.

A *saddle function* on Z is an everywhere defined function K with values in $[-\infty, +\infty]$ such that $K(\cdot, y)$ is convex on X for each $y \in Y$ and $K(x, \cdot)$ is concave on Y for each $x \in X$. We denote by $\operatorname{cl}_1 K$ the saddle function obtained by closing $K(x, y)$ as a convex function of x for each fixed y, i.e., $\operatorname{cl}_1 K(x, y) := \operatorname{cl} K(\cdot, y)(x)$. Similarly, we denote by $\operatorname{cl}_2 K$ the saddle function obtained by closing $(-K)(x, y)$ as a convex function of y for each fixed x and after that by taking its negative, i.e., $\mathrm{cl}_2 K(x, y) := -\mathrm{cl}(-K)(x, \cdot)(y)$. Clearly, $\mathrm{cl}_1 K \le K \le \mathrm{cl}_2 K$.

Two saddle functions K and L on Z are said to be *equivalent* (written $K \sim L$) if $\mathrm{cl}_1 K = \mathrm{cl}_1 L$ and $\mathrm{cl}_2 K = \mathrm{cl}_2 L$ (see Gossez [79] and Rockafellar [148, 149]). When $K \sim L$, we say that K and L belong to the same *equivalence class* and that K and L are *representatives* of the class. There exists another definition for the relation $K \sim L$ which is, of course, equivalent to the former. For the second one we need to introduce two more notions.

The function on $X \times Y^*$ obtained by taking the conjugate of $(-K)(x, y)$ in the second argument (its convex argument) when the first argument is fixed, i.e., $F(x, \cdot) := [-K(x, \cdot)]^*$, or

$$F(x, y^*) := \sup_{y \in Y} \{K(x, y) + \langle y, y^* \rangle\}, \tag{2.61}$$

will be called the *convex parent* of K (see Rockafellar [145]). It is a convex function over $X \times Y^*$. Dually, the *concave parent* of K is defined by $G(\cdot, y) := -[K(\cdot, y)]^*$, or

$$G(x^*, y) := \inf_{x \in X} \{K(x, y) - \langle x, x^* \rangle\}. \tag{2.62}$$

All parent functions that appear in the section are considered as functions of the *joint variable* belonging to the product Banach space, and their effective domains and subdifferentials are taken in this setting.

It is shown by Rockafellar in [148] that two saddle functions are equivalent exactly when they have the same parent functions. Here we give a proof for completeness and for the convenience of the reader.

Lemma 2.4.3. Let $K, L : X \times Y \to [-\infty, +\infty]$ be two saddle functions. Then $K \sim L$, if and only if, they have the same parent functions.

Proof. First, let us suppose that $K \sim L$, i.e., $\mathrm{cl}_1 K = \mathrm{cl}_1 L$ and $\mathrm{cl}_2 K = \mathrm{cl}_2 L$ according to the original definition. Denote by F the convex parent of K and by F' the convex parent of L. Then

$$F(x, y^*) := \sup_{y \in Y} \{K(x, y) + \langle y, y^* \rangle\} = \sup_{y \in Y} \{\mathrm{cl}_2 K(x, y) + \langle y, y^* \rangle\}, \tag{2.63}$$

where for the latter equality we use the fact that a convex function and its closure have the same conjugate function (see Moreau [129]). Hence,

$$F(x, y^*) = \sup_{y \in Y} \{\mathrm{cl}_2 K(x, y) + \langle y, y^* \rangle\} = \sup_{y \in Y} \{\mathrm{cl}_2 L(x, y) + \langle y, y^* \rangle\} = F'(x, y^*).$$

Analogously, for the concave parent G of K and the concave parent G' of L one obtains that $G(x^*, y) = G'(x^*, y)$.

Now, let us suppose that $F = F'$ and $G = G'$. From Fenchel duality (2.57) we have

$$-\mathrm{cl}_2 K(x, y) = \sup_{y^* \in Y^*} \{\langle y, y^* \rangle - F(x, y^*)\} = \sup_{y^* \in Y^*} \{\langle y, y^* \rangle - F'(x, y^*)\} = -\mathrm{cl}_2 L(x, y),$$

hence $\mathrm{cl}_2 K = \mathrm{cl}_2 L$. Analogously, $\mathrm{cl}_1 K = \mathrm{cl}_1 L$ and the proof is complete. □

If $\mathrm{cl}_1 K$ and $\mathrm{cl}_2 K$ are both equivalent to K (written $K \sim \mathrm{cl}_1 K \sim \mathrm{cl}_2 K$), then K is said to be a *closed* saddle function (see Gossez [79] and Rockafellar [148, 149]). A necessary and sufficient condition for K to be closed is

$$\mathrm{cl}_1 \mathrm{cl}_2 K = \mathrm{cl}_1 K \quad \text{and} \quad \mathrm{cl}_2 \mathrm{cl}_1 K = \mathrm{cl}_2 K. \tag{2.64}$$

It is easy to see that when K is a closed saddle function and L is a saddle function equivalent to K, i.e., $L \sim K$, then L is closed too.

The *effective domain* of a saddle function K (see Rockafellar [145, 146, 148]) is defined as the set $\mathrm{Dom}' K = C'_K \times D'_K$, where

$$C'_K = \{x \in X : K(x, y) < +\infty, \quad \forall y \in Y\},$$

$$D'_K = \{y \in Y : K(x, y) > -\infty, \quad \forall x \in X\}.$$

A basic disadvantage of this definition comes from the fact that $\mathrm{Dom}' K$ depends on the representative of the equivalence class of K, as it can be seen from an example due to Rockafellar (see Gossez [79]). Hence, one introduces the following more appropriate notion for the domain of a saddle function K that depends only on the equivalence class to which K belongs (and not on the representatives of the class). The *domain* of a saddle function K (see Gossez [79] and Rockafellar [149]) is defined by $\mathrm{Dom}\, K = C_K \times D_K$, where

$$C_K = \{x \in X : \mathrm{cl}_2 K(x, y) < +\infty, \quad \forall y \in Y\},$$

$$D_K = \{y \in Y : \mathrm{cl}_1 K(x, y) > -\infty, \quad \forall x \in X\}.$$

Clearly, $\mathrm{Dom}\, K \subset \mathrm{Dom}' K$. If K is closed, then $\mathrm{Dom}\, K$ is dense in $\mathrm{Dom}' K$, and $\mathrm{Dom}\, K = \mathrm{Dom}' \mathrm{cl}_1 K \cap \mathrm{Dom}' \mathrm{cl}_2 K$ (see Gossez [79, Proposition 1]). The saddle function K is said to be *proper* if $\mathrm{Dom}' K$ is a non-empty set. From the preceding, if K is closed, it is proper exactly when $\mathrm{Dom}\, K$ is a non-empty set.

Lemma 2.4.4. Let $K : X \times Y \to [-\infty, +\infty]$ be a saddle function. Then

$$\operatorname{Dom} K = \{(x, y) \in X \times Y : -\infty < \operatorname{cl}_1 K(x, y) \leq K(x, y) \leq \operatorname{cl}_2 K(x, y) < +\infty\}.$$

Proof. Let us denote

$$V := \{(x, y) \in X \times Y : -\infty < \operatorname{cl}_1 K(x, y) \leq K(x, y) \leq \operatorname{cl}_2 K(x, y) < +\infty\}.$$

The inclusion $V \subset \operatorname{Dom} K$ being obvious for $V = \emptyset$, we may suppose that $V \neq \emptyset$. Take any $(x_0, y_0) \in V$. Suppose that $y_0 \notin D_K$. Then there exists $\tilde{x} \in X$ such that $\operatorname{cl}_1 K(\tilde{x}, y_0) = -\infty$. Then, by the definition of the closure of a convex function, it follows that $\operatorname{cl}_1 K(x, y_0) = -\infty$ for all $x \in X$, which yields a contradiction, since $\operatorname{cl}_1 K(x_0, y_0)$ is finite. Hence, $y_0 \in D_K$. Analogously, one shows that $x_0 \in C_K$ and then $(x_0, y_0) \in \operatorname{dom} K$, so the inclusion $V \subset \operatorname{Dom} K$ is established.

The opposite inclusion, i.e., $\operatorname{Dom} K \subset V$, being obvious for $\operatorname{Dom} K = \emptyset$, we suppose it is non-empty and take any $(x_0, y_0) \in \operatorname{Dom} K$. Then

$$x_0 \in C_K \Longrightarrow \operatorname{cl}_2 K(x_0, y) < +\infty, \ \forall y \in Y \Longrightarrow \operatorname{cl}_2 K(x_0, y_0) < +\infty,$$

$$y_0 \in D_K \Longrightarrow \operatorname{cl}_1 K(x, y_0) > -\infty, \ \forall x \in X \Longrightarrow \operatorname{cl}_1 K(x_0, y_0) > -\infty,$$

hence $(x_0, y_0) \in V$. The proof is then complete. □

The following statement concerns a useful property of a proper closed saddle function.

Lemma 2.4.5. Let $K : X \times Y \to [-\infty, +\infty]$ be a proper closed saddle function. Then

(a) for any $y \in D_K$, the function $\operatorname{cl}_1 K(\cdot, y)$ is a proper lower semicontinuous convex function with $\operatorname{dom} \operatorname{cl}_1 K(\cdot, y) \supset C_K$ and, for any $y \notin D_K$, the function $\operatorname{cl}_1 K(\cdot, y) = -\infty$;

(b) for any $x \in C_K$, the function $(-\operatorname{cl}_2 K)(x, \cdot)$ is a proper lower semicontinuous convex function with $\operatorname{dom}(-\operatorname{cl}_2 K)(x, \cdot) \supset D_K$ and, for any $x \notin C_K$, the function $-\operatorname{cl}_2 K(x, \cdot) = -\infty$.

Proof. We will prove (a), the proof of (b) being similar. Since K is proper and closed, $\operatorname{Dom} K = C_K \times D_K$ is non-empty.

Take $\overline{y} \in D_K$. By definition, $\operatorname{cl}_1 K(x, \overline{y}) > -\infty$ for all $x \in X$ and for any $\overline{x} \in C_K$, $\operatorname{cl}_1 K(\overline{x}, \overline{y})$ is finite by Lemma 2.4.4. Hence, $\operatorname{cl}_1 K(\cdot, \overline{y})$ is a proper lower semicontinuous convex function and $C_K \subset \operatorname{dom} \operatorname{cl}_1 K(\cdot, \overline{y})$.

Take $\overline{y} \notin D_K$. Then there exists $\overline{x} \in X$ such that $\operatorname{cl}_1 K(\overline{x}, \overline{y}) = -\infty$ and the definition of the closure of a convex function implies that $\operatorname{cl}_1 K(\cdot, \overline{y}) = -\infty$. □

The following result is a simple extension of Rockafellar [145, Lemma 1].

Lemma 2.4.6. Let $K : X \times Y \to [-\infty, +\infty]$ be a proper closed saddle function and let F and G be its convex and concave parent, respectively. Then $F : X \times Y^* \to \mathbb{R} \cup \{+\infty\}$ and $(-G) : X^* \times Y \to \mathbb{R} \cup \{+\infty\}$ are proper lower semicontinuous convex functions, such that $\operatorname{dom} F \subset C_K \times Y^*$, $\operatorname{dom}(-G) \subset X^* \times D_K$. Moreover, F is the restriction of the conjugate function $(-G)^* : X^{**} \times Y^* \to \mathbb{R} \cup \{+\infty\}$ to the norm closed subspace $\widehat{X} \times Y^*$, i.e.,

$$F(x, y^*) = (-G)^*(\widehat{x}, y^*),$$

and $(-G)$ is the restriction of the conjugate function $F^* : X^* \times Y^{**} \to \mathbb{R} \cup \{+\infty\}$ to the norm closed subspace $X^* \times \widehat{Y}$, i.e.,

$$-G(x^*, y) = F^*(x^*, \widehat{y}).$$

Proof. Take a proper closed saddle function K. Then K, $\mathrm{cl}_1 K$ and $\mathrm{cl}_2 K$ have the same convex parent F according to Lemma 2.4.3. Hence, F is closed and convex because it is a supremum of closed convex functions $\mathrm{cl}_1 K(\cdot, y) + \langle \cdot, y \rangle$ on $X \times Y^*$. In a similar way one has that $(-G)$ is closed and convex.

The properness and closedness of K imply that the set $\operatorname{Dom} K = C_K \times D_K$ is non-empty. The inclusions of the domains hold from Lemma 2.4.5 and from the closedness of K. For $x \in C_K$, the function $(-\mathrm{cl}_2 K)(x, \cdot)$ is a proper lower semicontinuous convex function according to Lemma 2.4.5 and $F(x, \cdot)$ is its conjugate. Hence, $F(x, \cdot)$ also is a proper function (see Aubin and Ekeland [8, p. 201, Theorem 2]). This and the closedness of F imply that F is a proper function. Analogously one obtains that $(-G)$ is a proper function.

Further, since $F(x, \cdot)$ is the conjugate of the closed convex function $(-\mathrm{cl}_2 K)(x, \cdot)$ (see (2.63) and Lemma 2.4.5), from Fenchel duality (2.57) we have $[F(x, \cdot)]^* = -\mathrm{cl}_2 K(x, \cdot)$ on Y, i.e.,

$$-\mathrm{cl}_2 K(x, y) = \sup_{y^* \in Y^*} \{\langle y, y^* \rangle - F(x, y^*)\}.$$

For similar reasons,

$$\mathrm{cl}_1 K(x, y) = \sup_{x^* \in X^*} \{G(x^*, y) + \langle x, x^* \rangle\}. \tag{2.65}$$

As we said above, F is also the convex parent of $\mathrm{cl}_1 K$, i.e.,

$$F(x, y^*) = \sup_{y \in Y} \{\mathrm{cl}_1 K(x, y) + \langle y, y^* \rangle\}.$$

Combining the latter with (2.65) one obtains

$$F(x,y^*) = \sup_{(x^*,y)\in X^*\times Y} \{G(x^*,y) + \langle y,y^*\rangle + \langle x,x^*\rangle\}, \tag{2.66}$$

hence, $F(x,y^*)$ is the restriction of the conjugate function $(-G)^* : X^{**}\times Y^* \to \mathbb{R}\cup\{+\infty\}$ to the norm closed subspace $\widehat{X}\times Y^*$. Similarly one obtains that

$$-G(x^*,y) = \sup_{(x,y^*)\in X\times Y^*} \{-F(x,y^*) + \langle y,y^*\rangle + \langle x,x^*\rangle\}, \tag{2.67}$$

so $(-G)(x^*,y)$ is the restriction of the conjugate function $F^* : X^*\times Y^{**} \to \mathbb{R}\cup\{+\infty\}$ to the norm closed subspace $X^*\times\widehat{Y}$ and the proof is then complete. □

2.4.2 Subdifferential of a saddle function. Partially ball weakly inf-compact saddle functions. Definition and properties

The notion of the subdifferential of a saddle function $K : X\times Y \to [-\infty,+\infty]$ is introduced by Rockafellar as the multivalued mapping $\partial K : X\times Y \to 2^{X^*\times Y^*}$ defined by

$$\partial K(x,y) := \{(x^*,y^*)\in Z^* : \begin{array}{l} x^* \text{ is a subgradient of the convex function } K(\cdot,y) \text{ at } x \text{ and} \\ -y^* \text{ is a subgradient of the convex function } -K(x,\cdot) \text{ at } y\}. \end{array}$$

The (possibly empty) set $\partial K(x,y)$ is called the *subdifferential* of K at (x,y) (see Rockafellar [148, 142]). The *domain* of ∂K is defined by $\operatorname{Dom}\partial K := \{(x,y)\in X\times Y : \partial K(x,y)\neq\varnothing\}$. It is clear from the definitions and from (2.60) that when K is proper,

$$\operatorname{Dom}\partial K \subset \operatorname{Dom} K. \tag{2.68}$$

The original finite dimensional version of the following lemma is due to Rockafellar and can be found in Rockafellar [145, Lemma 4]. Our proof follows the same steps.

Lemma 2.4.7. For a proper closed saddle function $K : X\times Y \to [-\infty,+\infty]$ the following are equivalent:
(a) $(x^*,\widehat{y})\in\partial F(x,y^*)$;
(b) $(\widehat{x},y^*)\in\partial(-G)(x^*,y)$;
(c) $(x^*,-y^*)\in\partial K(x,y)$.

Any of these conditions implies that the values $F(x,y^*)$, $G(x^*,y)$ and $K(x,y)$ are finite.

Proof. First, we will show that (a) and (b) are equivalent. Since by Lemma 2.4.6 F is a proper closed convex function, we have by (2.59) that

$$(x^*,\widehat{y}) \in \partial F(x,y^*) \iff (\widehat{x},\widehat{y^*}) \in \partial F^*(x^*,\widehat{y}) \iff F^*(x^*,\widehat{y})+F(x,y^*) = \langle x,x^*\rangle+\langle y,y^*\rangle.$$

By Lemma 2.4.6 again, $F^*(x^*,\widehat{y}) = (-G)(x^*,y)$, hence

$$(x^*,\widehat{y}) \in \partial F(x,y^*) \iff (-G)(x^*,y) + F(x,y^*) = \langle x,x^*\rangle + \langle y,y^*\rangle,$$

which is

$$\text{(a)} \iff (-G)(x^*,y) + F(x,y^*) = \langle x,x^*\rangle + \langle y,y^*\rangle.$$

Since by Lemma 2.4.6 $(-G)$ is a proper closed convex function we have by (2.59) that

$$(\widehat{x},y^*)\in\partial(-G)(x^*,y) \iff (\widehat{x^*},\widehat{y})\in\partial(-G)^*(\widehat{x},y^*) \iff (-G)^*(\widehat{x},y^*)-G(x^*,y)=\langle x,x^*\rangle+\langle y,y^*\rangle.$$

By Lemma 2.4.6, $(-G)^*(\widehat{x},y^*) = F(x,y^*)$, hence

$$(\widehat{x},y^*)\in\partial(-G)(x^*,y) \iff F(x,y^*)-G(x^*,y)=\langle x,x^*\rangle+\langle y,y^*\rangle,$$

which is

$$\text{(b)} \iff F(x,y^*) - G(x^*,y) = \langle x,x^*\rangle+\langle y,y^*\rangle.$$

Finally, (a) and (b) are equivalent, since

$$\text{(a)} \iff F(x,y^*) - G(x^*,y) = \langle x,x^*\rangle + \langle y,y^*\rangle \iff \text{(b)} \tag{2.69}$$

and either of those implies the finiteness of $F(x,y^*)$ and $G(x^*,y)$.

Next, we will show that (c) implies (a) and (b).

When (c) holds it implies that $K(x,y)$ is finite and by the definition of ∂K we have that $x^* \in \partial_1 K(x,y)$, $y^* \in \partial_2(-K)(x,y)$, i.e.,

$$\begin{array}{rcll} K(z,y) & \geq & K(x,y) + \langle z-x,x^*\rangle, & \forall z \in X \\ -K(x,w) & \geq & -K(x,y) + \langle w-y,y^*\rangle, & \forall w \in Y \end{array}$$

$\iff$

$$\begin{array}{rcll} K(x,y) - \langle x,x^*\rangle & \leq & K(z,y) - \langle z,x^*\rangle, & \forall z \in X \\ K(x,y) + \langle y,y^*\rangle & \geq & K(x,w) + \langle w,y^*\rangle, & \forall w \in Y \end{array}$$

$\iff$

$$\begin{array}{l} K(x,y) - \langle x,x^*\rangle \leq G(x^*,y), \\ K(x,y) + \langle y,y^*\rangle \geq F(x,y^*). \end{array} \tag{2.70}$$

Adding the last two inequalities we obtain $F(x, y^*) - G(x^*, y) \leq \langle x, x^* \rangle + \langle y, y^* \rangle$. The opposite inequality being obvious, we have $F(x, y^*) - G(x^*, y) = \langle x, x^* \rangle + \langle y, y^* \rangle$. It follows from (2.69) that (a) and (b) hold.

Finally, let us suppose that (a) and (b) hold. The function $F(x, \cdot)$ is a proper lower semicontinuous convex function and $\widehat{y} \in \partial_2 F(x, y^*)$. Since $F(x, \cdot)$ is the conjugate of $(-\mathrm{cl}_2 K)(x, \cdot)$, it follows from (2.59) that

$$\widehat{y} \in \partial_2 F(x, y^*) \iff y^* \in \partial_2(-\mathrm{cl}_2 K)(x, y) \iff F(x, y^*) - \langle y, y^* \rangle = \mathrm{cl}_2 K(x, y).$$

The function $(-G)(\cdot, y)$ is a proper lower semicontinuous convex function and $\widehat{x} \in \partial_1(-G)(x^*, y)$. Since $(-G)(\cdot, y)$ is the conjugate of $\mathrm{cl}_1 K(\cdot, y)$, we have from (2.59)

$$\widehat{x} \in \partial_1(-G)(x^*, y) \iff x^* \in \partial_1 \mathrm{cl}_1 K(x, y) \iff G(x^*, y) + \langle x, x^* \rangle = \mathrm{cl}_1 K(x, y).$$

From (2.69), we have that the left hand sides of the last two equalities are equal, hence we obtain that $\mathrm{cl}_1 K(x, y) = \mathrm{cl}_2 K(x, y)$, and that, in particular, they are both equal to $K(x, y)$. Hence,

$$F(x, y^*) = K(x, y) + \langle y, y^* \rangle,$$

$$G(x^*, y) = K(x, y) - \langle x, x^* \rangle,$$

which entails (2.70). But it was already shown that (2.70) is equivalent to (c). The proof is then complete. □

Combining Lemma 2.4.3 and Lemma 2.4.7, one easily obtains that if K is a proper closed saddle function and L is a saddle function equivalent to K, i.e., $L \sim K$, then $\partial K = \partial L$. Hence, the subdifferential of a proper closed saddle function K depends only on the equivalence class to which K belongs and does not depend on its representatives.

It is established by Rockafellar that the domain of the subdifferential of a proper closed saddle function $K : X \times Y \to [-\infty, +\infty]$ is nonempty when one of the spaces X and Y is reflexive (see Rockafellar [146]). We will extend this property (see Theorem 2.4.11) for a class of closed saddle functions defined on product of Banach spaces introduced by the following

Definition 2.4.8. Let X, Y be Banach spaces and $K : X \times Y \to [-\infty, +\infty]$ be a saddle function.

We say that K is *X-bwc* if for some $y_0 \in D_K$ the function $\mathrm{cl}_1 K(\cdot, y_0)$ is bwc. Respectively, we say that K is *Y-bwc* if for some $x_0 \in C_K$ the function $-\mathrm{cl}_2 K(x_0, \cdot)$ is bwc.

The function K is said to be *partially ball weakly inf-compact* (*pbwc* for short) if it is X-bwc or Y-bwc.

When for any $(x_0, y_0) \in \operatorname{Dom} K$ it holds that $\operatorname{cl}_1 K(\cdot, y_0)$ or $-\operatorname{cl}_2 K(x_0, \cdot)$ is bwc, then K is said to be *totally partially ball weakly inf-compact* (*tpbwc* for short).

Let us note that when one of the spaces X and Y, say X, is reflexive then any saddle function K on $X \times Y$ is tpbwc because $\operatorname{cl}_1 K(\cdot, y)$ is X-bwc for all $y \in Y$ which is ensured by the weak compactness of the closed unit ball B_X.

From the very definition it is clear that K is pbwc whenever it is tpbwc, but when both spaces are not reflexive there exist pbwc saddle functions which are not tpbwc as we can see from the following

Example 2.4.9 (suggested by M. Ivanov). Let X and Y be two Banach spaces which are non reflexive. Fix any function $g : X \to \mathbb{R} \cup \{+\infty\}$ that is convex, lower semicontinuous and bwc function with $g(0) = 0$ and $g(x) > 0$ for all $x \in X \setminus \{0\}$, and such that for $C := \operatorname{dom} g$ the set $B_X \cap \operatorname{cl} C$ is not $w(X, X^*)$ compact and is non empty.

Define a function K from $X \times Y$ into $[-\infty, +\infty]$ by

$$K(x, y) = \begin{cases} (1 - \|y\|)g(x), & \text{if } x \in C,\ y \in B_Y \\ -\infty, & \text{if } x \in C,\ y \notin B_Y \\ +\infty, & \text{otherwise.} \end{cases}$$

Since $\operatorname{cl}_2 K(x, \cdot) = K(x, \cdot)$ for all $x \in X$ and since

$$\operatorname{cl}_1 K(\cdot, y) = \begin{cases} K(\cdot, y), & \text{if } \|y\| < 1 \\ \psi_{\operatorname{cl} C}(\cdot), & \text{if } \|y\| = 1 \\ -\infty, & \text{if } \|y\| > 1 \end{cases}$$

(here, for a subset S, ψ_S denotes the indicator function, i.e, $\psi_S(x) = 0$ if $x \in S$ and $\psi_S(x) = +\infty$ otherwise) it is not difficult to check that K is a proper closed saddle function with $\operatorname{Dom} K = C \times B_Y$.

As for $y_0 = 0 \in B_Y$ one has $\operatorname{cl}_1 K(\cdot, y_0) = K(\cdot, y_0) = g(\cdot)$ which is bwc, we get that the function K is pbwc.

We claim that K is not tpbwc. Fix any $(x_0, y_0) \in \operatorname{Dom} K = C \times B_Y$ with $\|y_0\| = 1$.

On the one hand, $\operatorname{cl}_1 K(\cdot, y_0) = \psi_{\operatorname{cl} C}(\cdot)$ which is not bwc since $B_X \cap \operatorname{cl} C$ is not $w(X, X^*)$ compact.

On the other hand,

$$\operatorname{cl}_2 K(x_0, y) = \begin{cases} (1 - \|y\|)g(x_0), & \text{if } y \in B_Y \\ -\infty, & \text{if } y \notin B_Y, \end{cases}$$

hence for $r = 0$

$$B_Y \cap \{y \in Y : -\mathrm{cl}_2 K(x_0, y) \le r\} = B_Y \cap B_Y = B_Y$$

which is not $w(Y, Y^*)$ compact. So the claim is established.

An example of such a function $g : X \to \mathbb{R} \cup \{+\infty\}$ with the properties listed above is given in $X := l_1(\mathbb{N})$ by

$$g(x) = \sum_{n=1}^{\infty} 2^n |x_n| \quad \text{for all } x \in l_1(\mathbb{N}),\ x = (x_n)_{n=1}^{\infty}.$$

Indeed, one has $\{g \le r\} = \emptyset$ for $r < 0$, $\{g \le 0\} = \{0\}$ for $r = 0$, and $\{g \le r\} \subset [-a, a]$ with $a := (r/2^n)_{n=1}^{\infty}$ and $[-a, a] := \{x \in l_1(\mathbb{N}) : -(r/2^n) \le x_n \le r/2^n, \forall n\}$ is easily seen to be totally bounded and closed and hence $\|.\|$ compact. So, the function g is a convex, lower semicontinuous bwc function with $g(0) = 0$ and $g(x) > 0$ for all $x \in X \setminus \{0\}$.

Concerning the above properties required for g, it remains to show that $L := B_X \cap \mathrm{cl}\, C$ is not $w(X, X^*)$ compact. Let e^k with $k \ge 1$ be the standard basis of $l_1(\mathbb{N}) = X$ and $c = (c_n)_{n=1}^{\infty} \in l_\infty(\mathbb{N}) = X^*$ with $0 < c_n < c_{n+1}$ and $c_n \to 1$, so $\|c\|_\infty = 1$. Since $e^k \in L$ and $\langle c, e^k \rangle = c_k \to 1$, one has $\sup_{x \in L} \langle c, x \rangle = 1$. On the other hand, for each $x \in L \setminus \{0\}$

$$\langle c, x \rangle \le \sum_{n=1}^{\infty} c_n |x_n| < \sum_{n=1}^{\infty} |x_n| \le 1,$$

so $\langle c, \cdot \rangle$ does not attain its maximum over L, which means that L is not $w(X, X^*)$ compact according to James theorem.

Below we will find out that the pbwc saddle functions defined on a product Banach space possess many of the well-known properties of the saddle functions defined on a product of Banach spaces one of which is reflexive.

Lemma 2.4.10. Let $X,\ Y$ be Banach spaces, let $K : X \times Y \to [-\infty, +\infty]$ be a proper closed X-bwc saddle function and let G be its concave parent. Then the following properties hold:
(a) $\mathrm{dom}(-G)^* \subset \widehat{X} \times Y^*$;
(b) the range of $\partial(-G)$ is a non-empty subset of $\widehat{X} \times Y^*$;
(c) $(x^*, y^{**}) \in \partial F(x, y^*) \iff (\widehat{x^*}, y^{**}) \in \partial(-G)^*(\widehat{x}, y^*)$.

Proof. (a) Take any $(x^{**}, y^*) \in \operatorname{dom}(-G)^*$ and set $r := (-G)^*(x^{**}, y^*)$. Then $r \in \mathbb{R}$ and the equality $(-G)^*(\cdot, y) = [\operatorname{cl}_1 K(\cdot, y)]^*$ yields

$$\begin{aligned} r := (-G)^*(x^{**}, y^*) &= \sup_{(x^*,y)\in X^*\times Y} \{\langle y, y^*\rangle + \langle x^*, x^{**}\rangle + G(x^*, y)\} \\ &= \sup_{(x^*,y)\in X^*\times Y} \{\langle y, y^*\rangle + \langle x^*, x^{**}\rangle - [\operatorname{cl}_1 K(\cdot, y)]^*\} \\ &= \sup_{(x^*,y)\in X^*\times D_K} \{\langle y, y^*\rangle + \langle x^*, x^{**}\rangle - [\operatorname{cl}_1 K(\cdot, y)]^*(x^*)\} \\ &= \sup_{y\in D_K}\{\langle y, y^*\rangle + \sup_{x^*\in X^*}\{\langle x^*, x^{**}\rangle - [\operatorname{cl}_1 K(\cdot, y)]^*(x^*)\}\} \\ &= \sup_{y\in D_K}\{\langle y, y^*\rangle + [\operatorname{cl}_1 K(\cdot, y)]^{**}(x^{**})\}, \end{aligned}$$

where the third equality is due to the fact that, for any $y \notin D_K$, one has $[\operatorname{cl}_1 K(\cdot, y]^* = +\infty$, since $\operatorname{cl}_1 K(\cdot, y) = -\infty$ according to Lemma 2.4.5. Hence, x^{**} is in the effective domain of the biconjugate of $\operatorname{cl}_1 K(\cdot, y)$ for any $y \in D_K$. Since for some $y_0 \in D_K$ the function $\operatorname{cl}_1 K(\cdot, y_0)$ is bwc, Theorem 2.4.2 (c) ensures that $x^{**} \in \widehat{X}$.
(b) From Lemma 2.4.6, $(-G)$ is a proper closed convex function, hence the range of $\partial(-G)$ is non-empty.

Take any $(x^{**}, y^*) \in \partial(-G)(x^*, y)$. By (2.59), $(\widehat{x^*}, \widehat{y}) \in \partial(-G)^*(x^{**}, y^*)$. From (a) we have $\operatorname{dom}\partial(-G)^* \subset \operatorname{dom}(-G)^* \subset \widehat{X} \times Y^*$. So, there exists some $x \in X$ such that $x^{**} = \widehat{x}$.
(c) follows from Lemma 2.4.6. □

Theorem 2.4.11. Let X, Y be Banach spaces and let $K : X \times Y \to [-\infty, +\infty]$ be a proper closed pbwc saddle function. Then $\operatorname{Dom}\partial K \neq \emptyset$.

Proof. Without loss of generality we suppose that K is X-bwc. For the proper closed convex function $(-G)$ there exists $(x^{**}, y^*) \in \partial(-G)(x^*, y)$. From Lemma 2.4.10 (b) we have that $x^{**} = \widehat{x}$ for some $x \in X$, i.e., we have that $(\widehat{x}, y^*) \in \partial(-G)(x^*, y)$. Lemma 2.4.7 ensures that this is equivalent to $(x^*, -y^*) \in \partial K(x, y)$. The latter says in particular that $\operatorname{Dom}\partial K$ is a non-empty set. The proof is then complete. □

From Theorem 2.4.11 we can derive the density of the domain of the subdifferential in the domain of a proper closed tpbwc saddle function. This result is conjectured for any proper closed saddle function by Rockafellar (see Rockafellar [146, p. 249]). It is established by him when X and Y are finite dimensional spaces (see Rockafellar [142, 148]) and extended by Gossez (see Gossez [79, Theorem 1]) to the case when one of the spaces X and Y is reflexive. The reflexivity assumption enters the latter proof essentially to ensure that in this case $\operatorname{Dom}\partial K$ is a non-empty set. We will follow the proof of Gossez. To this end let us recall Gossez [79, Corollary 1] which states: Let K be a closed saddle function on the product Banach space $Z = X \times Y$ and let $C \subset X$ and $D \subset Y$ be closed convex sets such that the interior of $(C \times D)$ meets $\operatorname{Dom} K$. Define

$$K_0(x,y) := \begin{cases} K(x,y), & \text{if } x \in C,\ y \in D \\ +\infty, & \text{if } x \notin C,\ y \in D \\ -\infty, & \text{if } y \notin D. \end{cases} \tag{2.71}$$

Then K_0 is a closed saddle function with

$$\operatorname{Dom} K_0 = \operatorname{Dom} K \cap [C \times D] \tag{2.72}$$

and for each $(x,y) \in X \times Y$

$$\partial K_0(x,y) = \partial K(x,y) + [N_C(x) \times (-N_D(y))], \tag{2.73}$$

where $N_C(x)$ (resp. $N_D(y)$) denotes the normal cone to C (resp. D) at x (resp. y).

Theorem 2.4.12. Let X, Y be Banach spaces and let $K : X \times Y \to [-\infty, +\infty]$ be a proper closed tpbwc saddle function. Then $\operatorname{Dom} \partial K$ is dense in $\operatorname{Dom} K$.

Proof. We will use Theorem 2.4.11 and will follow Gossez's proof of [79, Theorem 1].

Take any $(x_0, y_0) \in \operatorname{dom} K$. Without loss of generality we suppose that $\operatorname{cl}_1 K(\cdot, y_0)$ is bwc.

Fix $\varepsilon > 0$ and consider the function K_0 as above with $C := x_0 + \varepsilon B_X$ and $D := y_0 + \varepsilon B_Y$. Observe that (2.72) entails that $D_{K_0} \subset D_K$ and that, for any $y \in D_{K_0}$, we have by (2.71) that

$$K(\cdot, y) \le K_0(\cdot, y), \quad \text{and hence} \quad \operatorname{cl}_1 K(\cdot, y) \le \operatorname{cl}_1 K_0(\cdot, y).$$

So $\{\operatorname{cl}_1 K_0(\cdot, y) \le r\} \subset \{\operatorname{cl}_1 K(\cdot, y) \le r\}$ for all $r \in \mathbb{R}$ and $y \in D_{K_0}$. This says in particular that $\operatorname{cl}_1 K_0(\cdot, y_0)$ is bwc, hence K_0 is pbwc. Then Theorem 2.4.11 ensures that $\operatorname{Dom} \partial K_0 \neq \emptyset$ and hence (2.73) gives some $(x,y) \in C \times D$ such that $\partial K(x,y) \neq \emptyset$. The proof is then complete. □

Another interesting property of a closed proper pbwc saddle function is that the graphs of the subdifferentials of its parent functions are completely determined by the graph of its subdifferential.

Lemma 2.4.13. Let X, Y be Banach spaces and let $K : X \times Y \to [-\infty, +\infty]$ be a proper closed pbwc saddle function. Then $\operatorname{gph} \partial F$ and $\operatorname{gph} \partial(-G)$ are completely determined by the elements of $\operatorname{gph} \partial K$.

Proof. Suppose that K is X-bwc.

Lemma 2.4.7 shows that $(\widehat{x}, y^*) \in \partial(-G)(x^*, y)$ exactly when $(x^*, -y^*) \in \partial K(x, y)$, so by Lemma 2.4.10 (b) $\operatorname{gph} \partial(-G)$ is completely determined by $\operatorname{gph} \partial K$.

From Lemma 2.4.6 we know that F is a proper closed convex function. The density of $\operatorname{dom} \partial F$ in the non-empty set $\operatorname{dom} F$ implies that $\operatorname{dom} \partial F$ is non-empty. Let $(x^*, y^{**}) \in \partial F(x, y^*)$. From Lemma 2.4.10 (c) we have that $(x^*, y^{**}) \in \partial F(x, y^*)$ exactly when $(\widehat{x^*}, y^{**}) \in \partial(-G)^*(\widehat{x}, y^*)$. Set W to be the Banach space on which the concave parent G of K is defined, i.e., $W := X^* \times Y$. For the conjugate function $(-G)^*$ Lemma 2.4.10 (a) gives that $\operatorname{dom} \partial(-G)^* \subset \widehat{X} \times Y^*$. This combined with Theorem 2.4.1 ensures that:

$$
(2.74) \quad \begin{cases} (\widehat{x^*}, y^{**}) \in \partial(-G)^*(\widehat{x}, y^*) \iff \\ \text{there exists a net } \{(\widehat{x_\alpha}, {y_\alpha}^*)\}_{\alpha \in A} \in W^* \text{ converging to } (\widehat{x}, y^*) \\ \text{in the norm topology of } W^* \text{ and} \\ \text{a bounded net } \{(\widehat{x_\alpha^*}, \widehat{y_\alpha})\}_{\alpha \in A} \in \widehat{W} \text{ converging to } (\widehat{x^*}, y^{**}) \\ \text{in the } w(W^{**}, W^*) \text{ topology such that} \\ (\widehat{x_\alpha}, y_\alpha^*) \in \partial(-G)(x_\alpha^*, y_\alpha), \ \forall \alpha \in A. \end{cases}
$$

Lemma 2.4.7 implies that $(\widehat{x_\alpha}, y_\alpha^*) \in \partial(-G)(x_\alpha^*, y_\alpha)$ exactly when $(x_\alpha^*, -y_\alpha^*) \in \partial K(x_\alpha, y_\alpha)$. Hence, (2.74) may be rewritten as:

$$
(2.75) \quad \begin{cases} (x^*, y^{**}) \in \partial F(x, y^*) \iff \\ \text{there exists a net } \{(\widehat{x_\alpha}, {y_\alpha}^*)\}_{\alpha \in A} \in W^* \text{ converging to } (\widehat{x}, y^*) \\ \text{in the norm topology of } W^* \text{ and} \\ \text{a bounded net } \{(\widehat{x_\alpha^*}, \widehat{y_\alpha})\}_{\alpha \in A} \in \widehat{W} \text{ converging to } (\widehat{x^*}, y^{**}) \\ \text{in the } w(W^{**}, W^*) \text{ topology such that} \\ (x_\alpha^*, -y_\alpha^*) \in \partial K(x_\alpha, y_\alpha), \ \forall \alpha \in A. \end{cases}
$$

The latter says that $\operatorname{gph} F$ is completely determined by the elements of $\operatorname{gph} \partial K$. □

With any saddle function $K : X \times Y \to [-\infty, +\infty]$ one may associate a monotone (see below) multivalued operator $T_K : X \times Y \to 2^{X^* \times Y^*}$ given by

$$
T_K(x, y) := \{(x^*, y^*) \in X^* \times Y^* : (x^*, -y^*) \in \partial K(x, y)\}.
$$

When K is proper and closed, we saw in Lemma 2.4.7 that ∂K depends only on the equivalence class. The above definition ensures that in this case the same holds for the operator T_K. This operator was introduced in relation with minimax problems by Rockafellar [146] who proved that in the case when one of the Banach spaces involved is assumed to be reflexive, T_K is maximal monotone whenever K is proper and closed (see Rockafellar [146, Theorem 3]). We will show that this is still true when

K is a proper closed partially ball weakly inf-compact saddle function defined on a product Banach space. Closed relationship between proper closed saddle functions on $X \times X$ and maximal monotone operators on X is established by Krauss in [105].

Let us recall that a set-valued mapping S from a Banach space X to its dual X^* is said to be *monotone* if $x^* \in S(x)$, $y^* \in S(y)$ imply $\langle x - y, x^* - y^* \rangle \geq 0$. The operator S is said to be *maximal monotone* if S is monotone and S has no proper monotone extension, i.e., if $(x, x^*) \in X \times X^*$ is such that the monotone relation $\langle x - y, x^* - y^* \rangle \geq 0$ holds for all $(y, y^*) \in \operatorname{gph} S$, then $(x, x^*) \in \operatorname{gph} S$ (cf., e.g., Phelps [133], Rockafellar [147] and Simons [156]).

Theorem 2.4.14. Let X, Y be Banach spaces and let $K : X \times Y \to [-\infty, +\infty]$ be a closed proper pbwc saddle function. Then the operator T_K is maximal monotone.

Proof. Let K be X-bwc. By Lemma 2.4.10, (b) we know that the range of $\partial(-G)$ lies in $\widehat{X} \times Y^*$. By Lemma 2.4.7 and by the definition of T_K it is clear that $(x^*, y^*) \in T_K(x, y)$ exactly when $(\widehat{x}, y^*) \in \partial(-G)(x^*, y)$. The latter, being the subdifferential of a proper closed convex function, is maximal monotone (see Rockafellar [147]). This implies the maximal monotonicity of T_K. □

Let us note that in the paper of Pak [131] one can find the statement of the above result for arbitrary proper closed saddle function defined on a product of arbitrary Banach spaces, but the presented there proof implicitly presumes reflexivity.

2.4.3 Integrability of the subdifferential of a proper closed partially ball weakly inf-compact saddle function

Here we are interested in the integrability of the subdifferential of a saddle function on the product of Banach spaces. The result is established for Lipschitz saddle function by Correa and Thibault in [54]. Some generalizations for directionally Lipschitz saddle functions are already presented in Section 2.3.

We consider two proper closed saddle functions $K, L : X \times Y \to [-\infty, +\infty]$ defined on a product Banach space and we are interested whether the pbwc condition on one of the functions K and L and the inclusion $\partial L \subset \partial K$ entail that K and L are equivalent up to a finite additive constant.

First we will consider the case when the inside for the subdifferential inclusion function L is supposed to be pbwc.

Theorem 2.4.15. Let X, Y be Banach spaces and let $L : X \times Y \to [-\infty, +\infty]$ be a proper closed pbwc saddle function.

Then for any proper closed saddle function $K : X \times Y \to [-\infty, +\infty]$ the condition

$$\partial L(x, y) \subset \partial K(x, y), \quad \forall (x, y) \in \operatorname{Dom} \partial L$$

implies that K and L are equivalent up to a finite additive constant c, i.e., $K \sim (L + c)$.

Proof. Let K and L be saddle functions satisfying the assumptions of the theorem.

Let us denote by F' and F the convex parents of L and K, respectively and by G' and G the concave parents of L and K, respectively.

Since L and K are proper and closed we know by Lemma 2.4.6 that the functions $F', F, (-G')$ and $(-G)$ are proper and lower semicontinuous convex functions.

Since L is supposed to be pbwc, without loss of generality we suppose that L is X-bwc.

First, we will show that

$$\partial(-G') \subset \partial(-G). \tag{2.76}$$

Since L is X-bwc, by Lemma 2.4.10 (b) we have $\operatorname{Rge} \partial(-G') \subset \widehat{X} \times Y^*$. Take any $(\widehat{x}, y^*) \in \partial(-G')(x^*, y)$. Lemma 2.4.7 for L ensures that $(x^*, -y^*) \in \partial L(x, y)$, so $(x^*, -y^*) \in \partial K(x, y)$. Lemma 2.4.7 for K yields that $(\widehat{x}, y^*) \in \partial(-G)(x^*, y)$. Hence, we have proved (2.76).

Second, we will show that

$$\partial(F') \subset \partial(F). \tag{2.77}$$

Take any $(x^*, y^{**}) \in \partial F'(x, y^*)$. Since L is X-bwc, Lemma 2.4.10 (c) gives that the latter implies $(\widehat{x^*}, y^{**}) \in \partial(-G')^*(\widehat{x}, y^*)$. Since $(-G')^*$ is the conjugate function of the lower semicontinuous convex function $(-G')$, then Theorem 2.4.1 applied to it gives that (recalling the notation $W := X^* \times Y$)

$$\begin{cases} (\widehat{x^*}, y^{**}) \in \partial(-G')^*(\widehat{x}, y^*) \Rightarrow \\ \text{there exists a net } \{(x_\alpha{}^{**}, y_\alpha{}^*)\}_{\alpha \in A} \in W^* \text{ converging to } (\widehat{x}, y^*) \\ \text{in the norm topology of } W^* \text{ and} \\ \text{a bounded net } \{(\widehat{x^*_\alpha}, \widehat{y_\alpha})\}_{\alpha \in A} \in \widehat{W} \text{ converging to } (\widehat{x^*}, y^{**}) \\ \text{in the } w(W^{**}, W^*) \text{ topology such that} \\ (x_\alpha{}^{**}, y^*_\alpha) \in \partial(-G')(x^*_\alpha, y_\alpha), \ \forall \alpha \in A. \end{cases} \tag{2.78}$$

The X-bwc property of L yields according to Lemma 2.4.10 (b) that $x^{**}_\alpha = \widehat{x_\alpha}$ for some $x_\alpha \in X$. From Lemma 2.4.7 $(\widehat{x_\alpha}, y^*_\alpha) \in \partial(-G')(x^*_\alpha, y_\alpha)$ implies that $(x_\alpha{}^*, -y^*_\alpha) \in \partial L(x_\alpha, y_\alpha)$. From $\partial L \subset \partial K$ we have $(x_\alpha{}^*, -y^*_\alpha) \in \partial K(x_\alpha, y_\alpha)$ which by Lemma 2.4.7 yields that $(\widehat{x_\alpha}, y_\alpha{}^*) \in \partial(-G)(x^*_\alpha, y_\alpha)$.

Since $(\widehat{x_\alpha}, y_\alpha^*) \in \partial(-G)(x_\alpha^*, y_\alpha)$ for all $\alpha \in A$ and since the net $\{(\widehat{x_\alpha}, {y_\alpha}^*)\}_{\alpha\in A} \in W^*$ converges to $(\widehat{x}, y^*)$ in the norm topology of W^* and the bounded net $\{(\widehat{x_\alpha^*}, \widehat{y_\alpha})\}_{\alpha\in A} \in \widehat{W}$ converges to $(\widehat{x^*}, y^{**})$ in the $w(W^{**}, W^*)$ topology, Theorem 2.4.1 implies that $(\widehat{x^*}, y^{**}) \in \partial(-G)^*(\widehat{x}, y^*)$. As $F(u, v^*) = (-G)^*(\widehat{u}, v^*)$ for all $(u, v^*) \in X \times Y^*$ by Lemma 2.4.6, we deduce that $(x^*, y^{**}) \in \partial F(x, y^*)$. Hence, (2.77) is established.

The functions F and F' are both proper lower semicontinuous convex functions on the Banach space $X \times Y^*$ satisfying $\partial F' \subset \partial F$. Hence, we can apply Rockafellar [147, Theorem B] (see also Thibault and Zagrodny [160, Corollary 2.2]) to deduce that there exists a finite constant c such that

$$F(x, y^*) = F'(x, y^*) + c, \quad \forall (x, y^*) \in X \times Y^*. \tag{2.79}$$

The functions $(-G')$ and $(-G)$ are both proper lower semicontinuous convex functions satisfying $\partial(-G') \subset \partial(-G)$. Analogous reasoning gives a finite constant d such that

$$G(x^*, y) = G'(x^*, y) + d, \quad \forall (x^*, y) \in X^* \times Y. \tag{2.80}$$

Now using (2.66) from Lemma 2.4.6, (2.80) and (2.79) we obtain that

$$\begin{aligned} F(x, y^*) &= \sup_{(x^*, y) \in X^* \times Y} \{G(x^*, y) + \langle y, y^* \rangle + \langle x, x^* \rangle\} \\ &= \sup_{(x^*, y) \in X^* \times Y} \{G'(x^*, y) + \langle y, y^* \rangle + \langle x, x^* \rangle\} + d \\ &= F'(x, y^*) + d = F(x, y^*) - c + d, \end{aligned}$$

and $c = d$ because F' is finite at some point. From (2.79) and (2.80) it is clear that the functions K and $L + c$ have the same parent functions. By Lemma 2.4.3 one obtains that $K \sim (L + c)$ and K and L are equivalent up to the additive constant c. The proof is then complete. □

Second we consider the case when the outside for the subdifferential inclusion function K is supposed to be pbwc.

Theorem 2.4.16. Let X, Y be Banach spaces and let $K : X \times Y \to [-\infty, +\infty]$ be a proper closed pbwc saddle function. Then for any proper closed saddle function $L : X \times Y \to [-\infty, +\infty]$ the conditions

$$\partial_1 \mathrm{cl}_1 L(\cdot, y) \subset \partial_1 K(\cdot, y), \text{ and } \partial_2(-\mathrm{cl}_2 L)(x, \cdot) \subset \partial_2(-K)(x, \cdot), \forall (x, y) \in X \times Y \tag{2.81}$$

imply that K and L are equivalent up to a finite additive constant c, i.e., $K \sim (L + c)$.

Proof. Let $(\bar{x}, \bar{y}) \in \operatorname{Dom} L$. By Lemma 2.4.5 we know that the function $\operatorname{cl}_1 L(\cdot, \bar{y})$ is a proper lower semicontinuous convex function and that $\bar{x} \in \operatorname{dom} \operatorname{cl}_1 L(\cdot, \bar{y})$. By the density of the domain of the subdifferential of a closed convex function in its effective domain (see (2.58)) it follows that there exists $\tilde{x} \in \operatorname{dom} \partial_1 \operatorname{cl}_1 L(\cdot, \bar{y}) \subset \operatorname{dom} \partial_1 K(\cdot, \bar{y})$, the latter by (2.81). From (2.60) we have that $\tilde{x} \in \operatorname{dom} \operatorname{cl}_1 K(\cdot, \bar{y})$. Hence, $\operatorname{cl}_1 K(\cdot, \bar{y})$ is a proper closed convex function too and further the first inclusion of (2.81) yields $\partial_1 \operatorname{cl}_1 L(\cdot, \bar{y}) \subset \partial_1 \operatorname{cl}_1 K(\cdot, \bar{y})$. We can then apply Rockafellar [147, Theorem B] (see also Thibault and Zagrodny [160, Corollary 2.2]) to obtain that there exists a finite constant $c(\bar{y})$ such that

$$\begin{aligned} &\operatorname{cl}_1 L(x, \bar{y}) = \operatorname{cl}_1 K(x, \bar{y}) + c(\bar{y}), \quad \forall x \in X \quad \text{and} \\ &\operatorname{dom} \operatorname{cl}_1 L(\cdot, \bar{y}) \equiv \operatorname{dom} \operatorname{cl}_1 K(\cdot, \bar{y}). \end{aligned} \tag{2.82}$$

Analogously, one obtains that there exists a finite constant $d(\bar{x})$ such that

$$\begin{aligned} &\operatorname{cl}_2 L(\bar{x}, y) = \operatorname{cl}_2 K(\bar{x}, y) + d(\bar{x}), \quad \forall y \in Y \quad \text{and} \\ &\operatorname{dom}(-\operatorname{cl}_2 L)(\bar{x}, \cdot) \equiv \operatorname{dom}(-\operatorname{cl}_2 K)(\bar{x}, \cdot). \end{aligned} \tag{2.83}$$

Recall that (2.82) and (2.83) hold for all $(\bar{x}, \bar{y}) \in \operatorname{Dom} L$. From (2.82), (2.83) and Lemma 2.4.4 it follows that $(\bar{x}, \bar{y}) \in \operatorname{Dom} K$, so $\operatorname{Dom} L \subset \operatorname{Dom} K$.

To establish the opposite inclusion of domains, let us take any $(\bar{x}, \bar{y}) \in \operatorname{Dom} K$. Suppose that $\bar{y} \notin D_L$. Then $\operatorname{cl}_1 L(\cdot, \bar{y}) = -\infty$ and $\partial_1 \operatorname{cl}_1 L(x, \bar{y}) \equiv X^*$ for any $x \in X$. The first inclusion in (2.81) ensures that $K(\cdot, \bar{y})$ is everywhere subdifferentiable, hence from (2.60), $K(\cdot, \bar{y}) = \operatorname{cl}_1 K(\cdot, \bar{y})$, and $\partial_1 \operatorname{cl}_1 K(\cdot, \bar{y}) = \partial_1 K(\cdot, \bar{y}) = X^*$, which yields a contradiction, because $\operatorname{cl}_1 K(\cdot, \bar{y})$ is a proper closed convex function. By similar reasons we obtain that $\bar{x} \in C_L$, hence $\operatorname{Dom} K \subset \operatorname{Dom} L$ and, finally, $\operatorname{Dom} L = \operatorname{Dom} K$.

As K is pbwc, we may suppose that K is X-bwc. Take $y_0 \in D_K = D_L$ such that $\operatorname{cl}_1 K(\cdot, y_0)$ is bwc. Fix $r \in \mathbb{R}$ and $n \in \mathbb{N}$ and consider the sublevel set $P_{r,n} := \{\operatorname{cl}_1 L(\cdot, y_0) \leq r\} \cap nB_X$. Since the function $\operatorname{cl}_1 L(\cdot, y_0)$ is convex and lower semicontinuous the latter set is weakly closed. By (2.82), $P_{r,n} \subset \{\operatorname{cl}_1 K(\cdot, y_0) \leq r - c(y_0)\} \cap nB_X$ which is weakly compact. Hence, $P_{r,n}$ is weakly compact, which implies that the closed saddle function L is X-bwc too. Then by Theorem 2.4.11 $\operatorname{Dom} \partial L \neq \emptyset$. Take any $(x^*, y^*) \in \partial L(x, y)$. In particular, $x^* \in \partial_1 L(\cdot, y)(x)$. From (2.60) and the properness of $L(\cdot, y)$ we have that $L(x, y) = \operatorname{cl}_1 L(x, y)$ and that $\partial_1 \operatorname{cl}_1 L(\cdot, y)(x) = \partial_1 L(\cdot, y)(x)$. The latter says that $x^* \in \partial_1 \operatorname{cl}_1 L(\cdot, y)(x)$. The assumption (2.81) gives that $x^* \in \partial_1 K(\cdot, y)(x)$. Analogously one obtains that $-y^* \in \partial_2(-K(x, \cdot))(y)$. Hence, $(x^*, y^*) \in \partial K(x, y)$. The arbitrariness of (x^*, y^*) ensures that $\partial L(x, y) \subset \partial K(x, y)$ for all $(x, y) \in \operatorname{Dom} \partial L$. The assumptions of Theorem 2.4.15 being satisfied, we conclude that the result is established. □

Chapter 3

Variational analysis of multivalued maps

In this chapter we study certain multivalued maps, as well as, multivalued maps depending on parameter. Such maps are considered in optimization and are intensively studied last years.

In Section 3.1 we study Lipschitz continuity with respect to the parameter of the set of solutions of a parameterized minimax problem on a product Banach space. We present a sufficient condition, ensuring that the map which to any value of the parameter assigns the set of solutions of the problem (possibly multi-valued, and unbounded) possesses Aubin property. Results are published by Quincampoix and Zlateva in [140].

In Section 3.2 we present a derivative criterion for metric regularity of set-valued mappings that is based on works of J.-P. Aubin and co-authors. A related implicit mapping theorem is also obtained. As applications, we first show that Aubin criterion leads directly to the known fact that the mapping describing an equality/inequality system is metrically regular if and only if the Mangasarian-Fromovitz condition holds. We also derive a new necessary and sufficient condition for strong regularity of variational inequalities over polyhedral sets. A new proof of the radius theorem for metric regularity based on Aubin criterion is given as well. Results are published by Dontchev, Quincampoix and Zlateva in [63].

In Section 3.3 is proved Long orbit or empty value principle for a multivalued map ant it is applied to provide unified approach to several fixed point and surjectivity results. All of the latter are derived from a novel general result. Results are published by Ivanov and Zlateva in [96].

3.1 Parameterized minimax problem: on Lipschitz-like dependence of the solution with respect to parameter

In this section we establish quite general sufficient conditions for Aubin continuity of the saddle point map $\mathcal{S} : \lambda \rightrightarrows \mathcal{S}(\lambda)$ arising from the parameterized minimax problem $M(\lambda)$. Examples illustrating these conditions are presented. Several corollaries related to the case of convex-concave smooth data are also sketched.

Consider the parameterized minimax problem

$$M(\lambda) \qquad \inf_{x\in K} \sup_{y\in L} f(x,y,\lambda),$$

where $\lambda \in \Lambda$ is a parameter.

Here K and L are nonempty closed subsets of the Banach spaces X and Y, respectively; $\{f(\cdot,\cdot,\lambda) : X \times Y \to \mathbb{R}, \lambda \in \Lambda\}$ is a family of real-valued functions parameterized by $\lambda \in \Lambda$, where Λ is a subset of the Banach space Z.

Saddle point of $f(\cdot,\cdot,\lambda)$ on $K \times L$ is any point $(\overline{x},\overline{y}) \in K \times L$ that satisfies

$$f(\overline{x},y,\lambda) \le f(\overline{x},\overline{y},\lambda) \le f(x,\overline{y},\lambda) \quad \forall x \in K, \quad \forall y \in L.$$

A saddle point $(\overline{x},\overline{y})$ of $f(\cdot,\cdot,\lambda)$ on $K \times L$ can be considered as a solution of the minimax problem $M(\lambda)$ by reason of $(\overline{x},\overline{y}) \in K \times L$ and $f(\overline{x},\overline{y}) = \inf_{x\in K} \sup_{y\in L} f(x,y,\lambda)$. Let us denote the (possibly empty) set of all saddle points of the function $f(\cdot,\cdot,\lambda)$ on $K \times L$ by

$$\text{(3.1)} \qquad \mathcal{S}(\lambda) := \{(\overline{x},\overline{y}){\in}K{\times}L : f(\overline{x},y,\lambda) \le f(\overline{x},\overline{y},\lambda) \le f(x,\overline{y},\lambda),\ \forall x \in K,\ \forall y \in L\}.$$

That $\mathcal{S}(\lambda)$ is nonempty can be ensured in several cases. For example, if K and L are convex sets, $f(x,y,\lambda)$ is convex and lower semicontinuous in x, concave and upper semicontinuous in y, and there are $x_0 \in K$ and $y_0 \in L$ such that $f(\cdot,y_0,\lambda)$ is inf-compact and $f(x_0,\cdot,\lambda)$ is sup-compact, then $\mathcal{S}(\lambda) \neq \varnothing$ by a minimax result due to Hartung [81, Theorem 1] (see also Aubin and Ekeland [8, Theorem 6.2.8]). When, moreover, $f(x,y,\lambda)$ is strictly convex in x and strictly concave in y, then $\mathcal{S}(\lambda)$ is a singleton.

In the present section we presume the existence of saddle points for $M(\lambda)$ and focus our attention on studying Lipschitz-like dependence of the solution set $\mathcal{S}(\lambda)$ on the parameter λ. That is, we find sufficient conditions for Lipschitz-like continuity of the set-valued map

$$\mathcal{S} : \lambda \rightrightarrows \mathcal{S}(\lambda)$$

from Λ to nonempty subsets of $K \times L$.

Of course, when the map $\mathcal{S}$ is single-valued, the Lipschitz continuity is understood in the classical sense. However, the map $\mathcal{S}$ could be multivalued. Moreover, its values $\mathcal{S}(\lambda)$ could be unbounded sets. A notion of Lipschitz-like continuity very appropriate for such a case is due to Aubin [4, 5]:

The multivalued map $S : \Lambda \rightrightarrows X$ has *Aubin property*, or it is *Aubin continuous, near* $(\overline{\lambda}, \overline{x}) \in \operatorname{gph} S$, if there are positive constant κ and neighborhoods U of $\overline{x}$, and V of $\overline{\lambda}$, such that

$$e(S(\lambda) \cap U, S(\mu)) \leq \kappa\|\lambda - \mu\|, \quad \forall \lambda, \mu \in \Lambda \cap V, \tag{3.2}$$

where $e(A, B) := \sup_{x \in A} d(x, B)$ is the *excess* from set A to set B with $e(\emptyset, B) = +\infty$. S is said to be *Aubin continuous* if S is Aubin continuous near any point $(\overline{\lambda}, \overline{x}) \in \operatorname{gph} S$.

For various applications of Aubin continuity in the field of nonlinear analysis and optimization the reader is referred, e.g., to Aubin [4, 5], Aubin and Frankowska [10] and Rockafellar and Wets [152]. The Aubin property of a map S near $(\overline{\lambda}, \overline{x})$ is known to be equivalent to the metric regularity of S^{-1} near $(\overline{x}, \overline{\lambda})$ and was originally introduced in Aubin [5] under the name of *pseudo-Lipschitz continuity*. For bibliographical details see Rockafellar and Wets book [152].

Whenever S is locally bounded, Aubin continuity coincides with the classical notion for Lipschitz continuity of set-valued maps (see Aubin and Frankowska [10] and Rockafellar and Wets [152])

$$e(S(\lambda), S(\mu)) \leq \kappa\|\lambda - \mu\|, \quad \forall \lambda, \mu,$$

but Aubin property works without any boundedness imposed on the values of S. Aubin property is in fact Lipschitzean property localized in the range space, as well as in the domain space.

In Subsection 3.1.1, after a short subsection devoted to preliminaries, we formulate and prove a sufficient condition for Aubin continuity of the solution map $S : \Lambda \rightrightarrows X$ of a parameterized minimization problem

$$P(\lambda) \qquad \inf_{x \in K} f(x, \lambda).$$

Many authors study Lipschitz-like dependence on λ of the solutions of the associated generalized Euler equation

$$0 \in \nabla_x f(x, \lambda) + N_K(x);$$

see Bonnans and Shapiro [22], Dontchev and Rockafellar [65], Shapiro [155] and the references therein for recent developments. Here we do not follow that approach, because the map $St : \lambda \rightrightarrows St(\lambda)$, which to any λ assigns the set $St(\lambda)$ of solutions of the generalized Euler equation, does not inherit Aubin continuity property from S (see Example 3.1.6 for a parameterized problem such that the corresponding S is Aubin continuous while St is not).

In Subsection 3.1.2 we present our main result (Theorem 3.1.10), which is a sufficient condition for Aubin continuity of the saddle point map $\mathcal{S} : \Lambda \rightrightarrows X \times Y$ of a parameterized minimax problem $M(\lambda)$.

It is clear that the results of Subsection 3.1.1 are contained in the more general framework of Subsection 3.1.2. Nevertheless, we think that presenting the proof of the former simple case will help the understanding of the more technical proof of the latter general case.

Subsection 3.1.3 relates the obtained results to some questions in the field of two-player zero sum differential games.

3.1.1 Parameterized minimization problem

We work in a Banach space $(X, \|\cdot\|)$.

For $C \subset X$ the *distance function* to C is $d(x, C) := \inf_{c \in C} \|x - c\|$ if $C \neq \emptyset$, and $d(x, C) := +\infty$ if $C = \emptyset$.

Function $f : X \to \mathbb{R}$ is *Gâteaux differentiable* at $\overline{x} \in X$ if there exists $\nabla f(\overline{x}) \in X^*$, called the *Gâteaux derivative* of f at $\overline{x}$, such that for any $h \in X$,

$$\lim_{t \to 0} \frac{f(\overline{x} + th) - f(\overline{x})}{t} = \langle \nabla f(\overline{x}), h \rangle.$$

Also, f is said to be *strictly differentiable* at $\overline{x}$ whenever

$$\lim_{\substack{x \to \overline{x} \\ t \to 0}} \frac{f(x + th) - f(x)}{t} = \langle \nabla f(\overline{x}), h \rangle.$$

Given an open set $U \subset X$ we denote by $C^{1,\alpha}(U)$ the class of all Gâteaux differentiable functions $f : U \to \mathbb{R}$ such that $\nabla f : U \to X^*$ is *α-Hölder* on U, that is, for some constant $L > 0$,

$$\|\nabla f(x) - \nabla f(y)\| \leq L \|x - y\|^{\alpha}, \quad \forall x, y \in U.$$

Let Z be a Banach space, whose norm is also denoted by $\|\cdot\|$. Let S be a *map* from $\Lambda \subset Z$ to X. If not stated otherwise, *map* means *set-valued map*. In order to outline the multivaluedness we write $S : Z \rightrightarrows X$. The *inverse* $S^{-1} : X \rightrightarrows Z$ of S is defined by $\lambda \in S^{-1}(x) \iff x \in S(\lambda)$. The *graph*, *domain*, and *range* sets of S are given by

$$\operatorname{gph} S := \{(\lambda, x) \mid x \in S(\lambda)\}, \quad \operatorname{dom} S := \{\lambda \mid S(\lambda) \neq \emptyset\}, \quad \operatorname{rge} S := \operatorname{dom} S^{-1},$$

respectively.

Any product space $X \times Z$ of Banach spaces X and Z is considered with the supremum norm $\|(x, z)\| := \max\{\|x\|, \|z\|\}$.

Assumptions

Let $\{f(\cdot,\lambda) : X \to \mathbb{R}, \lambda \in \Lambda\}$ be a family of functions parameterized by $\lambda \in \Lambda \subset Z$. We look for sufficient conditions to ensure Aubin continuity of the solutions of the parameterized family of constrained minimization problems:

$$P(\lambda) \qquad \inf_{x\in K} f(x,\lambda),$$

where K is a given nonempty closed set in X.

For $\lambda \in \Lambda$, the (possibly empty) set of solutions of the minimization problem $P(\lambda)$ is denoted by

$$S(\lambda) := \left\{\overline{x} \in K : f(\overline{x},\lambda) = \inf_{x\in K} f(x,\lambda)\right\},$$

and its optimal value by

$$m(\lambda) := \inf_{x\in K} f(x,\lambda).$$

It is well known that even for smooth parameterized problem $P(\lambda)$ the solution $S : \Lambda \rightrightarrows X$ may fail Lipschitz continuity. For example, for $f(x,\lambda) = \frac{1}{4}x^4 - \lambda x$, where $x,\lambda \in \mathbb{R}$, and $K = [-1,1]$, we see that for $\lambda \in (-1,1)$ the solution is $S(\lambda) = \{\sqrt[3]{\lambda}\}$, and it is not Lipschitz continuous at $\lambda = 0$ (see Bonnans and Shapiro [22, Example 4.31]).

Hence, to establish Lipschitz behavior of S one needs something more than the standard requirements. We now turn to relevant analysis of $P(\lambda)$.

Definition 3.1.1. Let X and Z be Banach spaces. Let $U \subset X$, $V \subset Z$ be non-empty. We denote by $\mathfrak{L}^{\alpha,\beta}(U;V)$, $\alpha,\beta \in [0,1]$, the class of all functions $g : U \times U \times V \to \mathbb{R}$ such that there exists a constant $k_g > 0$ such that for all $x, x' \in U$ and all $\lambda, \lambda' \in V$,

$$|g(x,x',\lambda) - g(x,x',\lambda')| \le k_g \|x - x'\|^{\alpha} \|\lambda - \lambda'\|^{\beta}.$$

For example, $g \in \mathfrak{L}^{1,1}(U;V)$ means that $g(x,x',\cdot)$ is Lipschitz on V and its best Lipschitz constant $L(x,x')$ satisfies $L(x,x') \le k\|x - x'\|$ for some positive constant k and all $x, x' \in U$.

With the parameterized family of functions $\{f(\cdot,\lambda), \lambda \in \Lambda\}$ one may associate two difference functions: the function $f_1 : X \times X \times \Lambda \to \mathbb{R}$ defined by

$$f_1(x,x',\lambda) := f(x,\lambda) - f(x',\lambda),$$

and the function $f_2 : \Lambda \times \Lambda \times X \to \mathbb{R}$ defined by

$$f_2(\lambda,\lambda',x) := f(x,\lambda) - f(x,\lambda').$$

The above notions are linked through the following:

Proposition 3.1.2. For any $U \subset X$, $V \subset Z$ the function $f_1 \in \mathfrak{L}^{\alpha,\beta}(U;V)$ if and only if $f_2 \in \mathfrak{L}^{\beta,\alpha}(V;U)$.

Proof. Let $f_1 \in \mathfrak{L}^{\alpha,\beta}(U;V)$. Take any $x, x' \in U$, and any $\lambda, \lambda' \in V$. Since

$$\begin{aligned} f_2(\lambda,\lambda',x) - f_2(\lambda,\lambda',x') &= [f(x,\lambda) - f(x,\lambda')] - [f(x',\lambda) - f(x',\lambda')] \\ &= [f(x,\lambda) - f(x',\lambda)] - [f(x,\lambda') - f(x',\lambda')] \\ &= f_1(x,x',\lambda) - f_1(x,x',\lambda') \le k_{f_1}\|x - x'\|^{\alpha}\|\lambda - \lambda'\|^{\beta}, \end{aligned}$$

one can take $k_{f_2} := k_{f_1}$ to conclude that $f_2 \in \mathfrak{L}^{\beta,\alpha}(V;U)$. The proof of the other direction is similar. □

We are ready to present the sufficient condition for Aubin continuity of the solution map.

Given a $(\bar{\lambda}, \bar{x}) \in \operatorname{gph} S$, consider the following local assumption A at $(\bar{\lambda}, \bar{x})$:

$$(A) \qquad \begin{cases} \text{there exist neighborhoods } U \text{ of } \bar{x} \text{ and } V \text{ of } \bar{\lambda}, \text{ such that} \\ 1.\ \ S(\lambda) \cap U \neq \varnothing \text{ for all } \lambda \in \Lambda \cap V; \\ \text{and there exist constants } c > 0 \text{ and } \alpha \in [0,1] \text{ such that} \\ 2.\ \ f(x,\lambda') \ge m(\lambda') + cd^{1+\alpha}(x, S(\lambda')),\ \forall \lambda,\lambda' \in \Lambda \cap V,\ \forall x \in S(\lambda) \cap U; \\ 3.\ \ f_1 \in \mathfrak{L}^{\alpha,1}(K; \Lambda \cap V). \end{cases}$$

By Proposition 3.1.2 it is clear that assumption $A3$ could be replaced with $f_2 \in \mathfrak{L}^{1,\alpha}(\Lambda \cap V; K)$.

It is clear that $A1$ implies that $m(\lambda)$ is finite for all $\lambda \in \Lambda \cap V$.

In the case $\alpha = 1$, assumption $A2$ can be considered as a relaxed (in x) uniform (in λ') version of the so-called second-order growth condition. One says that the *second-order growth condition* holds for the problem

$$\inf_{x \in K} f(x)$$

in a neighborhood N of the solution set S_0, if there exists a constant $c > 0$ such that

$$f(x) \ge \inf_K f + cd^2(x, S_0), \quad \forall x \in K \cap N. \tag{3.3}$$

This condition is involved in a number of works (see Bonnans and Ioffe [20], Bonnans and Shapiro [21, 22], Klatte and Henrion [101], Shapiro [154]) in order to ensure Lipschitz stability of the solution map S of the constrained minimization problem. Let us recall that S is said to be *Lipschitz stable* or, equivalently, *upper Lipschitz* at a point $\bar{\lambda} \in \Lambda$, if there exist a constant $\kappa > 0$ and a neighborhood V of $\bar{\lambda}$ such that it holds that

$$e(S(\lambda), S(\bar{\lambda})) \le \kappa \left\|\lambda - \bar{\lambda}\right\|, \quad \forall \lambda \in \Lambda \cap V.$$

Let us note that Lipschitz stability is a point property: it holds for S at a fixed point $\bar{\lambda}$, while Aubin continuity we wish to obtain, is a local property, and it holds uniformly at all points μ in some neighborhood V of the referenced point $\bar{\lambda}$. Obviously, Aubin continuity of S near $(\bar{\lambda}, \bar{x})$ implies Lipschitz stability of $S \cap U$ at $\bar{\lambda}$ while the opposite implication is not always true.

A stronger version of uniform second-order growth condition than $A2$ with $\alpha = 1$ is given in Bonnans and Shapiro [22, Definition 5.16]. It implies single-valuedness and local Lipschitz continuity of S (cf. Bonnans and Shapiro [22, Theorem 5.17 and Remark 5.19]). In contrast, assumption $A2$ does not imply neither single-valuedness nor local boundedness of the solution map S (see Example 3.1.6).

We would now give a few examples of parameterized families of functions $\{f(\cdot, \lambda), \lambda \in \Lambda\}$ for which A holds, in this way showing the consistency of our main assumption.

Obviously, the compactness of K and lower semicontinuity of $f(\cdot, \lambda)$ on K are sufficient to ensure $A1$ (note that weak compactness and weak lower semicontinuity would do just as well).

$A2$ with $\alpha = 1$ is satisfied at any $(\bar{\lambda}, \bar{x}) \in \operatorname{gph} S$ provided that, for example, $U = X$, $V = Z$, and $K \subset X$ is a nonempty closed convex set, the functions $f(\cdot, \lambda)$ are lower semicontinuous, and uniformly on $\lambda \in \Lambda$ strongly convex on K, that is, for some constant $c > 0$ the inequality

$$f(tx' + (1-t)x'', \lambda) \le t f(x', \lambda) + (1-t) f(x'', \lambda) - ct(1-t)\|x' - x''\|^2$$

holds for every $t \in [0, 1]$, every $x', x'' \in K$ and every $\lambda \in \Lambda$.

Lemma 3.1.3 below provides examples of parameterized families of functions satisfying $A3$. However, we need a few more definitions before stating this lemma.

Recall that the *Clarke generalized derivative* of Lipschitz function $f : X \to \mathbb{R}$ at $\bar{x} \in X$ in direction $h \in X$ is

$$f^{\circ}(\bar{x}; h) := \limsup_{\substack{x \to \bar{x} \\ t \downarrow 0}} \frac{f(x + th) - f(x)}{t},$$

and the *Clarke subdifferential* at $\bar{x}$ is the nonempty w^* compact set

$$\partial f(\bar{x}) := \{x^* \in X^* : \langle x^*, h\rangle \le f^{\circ}(\bar{x}; h), \ \forall h \in X\};$$

see Clarke [44]. It is well known that for any $h \in X$ there exists some $x^* \in \partial f(\bar{x})$ such that

$$\langle x^*, h\rangle = f^{\circ}(\bar{x}; h).$$

Lipschitz function $f : U \to \mathbb{R}$ is said to be *regular* on an open set $U \subset X$ if for any $h \in X$ and any $\bar{x} \in U$ its *directional derivative*

$$f'(\bar{x}; h) := \lim_{t \downarrow 0} \frac{f(\bar{x} + th) - f(\bar{x})}{t}$$

exists and is equal to $f^\circ(\overline{x}; h)$. Convex continuous functions and strictly differentiable functions are examples of regular functions.

Let $f(x, \lambda)$ be Lipschitz on each variable bivariate function. Denote by $f_x^\circ(\overline{x}, \overline{\lambda}; h)$ and by $f_x'(\overline{x}, \overline{\lambda}; h)$ the generalized derivative and the directional derivative of $f(\cdot, \overline{\lambda})$ at $\overline{x}$ in direction h, respectively. Also, denote by $\partial_x f(\overline{x}, \overline{\lambda})$ the partial Clarke subdifferential of $f(\cdot, \overline{\lambda})$ at $\overline{x}$, and by $\partial_\lambda f(\overline{x}, \overline{\lambda})$ the partial Clarke subdifferential of $f(\overline{x}, \cdot)$ at $\overline{\lambda}$.

Lemma 3.1.3. Let $(\overline{\lambda}, \overline{x}) \in \operatorname{gph} S$ and let $U \subset X$ and $V \subset Z$ be convex neighborhoods of K and $\overline{\lambda}$, respectively. Consider the conditions:

$$(F1)\qquad \begin{cases} \text{for } \lambda \in \Lambda \cap V,\ f(\cdot, \lambda) \text{ is Lipschitz and regular on } U \text{ and} \\ \partial_x f(x, \cdot) : \Lambda \cap V \to X^* \text{ is a } k\text{-Lipschitz map on } \Lambda \cap V \\ \text{with } k \text{ that does not depend on } x \in U, \end{cases}$$

$$(F2)\qquad \begin{cases} \text{for } x \in K,\ f(x, \cdot) \text{ is Lipschitz and regular on V and} \\ \partial_\lambda f(\cdot, \lambda) : K \to Z^* \text{ is a } k\text{-Lipschitz map on } K \\ \text{with } k \text{ that does not depend on } \lambda \in V. \end{cases}$$

If f satisfies $F1$ or $F2$, then $A3$ holds with $\alpha = 1$.

Proof. Let f satisfy $F1$. Fix $x, y \in K$ and $\lambda, \mu \in \Lambda \cap V$. Consider the function $r(t) := f(y + t(x - y), \lambda)$ which is well-defined on an open interval I containing $[0, 1]$. Since the function $f(\cdot, \lambda)$ is assumed to be Lipschitz on U, we have that r is Lipschitz on I. By Rademacher's theorem, for almost all $t \in [0, 1]$ there exists

$$\begin{aligned} r'(t) &= \lim_{s \to 0} \frac{r(t + s) - r(t)}{s} = \lim_{s \downarrow 0} \frac{f(y + t(x - y) + s(x - y), \lambda) - f(y + t(x - y), \lambda)}{s} \\ &= f_x'(y + t(x - y), \lambda; x - y) = f_x^\circ(y + t(x - y), \lambda; x - y). \end{aligned}$$

The last equality holds because $f(\cdot, \lambda)$ is regular on U.

Hence,

$$f(x, \lambda) - f(y, \lambda) = r(1) - r(0) = \int_0^1 r'(t)\, dt = \int_0^1 f_x^\circ(y + t(x - y), \lambda; x - y)\, dt. \tag{3.4}$$

Similarly,

$$f(x, \mu) - f(y, \mu) = \int_0^1 f_x^\circ(y + t(x - y), \mu; x - y)\, dt. \tag{3.5}$$

There exists $x_\lambda^*(t) \in \partial_x f(y + t(x - y), \lambda)$ such that $f_x^\circ(y + t(x - y), \lambda; x - y) = \langle x_\lambda^*(t), x - y \rangle$, so (3.4) becomes

$$f(x, \lambda) - f(y, \lambda) = \int_0^1 \langle x_\lambda^*(t), x - y \rangle\, dt. \tag{3.6}$$

Since $x_\lambda(t) \in \partial_x f(y + t(x - y), \lambda)$ and the multivalued map $\partial_x f(x, \cdot) : \Lambda \cap V \to X^*$ is k–Lipschitz continuous with w^* compact images, there is $x^*_\mu(t) \in \partial_x f(y + t(x - y), \mu)$ such that $\|x^*_\lambda(t) - x^*_\mu(t)\| \le k\|\lambda - \mu\|$. Note that k does not depend on either $t \in [0, 1]$ or $x, y \in U$.

Let us use these for estimating $f_1(x, y, \lambda) - f_1(x, y, \mu)$.

From (3.6) we get

$$\begin{aligned} f(x,\lambda) - f(y,\lambda) &= \int_0^1 \langle x^*_\lambda(t) - x^*_\mu(t), x - y\rangle\, dt + \int_0^1 \langle x^*_\mu(t), x - y\rangle\, dt \\ &\le \int_0^1 \|x^*_\lambda(t) - x^*_\mu(t)\|\, \|x - y\|\, dt + \int_0^1 \langle x^*_\mu(t), x - y\rangle\, dt \\ &\le k\|\lambda - \mu\|\, \|x - y\| + \int_0^1 \langle x^*_\mu(t), x - y\rangle\, dt. \end{aligned}$$

Since $x^*_\mu(t) \in \partial_x f(y + t(x - y), \mu)$, it holds that $\langle x^*_\mu(t), x - y\rangle \le f^\circ_x(y + t(x - y), \mu; x - y)$, and by (3.5) we have

$$\int_0^1 \langle x^*_\mu(t), x - y\rangle\, dt \le \int_0^1 f^\circ_x(y + t(x - y), \mu; x - y)\, dt = f(x, \mu) - f(y, \mu).$$

Hence,

$$f(x, \lambda) - f(y, \lambda) \le f(x, \mu) - f(y, \mu) + k\|\lambda - \mu\|\|x - y\|;$$

that is, $f_1(x, y, \lambda) \le f_1(x, y, \mu) + k\|\lambda - \mu\|\|x - y\|$, or

$$f_1(x, y, \lambda) - f_1(x, y, \mu) \le k\|\lambda - \mu\|\|x - y\|,$$

which means that $f_1 \in \mathfrak{L}^{1,1}(K; \Lambda\cap V)$.

If f satisfies $F2$, then by the same reasoning one obtains that $f_2 \in \mathfrak{L}^{1,1}(\Lambda\cap V; K)$ and by Proposition 3.1.2, $A3$ holds. □

It is interesting to note here that the regularity (in particular, the differentiability) can be asked for the argument x as in $F1$, or for the parameter λ as in $F2$.

It is clear that both $F1$ and $F2$ hold whenever $f \in C^{1,1}(U \times V)$.

Lipschitz-like continuity of the solution map

Here we prove that given $(\overline{\lambda}, \overline{x}) \in \operatorname{gph} S$, assumption A is sufficient to ensure Aubin continuity of the solution map S near $(\overline{\lambda}, \overline{x})$.

Proposition 3.1.4. Assume that X and Z are Banach spaces and consider a family of constraint minimization problems $P(\lambda)$ parameterized by $\lambda \in \Lambda$, a nonempty subset of Z.

If for some $(\bar{\lambda}, \bar{x}) \in \operatorname{gph} S$ assumption A holds, then

$$e(S(\lambda) \cap U, S(\mu)) \leq \frac{k_{f_1}}{c} \|\lambda - \mu\|, \quad \forall \lambda, \mu \in \Lambda \cap V, \tag{3.7}$$

and the solution map S is Aubin continuous near $(\bar{\lambda}, \bar{x}) \in \operatorname{gph} S$.

Proof. Take any $\lambda \in \Lambda \cap V$ and any $x_\lambda \in S(\lambda) \cap U$ (which is a nonempty set thanks to $A1$). By $A2$, for arbitrary $\mu \in \Lambda \cap V$

$$f(x_\lambda, \mu) \geq m(\mu) + c d^{1+\alpha}(x_\lambda, S(\mu)). \tag{3.8}$$

Since by $A1$ the set $S(\mu) \cap U$ is nonempty, for any $\varepsilon > 0$ there exists some $x_\mu^\varepsilon \in S(\mu)$ such that

$$\|x_\lambda - x_\mu^\varepsilon\| \leq d(x_\lambda, S(\mu)) + \varepsilon. \tag{3.9}$$

As $x_\mu^\varepsilon \in S(\mu)$ we have $m(\mu) = f(x_\mu^\varepsilon, \mu)$ and inequality (3.8) reads

$$f(x_\lambda, \mu) \geq f(x_\mu^\varepsilon, \mu) + c d^{1+\alpha}(x_\lambda, S(\mu)). \tag{3.10}$$

Since $x_\lambda \in S(\lambda)$, we have that

$$f(x_\mu^\varepsilon, \lambda) \geq f(x_\lambda, \lambda). \tag{3.11}$$

By adding (3.10) and (3.11) and rearranging, we obtain

$$[f(x_\lambda, \mu) - f(x_\mu^\varepsilon, \mu)] - [f(x_\lambda, \lambda) - f(x_\mu^\varepsilon, \lambda)] \geq c d^{1+\alpha}(x_\lambda, S(\mu)).$$

That is,

$$f_1\left(x_\lambda, x_\mu^\varepsilon, \mu\right) - f_1\left(x_\lambda, x_\mu^\varepsilon, \lambda\right) \geq c d^{1+\alpha}(x_\lambda, S(\mu)). \tag{3.12}$$

Using $A3$, that is, $f_1 \in \mathfrak{L}^{\alpha,1}(K; \Lambda \cap V)$, we estimate the left-hand side of (3.12):

$$f_1\left(x_\lambda, x_\mu^\varepsilon, \mu\right) - f_1\left(x_\lambda, x_\mu^\varepsilon, \lambda\right) \leq k_{f_1} \|x_\lambda - x_\mu^\varepsilon\|^\alpha \|\lambda - \mu\|.$$

Hence, we have that $k_{f_1} \|\lambda - \mu\| \, \|x_\lambda - x_\mu^\varepsilon\|^\alpha \geq c d^{1+\alpha}(x_\lambda, S(\mu))$. From this and (3.9) it follows that

$$k_{f_1} \|\lambda - \mu\| [d(x_\lambda, S(\mu)) + \varepsilon]^\alpha \geq c d^{1+\alpha}(x_\lambda, S(\mu)).$$

Letting $\varepsilon \downarrow 0$ and then dividing by $d^\alpha(x_\lambda, S(\mu)) > 0$ (if $= 0$ the inequality below is trivial), we obtain $k_{f_1} \|\lambda - \mu\| \geq c d(x_\lambda, S(\mu))$, or

$$d(x_\lambda, S(\mu)) \leq \frac{k_{f_1}}{c} \|\lambda - \mu\|.$$

As x_λ was an arbitrary point in $S(\lambda) \cap U$, the latter yields

$$e(S(\lambda) \cap U, S(\mu)) \leq \frac{k_{f_1}}{c} \|\lambda - \mu\|,$$

completing the proof. □

Examples and corollaries

The following is a basic example of non-smooth parameterized minimization problem with Lipschitz continuous solution map with unbounded values. We show that it is within the scope of Proposition 3.1.4.

Example 3.1.5. Let $K = \mathbb{R}^2$ and

$$f(x_1, x_2, \lambda) := |x_1 - x_2 - \lambda|,$$

$x_1, x_2, \lambda \in \mathbb{R}$. Consider the parameterized family of unconstrained minimization problems over the plane

$$P(\lambda) \qquad \inf_{x_1, x_2} f(x_1, x_2, \lambda).$$

Then the solution map $S : \lambda \rightrightarrows S(\lambda)$ is Lipschitz continuous.

Proof. Obviously, for any $\lambda \in \mathbb{R}$ the solution set consists of a single line, i.e., $S(\lambda) = \{(x_1, x_2) : x_1 - x_2 = \lambda\}$. Moreover, for λ and μ the solution sets $S(\lambda)$ and $S(\mu)$ are parallel lines. The distance between $S(\lambda)$ and $S(\mu)$ is the distance from any point $(\bar{x}_1, \bar{x}_2) \in S(\lambda)$ to the line $x_1 - x_2 = \mu$ which is equal to $\frac{|\bar{x}_1 - \bar{x}_2 - \mu|}{\sqrt{2}} = \frac{|\lambda - \mu|}{\sqrt{2}}$, so the map S is Lipschitz continuous with Lipschitz constant $\frac{1}{\sqrt{2}}$.

Note that the sufficient condition A holds. Indeed

$A1$ holds with $U \equiv \mathbb{R}^2$;

$A2$ holds with $\alpha = 0$, $c = \sqrt{2}$, and $U = \mathbb{R}^2$, $V = \mathbb{R}$;

$A3$ holds because $f_1 \in \mathfrak{L}^{0,1}(\mathbb{R}^2, \mathbb{R})$ with $k_{f_1} = 2$.

The Lipschitz constant provided by Proposition 3.1.4 is $\frac{k_{f_1}}{c} = \sqrt{2}$. □

The next example shows that studying the generalized Euler equation may sometimes be inadequate for obtaining Aubin continuity of the solution map. This is because the set of the stationary points may be larger than the set of minima.

Example 3.1.6. Let $K = \mathbb{R}^2$ and

$$f(x_1, x_2, \lambda) := (x_1 + \lambda x_2 - 1)^2 (x_2 + \lambda x_1 + 1)^2,$$

$x_1, x_2, \lambda \in \mathbb{R}$. Consider the parameterized family of unconstrained minimization problems over the plane

$$P(\lambda) \qquad \inf_{x_1, x_2} f(x_1, x_2, \lambda).$$

Then at the point $\bar{\lambda} = 1$ the set of solutions $S(\lambda)$ is smaller than the set of stationary points $St(\lambda) := \{x \in \mathbb{R}^2 : 0 \in \nabla_x f(x, \lambda)\}$. Moreover, the map S is Aubin continuous near any point in his graph while St is not Aubin continuous near the point $(\bar{\lambda}, \bar{x}) \in \operatorname{gph} St$ where $\bar{\lambda} = 1$ and $\bar{x} = (0, 0)$.

Proof. Straightforward computations show that for any $\lambda \in \mathbb{R}$ the solution set

$$S(\lambda) = \{(x_1, x_2) : x_2 + \lambda x_1 = -1, \text{ or } x_1 + \lambda x_2 = 1\}$$

is the union of two lines—the line $p_1(\lambda)$ with equation $x_2 + \lambda x_1 = -1$ and the line $p_2(\lambda)$ with equation $x_1 + \lambda x_2 = 1$. Because of

$$\nabla_x f(x, \lambda) = [2(x_1 + \lambda x_2 - 1)(x_2 + \lambda x_1 + 1)(2\lambda x_1 + (1 + \lambda^2)x_2 + 1 - \lambda),\\ 2(x_1 + \lambda x_2 - 1)(x_2 + \lambda x_1 + 1)((1 + \lambda)^2 x_1 + 2\lambda x_2 + \lambda - 1)],$$

the set of the stationary points at $\overline{\lambda} = 1$ consists of three parallel lines

$$St(1) = \{(x_1, x_2) : x_1 + x_2 = 1, \text{ or } x_1 + x_2 = -1, \text{ or } x_1 + x_2 = 0\},$$

while for $\lambda \neq 1$, $St(\lambda) \equiv S(\lambda)$.

It is not difficult to see that S is Aubin continuous near an arbitrary point $(\lambda, x) \in \operatorname{gph} S$ (we note, by the way, that S is not Lipschitz continuous). Indeed, fix $\tilde{\lambda} \in \mathbb{R}$ and take $\tilde{x} = (\tilde{x}_1, \tilde{x}_2) \in S(\tilde{\lambda}) = p_1(\tilde{\lambda}) \cup p_2(\tilde{\lambda})$. Obviously, $\tilde{x} \neq 0$.

Take λ such that $|\lambda - \tilde{\lambda}| < 1/2$. If $\tilde{x} \in p_1(\tilde{\lambda})$, then

$$d(\tilde{x}, S(\lambda)) \leq d(\tilde{x}, p_1(\lambda)) = \frac{\left|\left(\lambda - \tilde{\lambda}\right)\tilde{x}_1\right|}{\sqrt{1 + \lambda^2}} \leq \left|\lambda - \tilde{\lambda}\right| |\tilde{x}_1|,$$

and if $\tilde{x} \in p_2(\tilde{\lambda})$, then

$$d(\tilde{x}, S(\lambda)) \leq d(\tilde{x}, p_2(\lambda)) = \frac{\left|\left(\lambda - \overline{\lambda}\right)\tilde{x}_2\right|}{\sqrt{1 + \lambda^2}} \leq \left|\lambda - \tilde{\lambda}\right| |\tilde{x}_2|,$$

which yields

$$d(\tilde{x}, S(\lambda)) \leq \left|\lambda - \tilde{\lambda}\right| \max\{|\tilde{x}_1|, |\tilde{x}_2|\} \leq \left|\lambda - \tilde{\lambda}\right| \|\tilde{x}\| < \|\tilde{x}\|/2.$$

This implies that for all λ such that $|\lambda - \tilde{\lambda}| < 1/2$ the intersection of $S(\lambda)$ with the neighborhood $U := \tilde{x} + \|\tilde{x}\|B^\circ$ is nonempty.

Take $x = (x_1, x_2) \in S(\lambda) \cap U$ and μ such that $|\mu - \tilde{\lambda}| < 1/2$. Similarly we get

$$d(x, S(\mu)) \leq |\lambda - \mu| \|x\| \leq |\lambda - \mu| [\|x - \tilde{x}\| + \|\tilde{x}\|] \leq 2\|\tilde{x}\| |\lambda - \mu|.$$

Hence,

$$e(S(\lambda) \cap U, S(\mu)) \leq 2\|\tilde{x}\| |\lambda - \mu|, \quad \forall \lambda, \mu \in \tilde{\lambda} + \frac{1}{2}B^\circ,$$

which means that S is Aubin continuous near $(\tilde{\lambda}, \tilde{x}) \in \operatorname{gph} S$.

In contrast, St is not Aubin continuous near the point, $(\overline{\lambda}, \overline{x}) \in \operatorname{gph} St$ where $\overline{\lambda} = 1$ and $\overline{x} = (0, 0)$. Indeed, if St is Aubin continuous near that point, then $d(\overline{x}, St(\lambda))$ tends to zero as λ tends to 1. But the distance

$$d(\overline{x}, St(\lambda)) = \min\{d(\overline{x}, p_1(\lambda)), d(\overline{x}, p_2(\lambda))\} = \frac{1}{\sqrt{1+\lambda^2}}$$

tends to $\frac{1}{\sqrt{2}}$ as λ tends to 1, which means that St is not Aubin continuous near $(\overline{\lambda}, \overline{x}) \in \operatorname{gph} St$. □

As an immediate consequence of Proposition 3.1.4 we get the following.

Corollary 3.1.7. Let for the parameterized family of minimization problems $P(\lambda)$ the following assumption hold

$$(A') \qquad \begin{cases} \text{for all } \lambda \in \Lambda, \text{ all } x \in K, \text{ and some } c > 0 \\ 1.\ S(\lambda) \neq \emptyset; \\ 2.\ f(x, \lambda) \geq m(\lambda) + cd^2(x, S(\lambda)); \\ 3.\ f \in C^{1,1}(X \times Z). \end{cases}$$

Then the solution map $S : \Lambda \rightrightarrows X$ is Lipschitz continuous on Λ.

In a Banach space X with separable dual X^* the notion of a second-order subdifferential for a function $f \in C^{1,1}(X)$ is introduced in Georgiev and Zlateva [78] (see also the previous work Hiriart-Urruty, Strodiot and Nguyen [85] for the finite dimensional case). For any $x \in X$ the *second-order subdifferential* $\partial^2 f(x)$ of f at x is a nonempty, convex, and w^* compact set in $\mathcal{L}(X \times X)$ (the Banach space of all bilinear continuous functionals $M : X \times X \to \mathbb{R}$ with the norm $\|M\| := \sup_{\|h_1\|=\|h_2\|=1} |M[h_1, h_2]|$), which is singleton exactly when f is twice strictly Gâteaux differentiable at x.

Setting a simple condition on the second subdifferential is sufficient to get a family of functions satisfying assumption A' in the above corollary.

Indeed, let X be a Banach space with separable dual. Let in the parameterized family of minimization problems $P(\lambda)$, $f \in C^{1,1}(X \times Z)$, and let the constraint set K be closed and convex. If there exist $c > 0$ with

$$(3.13) \qquad \langle M(y-x), y-x \rangle \geq c\|y-x\|^2 \quad \text{for all } \lambda \in \Lambda,\ x, y \in K,\ M \in \partial^2 f(\cdot, \lambda)(x),$$

then the solution map $S : \Lambda \to X$ will be single-valued and Lipschitz continuous on Λ.

It is easily seen that (3.13) implies uniform on $\lambda \in \Lambda$ strong convexity of $f(\cdot, \lambda)$ on K. By this and continuity of $f(\cdot, \lambda)$, for every λ the infimum of $f(\cdot, \lambda)$ is attained at unique $x_\lambda \in K$ and $A'1$ holds.

For any $x \in K$ and $\lambda \in \Lambda$ there exists some $z_\lambda \in K$ and $M_{z_\lambda} \in \partial^2 f(\cdot, \lambda)(z_\lambda)$ with

$$f(x, \lambda) = f(x_\lambda, \lambda) + \langle \nabla_x f(x_\lambda, \lambda), x - x_\lambda \rangle + \frac{1}{2} \langle M_{z_\lambda}(x - x_\lambda), x - x_\lambda \rangle$$

(see Georgiev and Zlateva [78]). Since x_λ is a minimum point for $f(\cdot, \lambda)$ on K and K is convex, then for all $x \in K$, $\langle \nabla_x f(x_\lambda, \lambda), x - x_\lambda \rangle \geq 0$ and from above equality and (3.13)

$$f(x, \lambda) \geq m(\lambda) + \frac{1}{2} c \|x - x_\lambda\|^2,$$

so $A'2$ holds.

We will use Corollary 3.1.7 to obtain existence and Lipschitz continuity of the optimal solution for a linearly perturbed optimization problem, assuming a slightly weaker version (see (3.14) below) of the uniform second-order growth condition (Definition 5.19 in Bonnans and Shapiro [22]), and $C^{1,1}$ data. In this way we extend Bonnans and Shapiro [22, Theorem 5.17] (see also Bonnans and Shapiro [22, Remark 5.19]), where C^2 data are assumed.

Recall that the Banach space X has *Radon–Nikodym property* (*RNP*) if for every bounded set C and every $\varepsilon > 0$, there exists an $x \in C$ that does not belong to the closed convex hull of $C \setminus \{x + \varepsilon B^\circ\}$. All Banach spaces which have separable dual and all reflexive Banach spaces have RNP. In Diestel and Uhl [59, p. 157] there is a long list of equivalent definitions of RNP. A good introductory survey on RNP is Diestel and Uhl [60].

An efficient tool in dealing with minimization problems on Banach space X with RNP is *Stegall's variational principle* [157] (see also Phelps [133, Theorem 5.15]): Let $C \subset X$ be a non-empty closed and bounded convex set and let $f : C \to \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous function, bounded below on C, then for every $\varepsilon > 0$, there exists $x^* \in X^*$ with $\|x^*\| \leq \varepsilon$ such that $f + x^*$ attains its strong minimum on C. Let us remind that $x_0 \in C$ is said to be a *strong minimum* for function $g : C \to \mathbb{R} \cup \{+\infty\}$ on the set C if $g(x_0) = \inf_C g$ and $\|x_n - x_0\| \to 0$ whenever $g(x_n) \to g(x_0)$.

Corollary 3.1.8. Let the Banach space X have Radon–Nikodym property. Consider a parameterized family of minimization problems $P(\lambda)$, where the parameter space is X^* and $f : X \times X^* \to \mathbb{R}$ is defined by $f(x, \lambda) := f(x) + \langle \lambda, x \rangle$.

Assume that the constraint set K is closed and convex, $f \in C^{1,1}(X)$, and $S(0)$ is nonempty.

Suppose that there exist neighborhood V of the origin 0 of X^* and a constant $c > 0$ such that for all $\lambda \in V$ and all $x_\lambda \in S(\lambda)$ it holds that

$$f(x, \lambda) \geq f(x_\lambda, \lambda) + c\|x - x_\lambda\|^2, \quad \forall x \in K. \tag{3.14}$$

Then there exists a neighborhood W of the origin 0 of X^* such that $S(\lambda)$ is single-valued and Lipschitz continuous on W.

Proof. From (3.14) it is clear that $S(\lambda)$ contains at most one point for $\lambda \in V$.

We will show that $S(\lambda)$ is nonempty for λ belonging to some neighborhood of 0. Fix $\gamma > 0$ such that $W := 2\gamma B^\circ \subset V$.

Given $\lambda \in \gamma B^\circ$, let ε_k be a sequence of positive numbers less than γ, tending to zero.

Thanks to (3.14) with $\lambda = 0$, $f(\cdot, \lambda)$ is bounded below on K.

If K is bounded we could apply directly Stegall's variational principle for the function $f(\cdot, \lambda) : K \to \mathbb{R}$ and ε_k to find $x_k^* \in X^*$ with $\|x_k^*\| \le \varepsilon_k$ and a strong minimum x_k of $f(\cdot, \lambda) + x_k^*$ on K.

If K is not bounded, a variant of Stegall's variational principle still holds thanks to (3.14). Indeed, (3.14) for $\lambda = 0$ reads

$$f(x) \ge f(x_0) + c\|x - x_0\|^2, \quad \forall x \in K,$$

which yields that for all $x \in K$,

$$\begin{aligned} f(x, \lambda) &\ge f(x_0) + \langle \lambda, x\rangle + c\|x - x_0\|^2 = f(x_0, \lambda) + \langle \lambda, x - x_0\rangle + c\|x - x_0\|^2 \\ &\ge f(x_0, \lambda) + \|x - x_0\|[c\|x - x_0\| - \|\lambda\|] \\ &\ge f(x_0, \lambda) + \|x - x_0\|[c\|x - x_0\| - \gamma]. \end{aligned} \tag{3.15}$$

Set $r := \frac{3\gamma}{c}$. Now, we apply Stegall's variational principle for the function $f(\cdot, \lambda)$ on the closed bounded set $K \cap \{x_0 + rB\}$ and ε_k. Thus, there exists $x_k^* \in X^*$, $\|x_k^*\| < \varepsilon_k$, and a point $x_k \in K \cap \{x_0 + rB\}$ such that $f(\cdot, \lambda) + x_k^*$ attains a strong minimum on $K \cap \{x_0 + rB\}$ at x_k. Moreover, x_k is a strong minimum of $f(\cdot, \lambda) + x_k^*$ on K. Indeed, if we assume that $x \in K$ is such that

$$f(x, \lambda) + \langle x_k^*, x\rangle \le f(x_k, \lambda) + \langle x_k^*, x_k\rangle = \inf_{K \cap \{x_0 + rB\}} f(\cdot, \lambda) + x_k^* \le f(x_0, \lambda) + \langle x_k^*, x_0\rangle,$$

then by (3.15) we will have

$$\|x - x_0\|[c\|x - x_0\| - \gamma] \le \|x_k^*\|\|x - x_0\| \le \varepsilon_k\|x - x_0\| < \gamma\|x - x_0\|,$$

or

$$\|x - x_0\| \le \frac{2\gamma}{c} < r,$$

which means that $x \in x_0 + rB$ and clearly entails $x = x_k$.

However, in both cases for any k we found $x_k^* \in X^*$ with $\|x_k^*\| \le \varepsilon_k$ and unique $x_k \in K$ satisfying

$$f(x) + \langle \lambda + x_k^*, x\rangle \ge f(x_k) + \langle \lambda + x_k^*, x_k\rangle, \quad \forall x \in K.$$

This means that $S(\lambda + x_k^*) = \{x_k\}$ and since $\lambda + x_k^* \in V$, (3.14) reads

$$f(x) + \langle \lambda + x_k^*, x\rangle \ge f(x_k) + \langle \lambda + x_k^*, x_k\rangle + c\|x - x_k\|^2, \quad \forall x \in K. \tag{3.16}$$

Substitute $x = x_n$ and rearrange to obtain

$$f(x_n) - f(x_k) \geq \langle \lambda + x_k^*, x_k - x_n \rangle + c\|x_n - x_k\|^2.$$

Also, swapping k and n we get

$$f(x_k) - f(x_n) \geq \langle \lambda + x_n^*, x_n - x_k \rangle + c\|x_n - x_k\|^2.$$

Adding the above two, we obtain $2c\|x_n - x_k\|^2 \leq \langle x_k^* - x_n^*, x_n - x_k \rangle \leq (\|x_k^*\| + \|x_n^*\|)\|x_n - x_k\|$. That is, $2c\|x_n - x_k\| \leq \varepsilon_k + \varepsilon_n$, which means that x_k is a Cauchy sequence. Let $x_\lambda \in K$ be its limit. Passing to limit in (3.16) we see that $x_\lambda \in S(\lambda)$.

A straightforward application of Corollary 3.1.7 completes the proof. □

3.1.2 Parameterized minimax problem

In this subsection we study the behavior of the saddle points set of a parameterized family of minimax problems.

Preliminaries and statement of the problem

Let X and Y be Banach spaces, and let $\{f(\cdot, \cdot, \lambda) : X \times Y \to \mathbb{R}, \lambda \in \Lambda\}$ be a family of functions defined on the product space $X \times Y$, parameterized by $\lambda \in \Lambda \subset Z$.

Let us consider the parameterized family of minimax problems

$$M(\lambda) \qquad \inf_{x \in K} \sup_{y \in L} f(x, y, \lambda),$$

where the constraints are nonempty closed sets $K \subset X$ and $L \subset Y$. Denote the optimal value of $M(\lambda)$ by $\mathfrak{m}(\lambda)$ and recall that the (possibly empty) set of saddle points of $f(\cdot, \cdot, \lambda)$ on $K \times L$ is given by (3.1).

For a set $C \subset X \times Y$ we denote by $\pi_X C$ and $\pi_Y C$ the *canonical projections* of C on the spaces X and Y, respectively. More precisely, $x \in \pi_X C$ whenever there exists some $y \in Y$ with $(x, y) \in C$ and $y \in \pi_Y C$ whenever there exists some $x \in X$ with $(x, y) \in C$.

It is well known that the saddle point set is a product set; that is,

$$\mathcal{S}(\lambda) = \pi_X \mathcal{S}(\lambda) \times \pi_Y \mathcal{S}(\lambda). \tag{3.17}$$

To the parameterized family of functions $\{f(\cdot, \cdot, \lambda), \lambda \in \Lambda\}$ one naturally associates three difference functions:

$$\begin{aligned}
\mathfrak{f}_1(x, x', \lambda, y) &:= f(x, y, \lambda) - f(x', y, \lambda), \\
\mathfrak{f}_2(y, y', \lambda, x) &:= f(x, y, \lambda) - f(x, y', \lambda), \\
\mathfrak{f}_3(\lambda, \lambda', x, y) &:= f(x, y, \lambda) - f(x, y, \lambda').
\end{aligned}$$

By analogy with Definition 3.1.1 we write $\mathfrak{f}_1 \in \mathfrak{L}_W^{\alpha,\beta}(U;V)$ whenever the functions $f_1^y(x,x',\lambda) := \mathfrak{f}_1(x,x',\lambda,y)$ are such that for all $y \in W$, $f_1^y \in \mathfrak{L}^{\alpha,\beta}(U;V)$, and $\sup_{y\in W} k_{f_1^y}$ is finite. We set $k_{\mathfrak{f}_1} := \sup_{y\in W} k_{f_1^y}$.

Easy computations as those done in Proposition 3.1.2 show that $\mathfrak{f}_1 \in \mathfrak{L}_W^{\alpha,\beta}(U;V)$ exactly when $\mathfrak{f}_3 \in \mathfrak{L}_W^{\beta,\alpha}(V;U)$ and that $\mathfrak{f}_2 \in \mathfrak{L}_U^{\alpha,\beta}(W;V)$ exactly when $\mathfrak{f}_3 \in \mathfrak{L}_U^{\beta,\alpha}(V;W)$.

Now we are ready to state the sufficient condition for Aubin continuity of the saddle points map $\mathcal{S} : \Lambda \rightrightarrows X \times Y$.

Let $(\bar{\lambda},\bar{x},\bar{y}) \in \operatorname{gph}\mathcal{S}$. We set the following local assumption $\mathcal{A}$ at $(\bar{\lambda},\bar{x},\bar{y})$:

$$(\mathcal{A}) \qquad \begin{cases} \text{there exist neighborhoods } U \text{ of } \bar{x},\ W \text{ of } \bar{y}, \text{ and } V \text{ of } \bar{\lambda}, \text{such that} \\ 1.\ \ \mathcal{S}(\lambda) \cap [U \times W] \neq \varnothing \text{ for all } \lambda \in \Lambda \cap V; \\ \text{and there exist constants } c > 0 \text{ and } \alpha \in [0,1] \text{ such that} \\ 2.\ \ f(x,y',\lambda') \geq \mathfrak{m}(\lambda') + c d^{1+\alpha}(x, \pi_X \mathcal{S}(\lambda')), \\ \quad\ f(x',y,\lambda') \leq \mathfrak{m}(\lambda') - c d^{1+\alpha}(y, \pi_Y \mathcal{S}(\lambda')), \\ \quad\ \forall \lambda,\lambda' \in \Lambda \cap V, \forall (x,y) \in \mathcal{S}(\lambda) \cap [U \times W], \forall (x',y') \in \mathcal{S}(\lambda'); \\ 3.\ \mathfrak{f}_1 \in \mathfrak{L}_{L\cap W}^{\alpha,1}(K; \Lambda \cap V) \text{ and } \mathfrak{f}_2 \in \mathfrak{L}_{K\cap U}^{\alpha,1}(L; \Lambda \cap V). \end{cases}$$

Clearly, condition $\mathcal{A}3$ could be replaced by

$$\mathfrak{f}_3 \in \mathfrak{L}_{L\cap W}^{1,\alpha}(\Lambda \cap V; K) \cap \mathfrak{L}_{K\cap U}^{1,\alpha}(\Lambda \cap V; L).$$

$\mathcal{A}1$ implies that $\mathfrak{m}(\lambda)$ is finite for $\lambda \in \Lambda \cap V$.

We would show the consistency of our main hypothesis by giving some examples of parameterized families of functions $\{f(\cdot,\cdot,\lambda), \lambda \in \Lambda\}$ for which $\mathcal{A}$ is satisfied.

One gets a parameterized family of functions $\{f(\cdot,\cdot,\lambda), \lambda{\in}\Lambda\}$ satisfying $\mathcal{A}2$, for example, by assuming that $K \subset X$ and $L \subset Y$ are nonempty closed convex sets; the function $f(\cdot,y,\lambda)$ is lower semicontinuous and uniformly on $(y,\lambda) \in L \times \Lambda$ strongly convex on K, i.e., such that for some constant $c > 0$ the inequality

$$f(tx + (1-t)x', y, \lambda) \leq t f(x,y,\lambda) + (1-t) f(x',y,\lambda) - ct(1-t)\|x - x'\|^2$$

holds for every $t \in [0,1]$, every $x, x' \in K$, and every $(y,\lambda) \in L \times \Lambda$; the function $f(x,\cdot,\lambda)$ is upper semicontinuous and uniformly on $(x,\lambda) \in K \times \Lambda$ strongly concave on L, i.e., such that the inequality

$$f(x, ty + (1-t)y', \lambda) \geq t f(x,y,\lambda) + (1-t) f(x,y',\lambda) + ct(1-t)\|y - y'\|^2$$

holds for every $t \in [0,1]$, every $y, y' \in L$, and every $(x,\lambda) \in K \times \Lambda$.

Then it is routine to see that $\mathcal{A}2$ holds at any $(\bar{\lambda},\bar{x},\bar{y}) \in \operatorname{gph}\mathcal{S}$ with $\alpha = 1$, $V = Z$, $U = X$, and $W = Y$.

Examples of parameterized families of functions satisfying $\mathcal{A}3$ are given by the following.

Lemma 3.1.9. Let $(\bar\lambda, \bar x, \bar y) \in \operatorname{gph} \mathcal{S}$ and let $U \subset X$, $W \subset Y$ and $V \subset Z$ be convex neighborhoods of K, L, and $\bar\lambda$, respectively. Consider the conditions:

$$(\mathcal{F}1) \quad \begin{cases} \text{for any } (y,\lambda) \in L \times [\Lambda \cap V],\ f(\cdot, y, \lambda) \text{ is Lipschitz and regular} \\ \text{function on } U \text{ and } \partial_x f(x,y,\cdot) : \Lambda \cap V \to X^* \text{ is a } k\text{–Lipschitz map on} \\ \Lambda \cap V \text{ with } k \text{ that does not depend on } (x,y) \in K \times L, \end{cases}$$

$$(\mathcal{F}2) \quad \begin{cases} \text{for any } (x,\lambda) \in K \times [\Lambda \cap V],\ f(x, \cdot, \lambda) \text{ is Lipschitz and regular} \\ \text{function on } W \text{ and } \partial_y f(x,y,\cdot) : \Lambda \cap V \to Y^* \text{ is a } k\text{–Lipschitz map on} \\ \Lambda \cap V \text{ with } k \text{ that does not depend on } (x,y) \in K \times L, \end{cases}$$

$$(\mathcal{F}3) \quad \begin{cases} \text{for any } (x,y) \in K \times L,\ f(x, y, \cdot) \text{ is Lipschitz and regular} \\ \text{function on } V \text{ and } \partial_\lambda f(\cdot,y,\lambda) : K \to \Lambda^* \text{ is a } k\text{–Lipschitz map on} \\ K \text{ with } k \text{ that does not depend on } (y,\lambda) \in L \times [\Lambda \cap V], \end{cases}$$

$$(\mathcal{F}4) \quad \begin{cases} \text{for any } (x,y) \in K \times L,\ f(x, y, \cdot) \text{ is Lipschitz and regular} \\ \text{function on } V \text{ and } \partial_\lambda f(x,\cdot,\lambda) : L \to \Lambda^* \text{ is a } k\text{–Lipschitz map on} \\ L \text{ with } k \text{ that does not depend on } (x,\lambda) \in K \times [\Lambda \cap V]. \end{cases}$$

If f satisfies $\mathcal{F}1 - \mathcal{F}2$ or $\mathcal{F}3 - \mathcal{F}4$, then $\mathcal{A}3$ holds with $\alpha = 1$.

Proof. We follow the same steps as in the proof of Lemma 3.1.3.
If f satisfies $\mathcal{F}1$, then $\mathfrak{f}_1 \in \mathfrak{L}^{1,1}_L(K; \Lambda \cap V)$.
If f satisfies $\mathcal{F}2$, then $\mathfrak{f}_2 \in \mathfrak{L}^{1,1}_K(L; \Lambda \cap V)$.
If f satisfies $\mathcal{F}3$, then $\mathfrak{f}_3 \in \mathfrak{L}^{1,1}_L(\Lambda \cap V; K)$.
If f satisfies $\mathcal{F}4$, then $\mathfrak{f}_3 \in \mathfrak{L}^{1,1}_K(\Lambda \cap V; L)$. □

Obviously, if $f \in C^{1,1}(U \times W \times V)$, then $\mathcal{F}1$ to $\mathcal{F}4$ hold.

Lipschitz-like continuity of the saddle point map

Here we will prove that assumption $\mathcal{A}$ is sufficient for Aubin continuity of the saddle point map $\mathcal{S}$. Let us note that the result cannot be derived (or at least not in an obvious manner) from the case of minimization only. Indeed, if $f(x,y,\lambda)$ satisfies assumption $\mathcal{A}$, then the function $f(x,\lambda) := \sup_{y \in L} f(x,y,\lambda)$ satisfies assumption $A2$ but $A3$ for this $f(x,\lambda)$ cannot be derived from $\mathcal{A}3$ since the differences of suprema involved do not yield themselves to rearrangement.

Theorem 3.1.10. Assume that for the parameterized family of minimax problems $M(\lambda)$ the assumption $\mathcal{A}$ holds at some $(\bar\lambda, \bar x, \bar y) \in \operatorname{gph} \mathcal{S}$. Then for all $\lambda, \mu \in \Lambda \cap V$

$$e(\mathcal{S}(\lambda) \cap [U \times W], \mathcal{S}(\mu)) \le \frac{2k}{c} \|\lambda - \mu\|, \tag{3.18}$$

where $k := \max\{k_{\mathfrak{f}_1}, k_{\mathfrak{f}_2}\}$, hence the saddle point map $\mathcal{S} : \Lambda \rightrightarrows X \times Y$ is Aubin continuous near $(\bar\lambda, \bar x, \bar y) \in \operatorname{gph} \mathcal{S}$.

Proof. By $\mathcal{A}1$ for all $\lambda \in \Lambda \cap V$ the set $\mathcal{S}(\lambda) \cap [U \times W]$ is nonempty.

Fix $\lambda \in \Lambda \cap V$ and take some $(x_\lambda, y_\lambda) \in \mathcal{S}(\lambda) \cap [U \times W]$.

Pick any other $\mu \in \Lambda \cap V$.

Since $\mathcal{S}(\mu)$ is a nonempty set we find some $x_\mu^\varepsilon \in \pi_X \mathcal{S}(\mu)$ such that

$$\|x_\lambda - x_\mu^\varepsilon\| \leq d(x_\lambda, \pi_X \mathcal{S}(\mu)) + \varepsilon.$$

Similarly, there is $y_\mu^\varepsilon \in \pi_Y \mathcal{S}(\mu)$ such that

$$\|y_\lambda - y_\mu^\varepsilon\| \leq d(y_\lambda, \pi_Y \mathcal{S}(\mu)) + \varepsilon.$$

By the product form of the saddle point set, $(x_\mu^\varepsilon, y_\mu^\varepsilon) \in \mathcal{S}(\mu)$. The first inequality of $\mathcal{A}2$ for $(x_\lambda, y_\lambda) \in \mathcal{S}(\lambda) \cap [U \times W]$ and $(x_\mu^\varepsilon, y_\mu^\varepsilon) \in \mathcal{S}(\mu)$ reads

$$f\left(x_\lambda, y_\mu^\varepsilon, \mu\right) \geq \mathfrak{m}(\mu) + c d^{1+\alpha}(x_\lambda, \pi_X \mathcal{S}(\mu)), \tag{3.19}$$

in particular,

$$f\left(x_\lambda, y_\mu^\varepsilon, \mu\right) \geq \mathfrak{m}(\mu), \tag{3.20}$$

while the second inequality of $\mathcal{A}2$ states

$$\mathfrak{m}(\mu) \geq f\left(x_\mu^\varepsilon, y_\lambda, \mu\right) + c d^{1+\alpha}(y_\lambda, \pi_Y \mathcal{S}(\mu)), \tag{3.21}$$

in particular,

$$\mathfrak{m}(\mu) \geq f\left(x_\mu^\varepsilon, y_\lambda, \mu\right). \tag{3.22}$$

Combining (3.19) with (3.22) and (3.20) with (3.21), we get

$$\begin{aligned} f\left(x_\lambda, y_\mu^\varepsilon, \mu\right) &\geq f\left(x_\mu^\varepsilon, y_\lambda, \mu\right) + c d^{1+\alpha}(x_\lambda, \pi_X \mathcal{S}(\mu)), \\ f\left(x_\lambda, y_\mu^\varepsilon, \mu\right) &\geq f\left(x_\mu^\varepsilon, y_\lambda, \mu\right) + c d^{1+\alpha}(y_\lambda, \pi_Y \mathcal{S}(\mu)), \end{aligned}$$

which yields

$$f\left(x_\lambda, y_\mu^\varepsilon, \mu\right) - f\left(x_\mu^\varepsilon, y_\lambda, \mu\right) \geq c[\max\{d(x_\lambda, \pi_X \mathcal{S}(\mu)), d(y_\lambda, \pi_Y \mathcal{S}(\mu))\}]^{1+\alpha}.$$

By the definition of the supremum norm and since $\mathcal{S}(\mu)$ is a product set, it is obvious that

$$\begin{aligned} d((x_\lambda, y_\lambda), \mathcal{S}(\mu)) &= d((x_\lambda, y_\lambda), \pi_X \mathcal{S}(\mu) \times \pi_Y \mathcal{S}(\mu)) \\ &= \max\{d(x_\lambda, \pi_X \mathcal{S}(\mu)), d(y_\lambda, \pi_Y \mathcal{S}(\mu))\}, \end{aligned} \tag{3.23}$$

and the above inequality can be rewritten as

$$f\left(x_\lambda, y_\mu^\varepsilon, \mu\right) - f\left(x_\mu^\varepsilon, y_\lambda, \mu\right) \geq c d^{1+\alpha}((x_\lambda, y_\lambda), \mathcal{S}(\mu)).$$

We transform the left-hand side to get

$$f\left(x_\lambda, y_\mu^\varepsilon, \mu\right) - f(x_\lambda, y_\lambda, \mu) + f(x_\lambda, y_\lambda, \mu) - f\left(x_\mu^\varepsilon, y_\lambda, \mu\right) \geq c d^{1+\alpha}((x_\lambda, y_\lambda), \mathcal{S}(\mu)),$$

which is

$$-\mathfrak{f}_2\left(x_\lambda, y_\lambda, y_\mu^\varepsilon, \mu\right) - \mathfrak{f}_1\left(x_\mu^\varepsilon, x_\lambda, y_\lambda, \mu\right) \geq c d^{1+\alpha}((x_\lambda, y_\lambda), \mathcal{S}(\mu)). \tag{3.24}$$

On the other hand, since $(x_\lambda, y_\lambda) \in \mathcal{S}(\lambda)$, the saddle point inequalities give

$$f(x, y_\lambda, \lambda) \geq f(x_\lambda, y_\lambda, \lambda) \geq f(x_\lambda, y, \lambda), \quad \forall x \in K, \forall y \in L.$$

In particular, for $x = x_\mu^\varepsilon \in K$ we have

$$f(x_\mu^\varepsilon, y_\lambda, \lambda) \geq f(x_\lambda, y_\lambda, \lambda),$$

which is

$$\mathfrak{f}_1(x_\mu^\varepsilon, x_\lambda, y_\lambda, \lambda) \geq 0, \tag{3.25}$$

and for $y = y_\mu^\varepsilon \in L$ we get

$$f(x_\lambda, y_\lambda, \lambda) \geq f(x_\lambda, y_\mu^\varepsilon, \lambda),$$

which is

$$\mathfrak{f}_2(x_\lambda, y_\lambda, y_\mu^\varepsilon, \lambda) \geq 0. \tag{3.26}$$

Adding the inequalities (3.24), (3.25), and (3.26) and rearranging we obtain

$$\begin{aligned}\left[\mathfrak{f}_1\left(x_\mu^\varepsilon, x_\lambda, y_\lambda, \lambda\right) - \mathfrak{f}_1\left(x_\mu^\varepsilon, x_\lambda, y_\lambda, \mu\right)\right] + \left[\mathfrak{f}_2\left(x_\lambda, y_\lambda, y_\mu^\varepsilon, \lambda\right) - \mathfrak{f}_2\left(x_\lambda, y_\lambda, y_\mu^\varepsilon, \mu\right)\right] \\ \geq c d^{1+\alpha}((x_\lambda, y_\lambda), \mathcal{S}(\mu)).\end{aligned} \tag{3.27}$$

Since by $\mathcal{A}3$, $\mathfrak{f}_1 \in \mathfrak{L}_{L\cap W}^{\alpha,1}(K; \Lambda \cap V)$, the term in first brackets in (3.27) is estimated by

$$\mathfrak{f}_1\left(x_\mu^\varepsilon, x_\lambda, y_\lambda, \mu\right) - \mathfrak{f}_1\left(x_\mu^\varepsilon, x_\lambda, y_\lambda, \lambda\right) \leq k_{\mathfrak{f}_1} \|x_\mu^\varepsilon - x_\lambda\|^\alpha \|\lambda - \mu\|, \tag{3.28}$$

and since $\mathfrak{f}_2 \in \mathfrak{L}_{K\cap U}^{\alpha,1}(L; \Lambda \cap V)$ the term in second brackets in (3.27) is estimated by

$$\mathfrak{f}_2\left(x_\lambda, y_\lambda, y_\mu^\varepsilon, \lambda\right) - \mathfrak{f}_2\left(x_\lambda, y_\lambda, y_\mu^\varepsilon, \mu\right) \leq k_{\mathfrak{f}_2} \|y_\lambda - y_\mu^\varepsilon\|^\alpha \|\lambda - \mu\|. \tag{3.29}$$

Using (3.29) and (3.28) in (3.27) and setting $k := \max\{k_{\bar{f}_1}, k_{\bar{f}_2}\}$, we get

$$k\|\lambda - \mu\|[\|x_\lambda - x^\varepsilon_\mu\|^\alpha + \|y_\lambda - y^\varepsilon_\mu\|^\alpha] \geq cd^{1+\alpha}((x_\lambda, y_\lambda), \mathcal{S}(\mu)).$$

By the choice of x^ε_μ and y^ε_μ, we have that

$$k\|\lambda - \mu\|[(d(x_\lambda, \pi_X\mathcal{S}(\mu)) + \varepsilon)^\alpha + (d(y_\lambda, \pi_Y\mathcal{S}(\mu)) + \varepsilon)^\alpha]$$
$$\geq cd^{1+\alpha}((x_\lambda, y_\lambda), \mathcal{S}(\mu)).$$

Passing to limit $\varepsilon \downarrow 0$ we obtain

$$k\|\lambda - \mu\|[d^\alpha(x_\lambda, \pi_X\mathcal{S}(\mu)) + d^\alpha(y_\lambda, \pi_Y\mathcal{S}(\mu))] \geq cd^{1+\alpha}((x_\lambda, y_\lambda), \mathcal{S}(\mu)). \tag{3.30}$$

By (3.23) we get

$$d^\alpha(y_\lambda, \pi_Y\mathcal{S}(\mu)) + d^\alpha(x_\lambda, \pi_X\mathcal{S}(\mu)) \leq 2\,[\max\{d(y_\lambda, \pi_Y\mathcal{S}(\mu)), d(x_\lambda, \pi_X\mathcal{S}(\mu))\}]^\alpha$$
$$= 2d^\alpha((x_\lambda, y_\lambda), \mathcal{S}(\mu)),$$

and from (3.30) we obtain

$$2k\|\lambda - \mu\|d^\alpha((x_\lambda, y_\lambda), \mathcal{S}(\mu)) \geq cd^{1+\alpha}((x_\lambda, y_\lambda), \mathcal{S}(\mu)).$$

This yields

$$\frac{2k}{c}\|\lambda - \mu\| \geq d((x_\lambda, y_\lambda), \mathcal{S}(\mu)),$$

and since (x_λ, y_λ) was an arbitrary element of $\mathcal{S}(\lambda) \cap [U \times W]$ the latter implies

$$e(\mathcal{S}(\lambda) \cap [U \times W], \mathcal{S}(\mu)) \leq \frac{2k}{c}\|\lambda - \mu\|.$$

The proof is completed. □

As an immediate consequence of Theorem 3.1.10 and Lemma 3.1.9 one deduces the following.

Corollary 3.1.11. Let for the parameterized family of minimax problems $M(\lambda)$ the following assumption hold:

$$(\mathcal{A}') \qquad \begin{cases} 1. & \mathcal{S}(\lambda) \neq \emptyset \text{ for any } \lambda \in \Lambda; \\ 2. & \text{for some constant } c > 0 \text{ and all } \lambda \in \Lambda, (x, y) \in \mathcal{S}(\lambda), (x', y') \in K \times L: \\ & f(x', y, \lambda) \geq \mathfrak{m}(\lambda) + cd^2(x', \pi_X\mathcal{S}(\lambda)), \\ & f(x, y', \lambda) \leq \mathfrak{m}(\lambda) - cd^2(y', \pi_Y\mathcal{S}(\lambda)); \\ 3. & f \in C^{1,1}(X \times Y \times Z). \end{cases}$$

Then the saddle point map $\mathcal{S} : \Lambda \to X \times Y$ is single-valued and Lipschitz continuous.

As we pointed out after Corollary 3.1.7 we could deduce the single-valuedness and Lipschitz continuity of the saddle point map $\mathcal{S}$ when X and Y has separable duals, the sets K and L are convex, $f \in C^{1,1}(X \times Y \times Z)$, and there exists a constant $c > 0$ such that for all $\lambda \in \Lambda$, $x, z \in K$, $y, w \in L$,

$$\langle M(z-x), z-x\rangle \geq c\|z-x\|^2, \quad \langle N(w-y), w-y\rangle \leq -c\|w-y\|^2$$

for all $M \in \partial^2 f(\cdot, y, \lambda)(x)$ and all $N \in \partial^2 f(x, \cdot, \lambda)(y)$.

3.1.3 Lipschitz continuity of the saddle points map in context of two-player zero sum differential games

In this subsection we briefly consider a differential game for which our result might be of relevance.

In differential games, open-loop strategies are of low interest in many examples. One major reason is that differential games with open-loop strategies do not satisfy, in general, the dynamic programming principle (see Cardaliaguet, Quincampoix and Saint-Pierre [39, 41], Plaskacz and Quincampoix [134]). It is well known now that to solve many problems in differential games (existence of a value, characterization of the game through Hamilton–Jacobi equations), one needs a more general class of strategies which contains the feedback strategies.[1] Such class of strategies is the class of nonanticipative strategies introduced by Elliot–Roxin–Varaiya–Kalton (cf. for instance Cardaliaguet, Quincampoix and Saint-Pierre [40]); another possible class of strategies are the positional strategies discussed in Krasovskii and Subbotin [104]. The class of nonanticipative strategies is nice enough to prove the existence of the value, but it is hard to implement for the players. So it is important to know when the nonanticipative strategies giving the value of the game can be reduced to feedback strategies. We will explain in this part how the main result of the paper can lead to a partial answer to this question.

We consider the following differential game with dynamic described by the differential equation:

$$\begin{cases} x'(t) = f(x(t), u(t)), \quad y'(t) \in g(y(t), v(t)), \\ u(t) \in U, \ v(t) \in V, \end{cases} \tag{3.31}$$

where $f : \mathbb{R}^n \times U \to \mathbb{R}^n$ and $g : \mathbb{R}^n \times V \to \mathbb{R}^n$ are (globally) Lipschitz, $U \subset \mathbb{R}^m$, $V \subset \mathbb{R}^p$ being the control sets of the players. The first player—Ursula, playing with u—wants

[1] It has been shown in Cardaliaguet [38] that the class of regular feedback is not rich enough to solve differential games at a satisfactory level of generality.

to minimize a given cost. The goal of the second player—Victor, playing with v—is to maximize the cost

$$J(x_0, y_0, u(\cdot), v(\cdot)) := \int_0^{\infty} e^{-rt} l(x(t), y(t))\, dt,$$

where $(x(\cdot), y(\cdot))$ is the unique solution starting at $t = 0$ from (x_0, y_0) and $r > 0$ is fixed. Observe that the game is in a separable form, i.e., each player acts in his own dynamics. This is the case, for instance, for pursuit games. Moreover, the integral cost does not depend directly on the control but only on the trajectories.

We work here in the framework of the *nonanticipative strategies* (also called Varaiya–Roxin–Elliot–Kalton strategies). Let

$$\mathcal{U} = L^1([0, +\infty[, U), \ \mathcal{V} = L^1([0, +\infty[, V) \tag{3.32}$$

be the sets of time-measurable controls of the first (Ursula) and the second (Victor) player, respectively. We denote $t \mapsto (x(t, x_0, u(t)), y(t, y_0, v(t)))$ the solution of (3.31) starting at $t = 0$ from (x_0, y_0).

Definition 3.1.12. A map $\alpha : \mathcal{V} \to \mathcal{U}$ is *a nonanticipative strategy* (*for Ursula*) if it satisfies the following condition: For any $s \geq 0$, for any $v_1(\cdot)$ and $v_2(\cdot)$ belonging to $\mathcal{V}$ such that $v_1(\cdot)$ and $v_2(\cdot)$ coincide almost everywhere on $[0, s]$, the images $\alpha(v_1(\cdot))$ and $\alpha(v_2(\cdot))$ coincide almost everywhere on $[0, s]$.

Nonanticipative strategies $\beta : \mathcal{U} \to \mathcal{V}$ (for Victor) are defined in the symmetric way.

Assume now that f and g are continuous and Lipschitz with respect to x and y. Then, we know that the game has a value (cf. Cardaliaguet, Quincampoix and Saint-Pierre [41]), namely,

$$V(x_0, y_0) = \inf_{\alpha} \sup_{v \in \mathcal{V}} J(x_0, y_0, \alpha(v(\cdot)), v(\cdot)) = \sup_{\beta} \inf_{u \in \mathcal{U}} J(x_0, y_0, u(\cdot), \beta(u(\cdot))).$$

Let us denote by $R(t)$ the attainable set of the dynamics (3.31) at moment t; i.e.,

$$\begin{aligned} R(t) = \{(x(t), y(t)) \in \mathbb{R}^n \times \mathbb{R}^n : \exists\ u \in \mathcal{U}, v \in \mathcal{V} \text{ such that} \\ (x(\cdot), y(\cdot)) \text{ is the solution of (3.31) starting at } t = 0 \text{ from } (x_0, y_0)\}. \end{aligned}$$

Now, suppose that U and V are convex and compact. Saddle point of the function $l(\cdot, \cdot)$ on $R(t)$ will be any point $(\bar{x}, \bar{y}) \in R(t)$ that satisfies

$$l(\bar{x}, y) \leq l(\bar{x}, \bar{y}) \leq l(x, \bar{y}), \quad \forall (x, y) \in R(t),$$

and, because of $e^{-rt} > 0$, the saddle points of $l(\cdot, \cdot)$ on $R(t)$ will be the same as the saddle points of $e^{-rt} l(\cdot, \cdot)$ on $R(t)$.

Let us denote the (possibly empty) set of all saddle points of the function $l(\cdot,\cdot)$ on $R(t)$ by $\mathcal{S}(t) := \{(\bar{x},\bar{y}) \in R(t) : l(\bar{x},y) \leq l(\bar{x},\bar{y}) \leq l(x,\bar{y}), \quad \forall (x,y) \in R(t)\}$.

Let us suppose that the parameterized by t family of functions $\{e^{-rt}l(\cdot,\cdot), t \in [0,\infty)\}$ satisfies an assumption slightly stronger than assumption $\mathcal{A}$, namely:

$$\begin{cases} 1. \;\; \mathcal{S}(t) \neq \emptyset, \;\; \forall t \geq 0; \\ \text{and there exist constants } k, c > 0 \text{ and } \alpha \in [0,1] \text{ such that} \\ \forall t, t' \geq 0, \forall (x,y) \in \mathcal{S}(t), \forall (x',y') \in \mathcal{S}(t') \text{ it holds :} \\ 2. \;\; l(x',y) \geq l(x,y) + ce^{rt}\|x'-x\|^{1+\alpha}, \\ \quad\;\; l(x,y') \leq l(x,y) - ce^{rt}\|y'-y\|^{1+\alpha}; \\ 3. \;\; |l(x,y) - l(x',y)| \leq k\|x-x'\|^{\alpha}, \\ \quad\;\; |l(x,y) - l(x,y')| \leq k\|y-y'\|^{\alpha}. \end{cases}$$

This assumption guarantees that for any $t \in [0,\infty)$ the saddle point mapping $\mathcal{S}(t)$ is single-valued and Lipschitz continuous; i.e., for all positive t, the function $e^{-rt}l(\cdot,\cdot)$ has a saddle point $(x(t), y(t))$ on the attainable set $R(t)$ of the dynamics (3.31), which depends in a Lipschitz way on t.

Therefore, if it turns out that so-obtained single valued saddle point mapping is a trajectory $(x(\cdot), y(\cdot))$ of (3.31), then it is an optimal feedback strategy of the game.

For example, under the above assumptions in the case when $m = p = n$ and $f(x,u) = u$, $g(y,v) = v$, the Lipschitz continuity on t of the saddle point map implies that the corresponding controls u and v belong to $\mathcal{U}$ and $\mathcal{V}$, respectively, and, hence, they generate a trajectory of the differential game.

3.2 Aubin criterion for metric regularity

In this section we investigate metric regularity of set-valued mappings. We present a derivative criterion for metric regularity of set-valued mappings that is based on works of J.-P. Aubin and co-authors. A related implicit mapping theorem is also obtained and several applications are given.

Throughout this section, X and Y are Banach spaces. The norms of both X and Y are denoted by $\|\cdot\|$; the closed ball centered at x with radius r by $B[x,r]$ and the open ball by $B(x,r)$; the closed unit ball is simply B and the open one B°. A neighborhood of a point x is any open set containing x. The distance from a point x to a set A is denoted by $d(x,A)$. By a mapping F from X to Y we generally mean a set-valued mapping and write $F : X \rightrightarrows Y$, having its inverse F^{-1} defined as $F^{-1}(y) = \{x \mid y \in F(x)\}$ and graph $\operatorname{gph} F = \{(x,y) \mid y \in F(x)\}$. When F is single-valued (a function) we write $F : X \to Y$.

A mapping $F : X \rightrightarrows Y$ is said to be *metrically regular* at $\bar{x}$ for $\bar{y}$ if $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ and there exist a constant $\kappa > 0$ and neighborhoods U of $\bar{x}$ and V of $\bar{y}$ such that

$$d(x, F^{-1}(y)) \leq \kappa d(y, F(x)) \text{ for all } (x, y) \in U \times V. \tag{3.33}$$

The metric regularity can be identified with the finiteness of the *regularity modulus* defined as

$$\operatorname{reg} F(\bar{x}|\bar{y}) = \inf\{\, \kappa \mid \text{ there exist neighborhoods } U \text{ and } V \text{ such that (3.33) holds}\}.$$

The absence of metric regularity is indicated by $\operatorname{reg} F(\bar{x}|\bar{y}) = \infty$.

The concept of metric regularity goes back to classical results by Banach, Lyusternik and Graves. More recently, its central role has been recognized in variational analysis for both theoretical developments such as obtaining necessary optimality conditions and also in numerically oriented studies, e.g., when deriving error bounds for solution approximations. Discussions of the property of metric regularity, its relations to other properties and characterizations by various approximations are presented in Rockafellar and Wets [152] and Ioffe [90].

Given a mapping $F : X \rightrightarrows Y$, the *graphical* (*contingent*) *derivative* of F at $(x, y) \in \operatorname{gph} F$ is the mapping $DF(x|y) : X \rightrightarrows Y$ whose graph is the tangent cone $T_{\operatorname{gph} F}(x, y)$ to $\operatorname{gph} F$ at (x, y):

$$v \in DF(x|y)(u) \iff (u, v) \in T_{\operatorname{gph} F}(x, y).$$

Recall that the tangent cone is defined as follows: $(u, v) \in T_{\operatorname{gph} F}(x, y)$ when there exist sequences $t_n \downarrow 0$, $u_n \to u$ and $v_n \to v$ such that $y + t_n v_n \in F(x + t_n u_n)$ for all n.

The mapping $DF(x|y)$ is positively homogeneous since its graph is a cone; specifically, one has $DF(x|y)(0) \ni 0$ and $DF(x|y)(\lambda u) = \lambda DF(x|y)(u)$ for all $u \in X$ for $\lambda > 0$. The *convexified graphical derivative* $D^{\star\star}F(x|y)$ of F at x for y is defined in a similar way:

$$v \in D^{\star\star}F(x|y)(u) \iff (u, v) \in \operatorname{co}\, T_{\operatorname{gph} F}(x, y)$$

where co stands for the closed convex hull. We also use the *inner* and the *outer "norms"* (see Rockafellar and Wets [152, Section 9D]) of a mapping $H : X \rightrightarrows Y$:

$$\|H\|^{-} = \sup_{x \in B} \inf_{y \in H(x)} \|y\| \quad \text{and} \quad \|H\|^{+} = \sup_{x \in B} \sup_{y \in H(x)} \|y\|.$$

Outer and inner norms can be related through adjoints. For a positively homogeneous mapping $F : X \rightrightarrows Y$ the *upper adjoint* $F^{*+} : Y^* \rightrightarrows X^*$ is defined by

$$(y^*, x^*) \in \operatorname{gph} F^{*+} \iff \langle x^*, x \rangle \leq \langle y^*, y \rangle \text{ for all } (x, y) \in \operatorname{gph} F,$$

while the *lower adjoint* $F^{*-} : Y^* \rightrightarrows X^*$ is

$$(y^*, x^*) \in \operatorname{gph} F^{*-} \iff \langle x^*, x \rangle \geq \langle y^*, y \rangle \text{ for all } (x, y) \in \operatorname{gph} F,$$

where X^* and Y^* are the dual spaces of X and Y. Borwein derived in [23] the following duality relations between outer and inner norms for a sublinear mapping F with closed graph:

$$\|F\|^+ = \|F^{*+}\|^- = \|F^{*-}\|^- \quad \text{and} \quad \|F\|^- = \|F^{*+}\|^+ = \|F^{*-}\|^+.$$

Recall that a mapping $F : X \rightrightarrows Y$ is said to have a locally closed graph at $(\bar{x}, \bar{y})$ when $\operatorname{gph} F \cap (B[\bar{x}, r] \times B[\bar{y}, r])$ is a closed set for some $r > 0$.

The central result in this section is the following theorem:

Theorem 3.2.1 (Aubin criterion). Consider two Banach spaces X and Y, and a mapping $F : X \rightrightarrows Y$ which graph is locally closed at $(\bar{x}, \bar{y}) \in \operatorname{gph} F$. Then

$$\operatorname{reg} F(\bar{x}|\bar{y}) \le \limsup_{\substack{(x,y)\to(\bar{x},\bar{y})\\ (x,y)\in\operatorname{gph} F}} \|DF(x|y)^{-1}\|^-, \tag{3.34}$$

and hence F is metrically regular at $\bar{x}$ for $\bar{y}$ provided that

$$\limsup_{\substack{(x,y)\to(\bar{x},\bar{y})\\ (x,y)\in\operatorname{gph} F}} \|DF(x|y)^{-1}\|^- < \infty. \tag{3.35}$$

If X is finite dimensional, then (3.34) becomes an equality,

$$\operatorname{reg} F(\bar{x}|\bar{y}) = \limsup_{\substack{(x,y)\to(\bar{x},\bar{y})\\ (x,y)\in\operatorname{gph} F}} \|DF(x|y)^{-1}\|^-, \tag{3.36}$$

and hence F is metrically regular at $\bar{x}$ for $\bar{y}$ if and only if (3.35) holds. Moreover, when both spaces X and Y are finite dimensional one has

$$\operatorname{reg} F(\bar{x}|\bar{y}) = \limsup_{\substack{(x,y)\to(\bar{x},\bar{y})\\ (x,y)\in\operatorname{gph} F}} \|D^{\star\star} F(x|y)^{-1}\|^-. \tag{3.37}$$

Theorem 3.2.1 can be viewed as a partial extension of Theorem 5.4.3 in Aubin and Frankowska book [10] where a sufficient condition for the Aubin property of the inverse F^{-1} is shown, for predecessors of this result see Aubin [3] and Aubin and Frankowska [9]. That condition is in general weaker than (3.35) but, as we see here, for a finite-dimensional X is actually equivalent to it. Recall that a mapping $S : Y \rightrightarrows X$ has *the Aubin property* at $\bar{y}$ for $\bar{x}$ when $(\bar{y}, \bar{x}) \in \operatorname{gph} S$ and there exist neighborhoods V of $\bar{y}$ and U of $\bar{x}$ such that

$$e(S(y) \cap U, S(y')) \le \kappa \|y - y'\| \quad \text{for all} \quad y, y' \in V, \tag{3.38}$$

where $e(A, B) = \sup_{x\in A} d(x, B)$ is the *excess* from A to B. The Aubin property of a mapping S is known to be equivalent to the metric regularity of S^{-1} and was introduced in Aubin [5] under the name "pseudo-Lipschitz" continuity, it was studied in Aubin and Frankowska [9] in the infinite dimensional case, for more bibliographical details see Rockafellar and Wets [152]. Moreover, the infimum of the constant κ in (3.38) is equal to $\operatorname{reg} S^{-1}(\bar{x}|\bar{y})$.

The equality (3.37) was stated recently in Aubin, Bayen, Bonneuil, and Saint-Pierre [7] with a proof based on viability theory. The given here proof of (3.37) is inspired by the proof of Theorem 3.2.4 in Aubin [6] due to Frankowska.

The characterization of metric regularity exhibited in Theorem 3.2.1 complements, and in some sense also completes, results previously displayed by J.-P. Aubin and co-authors; therefore, we call it the *Aubin criterion for metric regularity*.

In a sense "dual" to the Aubin criterion is the known Mordukhovich criterion in finite dimensions, see Mordukhovich [122] and Rockafellar and Wets [152] which uses the *coderivative* $D^*F(x|y)$ defined as

$$v \in D^*F(x|y)(u) \iff (v, -u) \in N_{\operatorname{gph} F}(x, y),$$

where $N_C(x)$ is the (nonconvex, limiting) normal cone to the set C at x. The Mordukhovich criterion says that F is metrically regular at $\bar{x}$ for $\bar{y}$ if and only if

$$\|D^*F^{-1}(\bar{y}|\bar{x})\|^+ < \infty.$$

One would expect that each of these criteria could be directly derived from the other one, and this is clearly so when $\operatorname{gph} F$ is Clarke regular, see Rockafellar and Wets [152, 8.40 and 11.29]. In infinite dimensions, the characterizations of the metric regularity via coderivatives assume some more from the spaces, e.g., to be Asplund, see Mordukhovich [123], Mordukhovich and Shao [125]. When the domain space X is finite dimensional and Y is any Banach space, a necessary and sufficient condition for metric regularity in terms of the Ioffe approximate coderivative is given in Jourani and Thibault [99]. This latter result is the closest to Theorem 3.2.1 from the literature known to the author. We also refer to the book Klatte and Kummer [102] as a source of information on criteria for metric regularity.

If $F : X \to Y$ is a bounded linear mapping, denoted $F \in L(X, Y)$, then the Aubin criterion (3.36) is also valid for X and Y Banach spaces and reduces to $\operatorname{reg} F(\bar{x}|\bar{y}) = \|F^{-1}\|^-$. Equivalently, F is metrically regular (at any point) if and only if F is surjective; this covers the classical case of the Banach open mapping principle. The equivalence among metric regularity at the origin, the finiteness of the inner norm, and the surjectivity holds also for mappings acting in Banach spaces whose graphs are closed and convex cones. Specifically, we have the following result proved in Dontchev, Lewis and Rockafellar [62, Example 2.1] :

Proposition 3.2.2. Let X and Y be Banach spaces and let $F : X \rightrightarrows Y$ be such that $\operatorname{gph} F$ is a closed and convex cone. Then the modulus of regularity of F at the origin satisfies

$$\operatorname{reg} F(0|0) = \|DF(0|0)^{-1}\|^- = \|F^{-1}\|^-.$$

Moreover, $\operatorname{reg} F(0|0) < \infty$ if and only if F is surjective and then F is metrically regular at any point in its graph.

In Subsection 3.2.2 we give a proof of Theorem 3.2.1 by first obtaining the sufficiency part of the Aubin criterion as a corollary from a more general "implicit mapping theorem" (Theorem 3.2.3) in Subsection 3.2.1, which is about the solution mapping of a generalized equation of the form

$$0 \in G(p, x), \tag{3.39}$$

where p is a parameter. We show that if the partial graphical derivative with respect to x of the mapping G is bounded in the sense of (3.35), then G has a property of "partial metric regularity".

In a related paper [110], Ledyaev and Zhu obtained an implicit mapping theorem for a general inclusion of the form (3.39) in terms of coderivatives in Banach spaces assumed to have Fréchet-smooth Lipschitz bump functions. Putting aside the derivative condition in our Theorem 3.2.3 and the coderivative condition in Ledyaev and Zhu [110, Theorem 3.7] which are independent from each other and can not be compared, we impose weaker conditions on the mapping G and allow for arbitrary Banach spaces X and Y.

In Subsection 3.2.3 we present applications of the Aubin criterion to systems of inequalities and to variational inequalities, obtaining a new characterization of strong regularity of variational inequalities over polyhedral sets. We also give a new proof of the radius (Eckart-Young) theorem first proven in Dontchev, Lewis and Rockafellar [62] with the help of Mordukhovich criterion; for history and recent developments, see Dontchev, Lewis and Rockafellar [62], Dontchev and Rockafellar [66] and Dontchev and Lewis [61].

In addition to the Aubin criterion, in Subsection 3.2.3 we use a fundamental result in the modern nonlinear analysis, commonly known as the Lyusternik-Graves theorem, for more see, e.g., Aubin [5], Aubin and Frankowska [9], Dontchev, Lewis and Rockafellar [62] and Ioffe [90]. First, we need some terminology. For a function $g : X \to Y$ and a point $\bar{x} \in \operatorname{int} \operatorname{dom} g$, we introduce Lipschitz modulus of g at $\bar{x}$ as follows:

$$\operatorname{lip} g(\bar{x}) = \limsup_{\substack{x,x' \to \bar{x} \\ x \neq x'}} \frac{\|g(x) - g(x')\|}{\|x - x'\|}.$$

Recall that a function $g : X \to Y$ is strictly differentiable at $\bar{x} \in \operatorname{int} \operatorname{dom} g$ with a strict derivative mapping $\nabla g(\bar{x}) \in L(X, Y)$, the space of linear bounded mappings from X to Y, if

$$\operatorname{lip}(g - \nabla g(\bar{x}))(\bar{x}) = 0.$$

We use the following form of the well-known

Lyusternik-Graves theorem. Let X and Y be Banach spaces and consider a mapping $\mathcal{F} : X \rightrightarrows Y$ and a point $(\bar{x}, \bar{y}) \in \operatorname{gph}\mathcal{F}$ at which $\operatorname{gph}\mathcal{F}$ is locally closed. Then for any function $g : X \to Y$ which is strictly differentiable at $\bar{x}$ one has

$$\operatorname{reg}(g + \mathcal{F})(\bar{x}|\bar{y} + g(\bar{x})) = \operatorname{reg}(\nabla g(\bar{x}) + \mathcal{F})(\bar{x}|\bar{y} + \nabla g(\bar{x})\bar{x}).$$

3.2.1 An implicit mapping theorem

In this subsection we study the inclusion (generalized equation)

$$0 \in G(p, x),$$

where $G : P \times X \rightrightarrows Y$, X and Y are Banach spaces, P is a metric space, $x \in X$ is the variable we are solving for and $p \in P$ is a parameter. Let us denote by $S : P \rightrightarrows X$ the *solution mapping* which associates to a value p the set of solutions

$$S(p) := \{x \in X \mid G(p, x) \ni 0\}. \tag{3.40}$$

We will show that the local boundedness of the partial graphical derivative of the mapping G in x, of the kind displayed in (3.35), implies partial metric regularity of G. The partial graphical derivative $D_x G(p, x|y)$ of G is defined as the graphical derivative of the mapping $x \mapsto G(p, x)$ with p fixed.

Theorem 3.2.3 (Implicit mapping theorem). Let X and Y be Banach spaces, and let P be a metric space. Consider a mapping $G : P \times X \rightrightarrows Y$ and a point $(\bar{p}, \bar{x}, 0) \in \operatorname{gph} G$ such that the graph of G is locally closed near $(\bar{p}, \bar{x}, 0)$ and the function $p \to d(0, G(p, \bar{x}))$ is upper semicontinuous at $\bar{p}$. Then for every positive scalar c satisfying

$$\limsup_{\substack{(p,x,y)\to(\bar{p},\bar{x},0)\\(p,x,y)\in\operatorname{gph} G}} \|D_x G(p, x|y)^{-1}\|^- < c \tag{3.41}$$

there exist neighborhoods V of $\bar{p}$ and U of $\bar{x}$ such that one has

$$d(x, S(p)) \le c\,d(0, G(p, x)) \quad \text{for } x \in U \text{ and } p \in V. \tag{3.42}$$

Proof. On the product space $Z := X \times Y$ we consider the norm

$$|||(x, y)||| := \max\{\|x\|, c\|y\|\},$$

which makes $(Z, |||\cdot|||)$ a Banach space, and on the space $P \times Z$ we introduce the metric

$$\sigma((p, z), (q, w)) := \max\{\rho(p, q), |||z - w|||\} \quad \text{for } p, q \in P,\ z, w \in Z,$$

where ρ stands for the metric of P.

A constant c satisfies (3.41) if and only if there exists $\eta > 0$ such that

$$\text{(3.43)} \quad \begin{array}{l} \text{for every } (p, x, y) \in \operatorname{gph} G \text{ with } \sigma((p, x, y), (\bar{p}, \bar{x}, 0)) \le 3\eta, \\ \text{and for every } v \in Y \text{ there exists } u \in D_x G(p, x|y)^{-1}(v) \text{ with } \|u\| \le c\|v\|. \end{array}$$

We can always choose η smaller so that the set $\operatorname{gph} G \cap (\bar{p}, \bar{x}, 0) + 3\eta B$ is closed. Next, let us pick $\varepsilon > 0$ such that

$$\text{(3.44)} \qquad c\varepsilon < 1.$$

In the proof we use the following lemma:

Lemma 3.2.4. For η and ε as above, choose any $(p, \omega, v) \in \operatorname{gph} G$ with $(p, \omega, v) \in (\bar{p}, \bar{x}, 0) + \eta B$ and any s, $0 < s \le \varepsilon\eta$. Then for every $y' \in v + sB^\circ$ there exists $\hat{x}$ with $(p, \hat{x}, y') \in \operatorname{gph} G$ such that

$$\text{(3.45)} \qquad \|\hat{x} - \omega\| \le \frac{1}{\varepsilon}\|y' - v\|.$$

Proof of Lemma 3.2.4. Pick $(p, \omega, v) \in \operatorname{gph} G$ and s as required. The set $E_p := \{(x, y) | (p, x, y) \in \operatorname{gph} G \cap ((\bar{p}, \bar{x}, 0) + 3\eta B)\} \subset X \times Y$ equipped with the metric induced by the norm $||| \cdot |||$ is a complete metric space. The function $V_p : E_p \to \boldsymbol{R}$ defined as

$$V_p(x, y) := \|y' - y\| \quad \text{for}(x, y) \in E_p$$

is continuous in its domain E_p. We apply the Ekeland variational principle to V_p for (x, y) near (ω, v) and the ε chosen in (3.44) to obtain the existence of $(\hat{x}, \hat{y}) \in E_p$ such that

$$\text{(3.46)} \qquad V_p(\hat{x}, \hat{y}) + \varepsilon|||(\omega, v) - (\hat{x}, \hat{y})||| \le V_p(\omega, v)$$

and

$$\text{(3.47)} \qquad V_p(\hat{x}, \hat{y}) \le V_p(x, y) + \varepsilon|||(x, y) - (\hat{x}, \hat{y})||| \quad \text{for all } (x, y) \in E_p.$$

The relations (3.46) and (3.47) come down as

$$\text{(3.48)} \qquad \|y' - \hat{y}\| + \varepsilon|||(\omega, v) - (\hat{x}, \hat{y})||| \le \|y' - v\|$$

and

$$\text{(3.49)} \qquad \|y' - \hat{y}\| \le \|y' - y\| + \varepsilon|||(x, y) - (\hat{x}, \hat{y})||| \quad \text{for all}(x, y) \in E_p.$$

From (3.48) we obtain

$$\text{(3.50)} \qquad |||(\omega, v) - (\hat{x}, \hat{y})||| \le \frac{1}{\varepsilon}\|y' - v\|.$$

Since $y' \in v + sB^{\circ}$, we then have

$$|||(\omega, v) - (\hat{x}, \hat{y})||| < \frac{s}{\varepsilon},$$

and hence, from the choice of (p, ω, v),

$$\begin{array}{rcl}\sigma((p,\hat{x},\hat{y}),(\bar{p},\bar{x},0)) & \leq & \sigma((p,\hat{x},\hat{y}),(p,\omega,v)) + \sigma((p,\omega,v),(\bar{p},\bar{x},0)) \\ & \leq & \eta + |||(\omega,v)-(\hat{x},\hat{y})||| < \eta + \dfrac{s}{\varepsilon} \leq 2\eta.\end{array}$$

Thus, $(p, \hat{x}, \hat{y}) \in \operatorname{gph} G$ with $\sigma((\bar{p}, \bar{x}, 0), (p, \hat{x}, \hat{y})) \leq 2\eta$, and then (3.43) implies that there exists $u \in X$ such that

$$y' - \hat{y} \in D_x G(p, \hat{x}|\hat{y})(u) \quad \text{and} \quad \|u\| \leq c\|y' - \hat{y}\|. \tag{3.51}$$

By the definition of the partial graphical derivative, there exist sequences $t_n \downarrow 0$, $u_n \to u$, and $v_n \to y' - \hat{y}$ such that

$$\hat{y} + t_n v_n \in G(p, \hat{x} + t_n u_n) \text{ for all } n,$$

meaning that, for sufficiently large n, $(\hat{x} + t_n u_n, \hat{y} + t_n v_n) \in E_p$. Now, if we plug the point $(\hat{x} + t_n u_n, \hat{y} + t_n v_n)$ into (3.49), we obtain

$$\|y' - \hat{y}\| \leq \|y' - (\hat{y} + t_n v_n)\| + \varepsilon|||(\hat{x} + t_n u_n, \hat{y} + t_n v_n) - (\hat{x}, \hat{y})|||$$

resulting in

$$\|y' - \hat{y}\| \leq (1 - t_n)\|y' - \hat{y}\| + t_n\|v_n - (y' - \hat{y})\| + \varepsilon t_n|||(u_n, v_n)|||.$$

After obvious simplifications, this gives

$$\|y' - \hat{y}\| \leq \varepsilon|||(u_n, v_n)||| + \|v_n - (y' - \hat{y})\|.$$

Passing to the limit with $n \to \infty$ we obtain

$$\|y' - \hat{y}\| \leq \varepsilon|||(u, y' - \hat{y})|||$$

and hence, taking into account the second relation in (3.51) we conclude that

$$\|y' - \hat{y}\| \leq \varepsilon c\|y' - \hat{y}\|.$$

Since by (3.44) $\varepsilon c < 1$, we finally obtain that $y' = \hat{y}$. Then (3.50) yields (3.45) and the proof of the lemma is complete. □

We continue with the proof of Theorem 3.2.3.

Fix $s \in (0, \varepsilon\eta/2]$. Since the function $p \to d(0, G(p, \bar{x}))$ is upper semicontinuous at $\bar{p}$, there exists $\delta > 0$ such that $d(0, G(p, \bar{x})) \leq s/2$ for all p with $\rho(p, \bar{p}) < \delta$. Of course, we can take smaller δ, e.g., $\delta \leq s/\varepsilon$. For such p we can find y such that $y \in G(p, \bar{x})$ with $\|y\| \leq d(0, G(p, \bar{x})) + s/3 < s$. Then we apply Lemma 3.2.4 with s, $y' = 0$ and $(p, \omega, v) = (p, \bar{x}, y)$ inasmuch as

$$\sigma((p, \bar{x}, y), (\bar{p}, \bar{x}, 0)) = \max\{\rho(p, \bar{p}), c\|y\|\} \leq \max\{\delta, cs\} \leq \max\left\{\frac{s}{\varepsilon}, cs\right\} = \frac{s}{\varepsilon} \leq \frac{\varepsilon\eta}{\varepsilon} = \eta,$$

obtaining the existence of $\hat{x}$ such that $(p, \hat{x}, 0) \in \operatorname{gph} G$; that is, $\hat{x} \in S(p)$. Also, from the estimate (3.45) with $\omega = \bar{x}$ we have that $\hat{x} \in \bar{x} + \frac{s}{\varepsilon}B^\circ$.

Set $V := B(\bar{p}, \delta)$, $U := B\left(\bar{x}, \frac{s}{\varepsilon}\right)$ and pick $p \in V$ and $x \in U$. We consider two cases.

Case 1. $d(0, G(p, x)) \geq 2s$.

We just proved that there exists $\hat{x} \in S(p)$ with $\hat{x} \in B\left(\bar{x}, \frac{s}{\varepsilon}\right)$; then

$$\begin{aligned} d(x, S(p)) &\leq d(\bar{x}, S(p)) + \|x - \bar{x}\| \leq \|\bar{x} - \hat{x}\| + \|x - \bar{x}\| \\ &\leq \frac{s}{\varepsilon} + \frac{s}{\varepsilon} = \frac{2s}{\varepsilon} \leq \frac{1}{\varepsilon} d(0, G(p, x)). \end{aligned} \tag{3.52}$$

Case 2. $d(0, G(p, x)) < 2s$.

In this case, for a sufficiently small $\gamma > 0$ we can find $y_\gamma \in G(p, x)$ such that

$$\|y_\gamma\| \leq d(0, G(p, x)) + \gamma < 2s.$$

Then

$$c\|y_\gamma\| < 2cs \leq 2c\frac{\varepsilon\eta}{2} < \eta$$

and, hence, $(p, x, y_\gamma) \in \operatorname{gph} G$ is such that $\sigma((p, x, y_\gamma), (\bar{p}, \bar{x}, 0)) \leq \eta$.

Applying Lemma 3.2.4 for $(p, \omega, v) = (p, x, y_\gamma)$, $y' = 0$ and $2s$ in place of s, we find $\hat{x}_\gamma \in S(p)$ such that

$$\|x - \hat{x}_\gamma\| \leq \frac{1}{\varepsilon}\|y_\gamma\|.$$

Then, by the choice of y_γ,

$$d(x, S(p)) \leq \|x - \hat{x}_\gamma\| \leq \frac{1}{\varepsilon}\|y_\gamma\| \leq \frac{1}{\varepsilon}(d(0, G(p, x)) + \gamma),$$

thus

$$d(x, S(p)) \leq \frac{1}{\varepsilon}(d(0, G(p, x)) + \gamma).$$

The left-hand side of this inequality does not depend on γ, hence letting $\gamma \downarrow 0$ leads to

$$d(x, S(p)) \leq \frac{1}{\varepsilon} d(0, G(p, x)). \tag{3.53}$$

We obtained this inequality also in the Case 1 in (3.52), hence it holds for any p in V and $x \in U$. Since $1/\varepsilon$ can be arbitrarily close to c, this gives us (3.42) which completes the proof of the theorem. □

The relation (3.42), obtained in Theorem 3.2.3 can be considered as *metric regularity of* G *with respect to* x at $(\bar{p}, \bar{x})$ for 0. Parallel to the partial metric regularity of G in x, we can define the partial Aubin property for G in p in the following way: $G : P \times X \rightrightarrows Y$ is said to have the *Aubin property with respect to* p *uniformly in* x at $(\bar{p}, \bar{x})$ for 0 if $0 \in G(\bar{p}, \bar{x})$ and there exist a constant $\kappa > 0$ and neighborhoods O of 0, Q for $\bar{p}$ and U of $\bar{x}$ such that

$$e(G(p, x) \cap O, G(p', x)) \le \kappa\rho(p, p') \text{ for all } \; p, p' \in Q \text{ and } x \in U.$$

By combining this definition with (3.42) one obtains (see also Ledyaev and Zhu [110, Corollary 3.9])

Proposition 3.2.5. Let $G : P \times X \rightrightarrows Y$ be both metrically regular with respect to x and have the Aubin property with respect to p uniformly in x at $(\bar{p}, \bar{x})$ for 0. Then the solution mapping S has the Aubin property at $\bar{p}$ for $\bar{x}$.

Proof. Take p, p' near $\bar{p}$ and $x \in S(p)$ near $\bar{x}$. Then we have

$$d(x, S(p')) \le \kappa' d(0, G(p', x)) \le \kappa'\kappa\rho(p, p'),$$

where κ' and κ are the constants of the assumed metric regularity and Aubin property, respectively. Since x is arbitrarily chosen in $S(p)$ near $\bar{x}$, we are done. □

3.2.2 Proof of Aubin criterion

This subsection contains the proof of Theorem 3.2.1.

For short, denote

$$d^-_{DF}(\bar{x}|\bar{y}) := \limsup_{\substack{(x,y)\to(\bar{x},\bar{y})\\(x,y)\in \operatorname{gph} F}} \|DF(x|y)^{-1}\|^-.$$

Step 1. Proof of the inequality $\operatorname{reg} F(\bar{x}|\bar{y}) \le d^-_{DF}(\bar{x}|\bar{y})$.

If $d^-_{DF}(\bar{x}|\bar{y}) = +\infty$ there is nothing to prove. Let $d^-_{DF}(\bar{x}|\bar{y}) < \infty$. Applying Theorem 3.2.3 with $P = Y$ and $G(p, x) = F(x) - p$, for y in the place of p and $\bar{y} = \bar{p}$, we have that $S(y) = F^{-1}(y)$ and $d(0, G(y, x)) = d(y, F(x))$. Then for any $c > d^-_{DF}(\bar{x}|\bar{y})$ from (3.42) we obtain that F is metrically regular at $\bar{x}$ for $\bar{y}$ with a constant c. Thus, $\operatorname{reg} F(\bar{x}|\bar{y}) \le c$ and therefore $\operatorname{reg} F(\bar{x}|\bar{y}) \le d^-_{DF}(\bar{x}|\bar{y})$ which gives us (3.34).

Step 2. Proof of $\operatorname{reg} F(\bar{x}|\bar{y}) = d^-_{DF}(\bar{x}|\bar{y})$ when X is finite dimensional.

If $\operatorname{reg} F(\bar{x}|\bar{y}) = +\infty$ we are done. Let $\operatorname{reg} F(\bar{x}|\bar{y}) < \kappa < \infty$. Then there are neighborhoods U of $\bar{x}$ and V of $\bar{y}$ such that

$$d(x, F^{-1}(y)) \leq \kappa d(y, F(x)) \text{ whenever } x \in U,\ y \in V. \tag{3.54}$$

It is obvious that when F satisfies (3.54) one can choose V so small that $F^{-1}(y) \cap U \neq \emptyset$ for all $y \in V$. Pick any $y \in V$ and $x \in F^{-1}(y) \cap U$, and let $v \in B$. Take a sequence $t_n \downarrow 0$ such that $y_n := y + t_n v \in V$ for all n. By (3.54) there exists $x_n \in F^{-1}(y + t_n v)$ such that

$$\|x - x_n\| = d(x, F^{-1}(y_n)) \leq \kappa d(y_n, F(x)) \leq \kappa \|y_n - y\| = \kappa t_n \|v\|.$$

For $u_n := (x_n - x)/t_n$ we obtain

$$\|u_n\| \leq \kappa \|v\|; \tag{3.55}$$

thus the sequence u_n is bounded and hence $u_n \to u$ for a subsequence. Since $(x_n, y + t_n v) \in \operatorname{gph} F$, by the definition of the tangent cone, we obtain $(u, v) \in T_{\operatorname{gph} F}(x, y)$ and hence, by the definition of the graphical derivative, we have $u \in DF(x|y)^{-1}(v)$. From (3.55) it follows

$$\|DF(x|y)^{-1}\|^{-} \leq \kappa.$$

Since $(x, y) \in \operatorname{gph} F$ is arbitrarily chosen near $(\bar{x}, \bar{y})$, we conclude that $d^{-}_{DF}(\bar{x}|\bar{y}) \leq \kappa$. Finally, since κ can be arbitrarily close to $\operatorname{reg} F(\bar{x}|\bar{y})$ we obtain $d^{-}_{DF}(\bar{x}|\bar{y}) \leq \operatorname{reg} F(\bar{x}|\bar{y})$. This, combined with (3.34), gives us (3.36) and Step 2 of the proof is complete.

Step 3. Proof of

$$\limsup_{\substack{(x,y)\to(\bar{x},\bar{y})\\ (x,y)\in \operatorname{gph} F}} \|D^{\star\star} F(x|y)^{-1}\|^{-} = d^{-}_{DF}(\bar{x}|\bar{y})$$

when both X and Y are finite dimensional.

Since $\operatorname{gph} D^{\star\star} F(x|y) = \operatorname{co} \operatorname{gph} DF(x|y)$, we have $D^{\star\star} F(x|y)^{-1}(v) \supset DF(x|y)^{-1}(v)$ for any v, which implies

$$\inf_{u \in D^{\star\star} F(x|y)^{-1}(v)} \|u\| \leq \inf_{u \in DF(x|y)^{-1}(v)} \|u\|,$$

consequently

$$\|D^{\star\star} F(x|y)^{-1}\|^{-} \leq \|DF(x|y)^{-1}\|^{-},$$

and then

$$\limsup_{\substack{(x,y)\to(\bar{x},\bar{y})\\ (x,y)\in \operatorname{gph} F}} \|D^{\star\star} F(x|y)^{-1}\|^{-} \leq d^{-}_{DF}(\bar{x}|\bar{y}).$$

Therefore, we only need to prove the opposite inequality

$$d^{-}_{DF}(\bar{x}|\bar{y}) \leq \limsup_{\substack{(x,y)\to(\bar{x},\bar{y})\\ (x,y)\in \operatorname{gph} F}} \|D^{\star\star} F(x|y)^{-1}\|^{-} \tag{3.56}$$

If the right hand side in (3.56) is finite, pick λ such that

$$\limsup_{\substack{(x,y)\to(\bar{x},\bar{y})\\(x,y)\in\operatorname{gph}F}} \|D^{\star\star}F(x|y)^{-1}\|^{-} < \lambda < +\infty.$$

Let $X\times Y$ be equipped with the Euclidian norm, and let $r > 0$ be small enough to ensure that

$$\max_{v\in B}\ \min_{u\in D^{\star\star}F(x|y)^{-1}(v)} \|u\| \le \lambda \quad \text{for all } (x,y)\in \operatorname{gph}F\cap B[(\bar{x},\bar{y}),r], \tag{3.57}$$

and that $\operatorname{gph}F\cap B[(\bar{x},\bar{y}),r]$ is a closed set. We will prove that

$$\max_{v\in B}\ \min_{u\in DF(x|y)^{-1}(v)} \|u\| \le \lambda \quad \text{for all } (x,y)\in \operatorname{gph}F\cap B((\bar{x},\bar{y}),r). \tag{3.58}$$

Fix $v\in B$. For any sets A, C denote by $d(A,C) := \inf\{\|a-c\| \mid a\in A, c\in C\}$. Let us fix $(x,y)\in\operatorname{gph}F\cap((\bar{x},\bar{y})+rB^{\circ})$. Let $(u^*,v^*)\in\operatorname{gph}DF(x|y)$ and $w\in\lambda B$ be such that

$$\|(w,v)-(u^*,v^*)\| = d(\lambda B\times\{v\},\operatorname{gph}DF(x|y)).$$

Observe that the point (u^*,v^*) is the unique projection of any point in the open segment $((u^*,v^*),(w,v))$ on $\operatorname{gph}DF(x|y)$. We will prove that $(u^*,v^*)=(w,v)$ and this will be enough to have (3.58) and hence (3.56).

By the definition of the graphical derivative, there exist sequences $t_n\downarrow 0$, $u_n\to u^*$, and $v_n\to v^*$ such that $y+t_nv_n\in F(x+t_nu_n)$ for all n. Let (x_n,y_n) be a point in $\operatorname{cl}\operatorname{gph}F$ which is closest to $(x,y)+\frac{t_n}{2}(u^*+w,v^*+v)$ (a projection, not necessarily unique, of the latter point on the closure of $\operatorname{gph}F$). Since $(x,y)\in\operatorname{gph}F$ we have

$$\left\|(x,y)+\frac{t_n}{2}(u^*+w,v^*+v)-(x_n,y_n)\right\| \le \frac{t_n}{2}\,\|(u^*+w,v^*+v)\|,$$

and hence

$$\begin{aligned}\|(x,y)-(x_n,y_n)\| &\le \left\|(x,y)+\frac{t_n}{2}(u^*+w,v^*+v)-(x_n,y_n)\right\|\\ &+\frac{t_n}{2}\,\|(u^*+w,v^*+v)\| \le t_n\,\|(u^*+w,v^*+v)\|\end{aligned}$$

Thus, for n sufficiently large, we have $(x_n,y_n)\in((\bar{x},\bar{y})+rB^{\circ})$ and hence $(x_n,y_n)\in\operatorname{gph}F\cap((\bar{x},\bar{y})+rB^{\circ})$. Setting $(\bar{u}_n,\bar{v}_n)=(x_n-x,y_n-y)/t_n$, we deduce by the usual property of a projection that

$$\frac{1}{2}(u^*+w,v^*+v)-(\bar{u}_n,\bar{v}_n)\in[T_{\operatorname{gph}F}(x_n,y_n)]^0=[\operatorname{gph}D^{\star\star}F(x_n|y_n)]^0,$$

where K^0 stands for the negative polar cone of a set K. Then, by (3.57), there exists $w_n\in\lambda B$ such that $v\in D^{\star\star}F(x_n|y_n)(w_n)$ and from the above relation

$$\left\langle\frac{u^*+w}{2}-\bar{u}_n,w_n\right\rangle+\left\langle\frac{v^*+v}{2}-\bar{v}_n,v\right\rangle\le 0. \tag{3.59}$$

We claim that $(\bar{u}_n, \bar{v}_n)$ converges to (u^*, v^*) as $n \to \infty$. Indeed,

$$\begin{aligned}\left\|\left(\frac{u^*+w}{2}, \frac{v^*+v}{2}\right) - (\bar{u}_n, \bar{v}_n)\right\| &= \frac{1}{t_n}\left\|(x,y) + t_n\left(\frac{u^*+w}{2}, \frac{v^*+v}{2}\right) - (x_n, y_n)\right\| \\ &\le \frac{1}{t_n}\left\|(x,y) + t_n\left(\frac{u^*+w}{2}, \frac{v^*+v}{2}\right) - (x,y) - t_n(u_n, v_n)\right\| \\ &= \left\|\left(\frac{u^*+w}{2}, \frac{v^*+v}{2}\right) - (u_n, v_n)\right\|.\end{aligned}$$

Therefore, $(\bar{u}_n, \bar{v}_n)$ is a bounded sequence and then, since $y_n = y + t_n\bar{v}_n \in F(x_n) = F(x + t_n\bar{u}_n)$, every cluster point $(\bar{u}, \bar{v})$ of it belongs to $\operatorname{gph} DF(x|y)$. Moreover, $(\bar{u}, \bar{v})$ satisfies

$$\left\|\left(\frac{u^*+w}{2}, \frac{v^*+v}{2}\right) - (\bar{u}, \bar{v})\right\| \le \left\|\left(\frac{u^*+w}{2}, \frac{v^*+v}{2}\right) - (u^*, v^*)\right\|.$$

The above inequality together with the fact that (u^*, v^*) is the unique closest point to $\frac{1}{2}(u^* + w, v^* + v)$ in $\operatorname{gph} DF(x|y)$ implies that $(\bar{u}, \bar{v}) = (u^*, v^*)$. Our claim is proved.

Up to a subsequence, w_n satisfying (3.59) converges to some $\bar{w} \in \lambda B$. Passing to the limit in (3.59) one obtains

$$\langle w - u^*, \bar{w}\rangle + \langle v - v^*, v\rangle \le 0. \tag{3.60}$$

Since (w, v) is the unique closest point of (u^*, v^*) to the closed convex set $\lambda B \times \{v\}$, we have

$$\langle w - u^*, w - \bar{w}\rangle \le 0. \tag{3.61}$$

Finally, since (u^*, v^*) is the unique closest point to $\frac{1}{2}(u^* + w, v^* + v)$ in $\operatorname{gph} DF(x|y)$ which is a closed cone, we get

$$\langle w - u^*, u^*\rangle + \langle v - v^*, v^*\rangle = 0. \tag{3.62}$$

n view of (3.60), (3.61) and (3.62), we obtain

$$\begin{aligned}\|(w,v) - (u^*, v^*)\|^2 &= \langle w - u^*, w - \bar{w}\rangle \\ &+ (\langle w - u^*, \bar{w}\rangle + \langle v - v^*, v\rangle) - (\langle w - u^*, u^*\rangle + \langle v - v^*, v^*\rangle) \le 0.\end{aligned}$$

Hence $w = u^*$ and $v = v^*$ and the proof is complete. □

3.2.3 Applications of the Aubin criterion

As a first specific application of the Aubin criterion we consider the constraint system

$$\text{Find } x \in \boldsymbol{R}^n \text{ such that } f_i(x)\begin{cases} = 0 & \text{for } i = 1, \dots, r, \\ \le 0 & \text{for } i = r+1, \dots, m,\end{cases} \tag{3.63}$$

where $f_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \dots, m$. This system can also be written as the inclusion $0 \in F(x)$ with $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ given by

$$F(x) = f(x) + K, \tag{3.64}$$

where $f = (f_1, \dots, f_m)$ and $K = \{0\}^r \times R_+^{m-r}$. Let $\bar{x}$ be a solution of (3.63) and f be strictly differentiable at $\bar{x}$. We denote the index set of active inequality constraints at $\bar{x}$ as

$$\bar{J} = \{i \in \{r+1, \dots, m\} \mid f_i(\bar{x}) = 0\}.$$

We will now show that Aubin criterion directly leads to the following well-known result:

Theorem 3.2.6. The mapping F in (3.64) is metrically regular at $\bar{x}$ for 0 if and only if the Mangasarian-Fromovitz condition holds: the vectors $\nabla f_i(\bar{x})$, $i = 1, \dots, r$ are linearly independent and also there exists $w \in \mathbb{R}^n$ such that

$$\begin{cases} \nabla f_i(\bar{x})w = 0 & \text{for } i = 1, \dots, r, \\ \nabla f_i(\bar{x})w < 0 & \text{for } i \in \bar{J}. \end{cases} \tag{3.65}$$

Proof. By Lyusternik-Graves theorem with $\mathcal{F} = K$ and $g = f$, the metric regularity of the mapping F at $\bar{x}$ for 0 is equivalent to the metric regularity at $\bar{x}$ for 0 of its "partial linearization"

$$F_0(x) = f(\bar{x}) + A(x - \bar{x}) + K \text{ where } A = \nabla f(\bar{x}).$$

Also, by the specific form of K,

$$v \in DK(x|y)(u) \iff \begin{cases} v_i = 0 & \text{for } i \in I(y), \\ v_i \geq 0 & \text{for } i \in J(y), \end{cases}$$

where

$$I(y) = \{i \in \{1, \dots, r\} \mid y_i = 0\} \text{ and } J(y) = \{i \in \{r+1, \dots, m\} \mid y_i = 0\}. \tag{3.66}$$

Then, of course, $\bar{J} = J(f(\bar{x}))$. Since $f_i(\bar{x}) < 0$ for $i \in \{r+1, \dots, m\} \setminus \bar{J}$, we have that $y_i - f_i(x) > 0$ for all such i and for (x, y) close to $(\bar{x}, 0)$. This means that for such (x, y) the set $J(y)$ in (3.66) is always a subset of $\bar{J}$. Then the Aubin criterion for metric regularity of F_0 becomes the following condition: for every $I \subset \{1, \dots, r\}$ and for every $J \subset \bar{J}$ we have:

$$\forall v \in \mathbb{R}^{I \cup J}\ \exists\, u \in \mathbb{R}^n \text{ such that } (v - Au)_i = 0 \text{ for } i \in I \text{ and } (v - Au)_i \geq 0 \text{ for } i \in J. \tag{3.67}$$

Assume that Mangasarian-Fromovitz condition holds and let $I \subset \{1, \dots, r\}$ and $J \subset \bar{J}$. If either $I = \emptyset$ or $J = \emptyset$ we skip the corresponding step of the proof. Let

$I \neq \emptyset$. Then the matrix $H = [\nabla f_i(\bar{x})]_{i\in I}$ is onto and hence, by the metric regularity of H there exists a constant κ such that

$$\forall v \in \mathbb{R}^I \quad \exists\, u \in \mathbb{R}^n \text{ such that } v - Hu = 0 \text{ and } \|u\| \leq \kappa\|v\|. \tag{3.68}$$

This means in particular that taking v with a norm small enough we can have the corresponding u in (3.68) with arbitrarily small norm. Then, since $\nabla f_i(\bar{x})w < -\alpha$ for all $i \in J$ and some $\alpha > 0$, we end up having that for any $v \in \mathbb{R}^{I\cup J}$ with sufficiently small norm

$$v_i - \nabla f_i(\bar{x})(u + w) \begin{cases} = 0 & \text{for } i \in I, \\ \geq 0 & \text{for } i \in J. \end{cases} \tag{3.69}$$

By the positive homogeneity, from (3.68) and (3.69) we obtain (3.67).

Conversely, if (3.67) holds, then taking $I = \{1, \ldots, r\}$ and $J = \emptyset$ we conclude that $\nabla f_i(\bar{x}), i = 1, \ldots, r$ must be linearly independent. Next, taking $I = \{1, \ldots, r\}$ and $J = \bar{J}$, for

$$v_i = \begin{cases} 0 & \text{for } i = 1, \ldots, r, \\ -\varepsilon & \text{for } i \in \bar{J} \end{cases}$$

with some $\varepsilon > 0$ we obtain (3.65). □

Our second application is for a mapping describing the variational inequality

$$\langle f(x), u - x\rangle \geq 0 \text{ for all } u \in C, \tag{3.70}$$

where $f : \mathbb{R}^n \to \mathbb{R}^n$ and C a nonempty convex closed set in $\mathbb{R}^n$ that is *polyhedral*. In terms of the normal cone mapping

$$N_C(x) = \begin{cases} \{y \mid \langle y, u - x\rangle \leq 0 \text{ for all } u \in C\} & \text{for } x \in C, \\ \emptyset & \text{otherwise,} \end{cases}$$

we can write the variational inequality (3.70) as the inclusion $0 \in F(x)$ where

$$F(x) = f(x) + N_C(x). \tag{3.71}$$

We assume that $\bar{x}$ is a solution of (3.70) and f is strictly differentiable at $\bar{x}$. Then, again, the Lyusternik-Graves theorem, this time with $\mathcal{F} = N_C$ and $g = f$, allows us to restrict our attention to the linearized mapping

$$F_0(x) = f(\bar{x}) + A(x - \bar{x}) + N_C(x) \quad \text{where } A = \nabla f(\bar{x}).$$

Let $[v]$ be the subspace of dimension one (or zero for $v = 0$) spanned on a vector $v \in \mathbb{R}^n$, that is, $[v] = \{\tau v \mid \tau \in \mathbb{R}\}$, and let $[v]^\perp$ be its orthogonal complement. The form of the graphical derivative of F_0 will be obtained by introducing the *critical cone* $K(x, v)$ to the set C at $x \in C$ for $v \in N_C(x)$,

$$K(x, v) = T_C(x) \cap [v]^{\perp},$$

via the following

Reduction lemma in Dontchev and Rockafellar [64]. Let C be a convex polyhedral set in $\mathbb{R}^n$. For any $(x, v) \in \operatorname{gph} N_C$ there is a neighborhood O of the origin in $\mathbb{R}^n \times \mathbb{R}^n$ such that for $(x', v') \in O$ one has

$$v + v' \in N_C(x + x') \iff v' \in N_{K(x,v)}(x').$$

Consequently,

$$(x', v') \in T_{\operatorname{gph} N_C}(x, v) \iff v' \in N_{K(x,v)}(x'),$$

and hence, for $(x, y) \in \operatorname{gph} F_0$ and $v = y - f(\bar{x}) - A(x - \bar{x})$, where $A = \nabla f(\bar{x})$, we have

$$DF_0(x|y)(u) = Au + N_{K(x,v)}(u). \tag{3.72}$$

For any cone K, a set of the form

$$F = K \cap [v]^{\perp} \text{ for some } v \in K^0,$$

where K^0 is the polar to K, is said to be a *face* of K. The largest of the faces is K itself while the smallest is the set $K \cap (-K)$ which is the largest subspace contained in K. Every polyhedral cone has finitely many faces.

It was proved in Dontchev and Rockafellar [64, Theorem 1] that the metric regularity of a mapping F of the form (3.71) with a polyhedral set C implies a sharper property called strong regularity. A mapping $F : X \rightrightarrows Y$ is said to be *strongly regular* at $\bar{x}$ for $\bar{y}$ is it is metrically regular there and, in addition, the graphical localization of its inverse F^{-1} near $(\bar{y}, \bar{x})$ is single-valued. In other words, F is strongly regular at $\bar{x}$ for $\bar{y}$ when there are neighborhoods U of $\bar{x}$ and V of $\bar{y}$ such that the mapping $V \ni y \mapsto F^{-1}(y) \cap U$ is a Lipschitz continuous function.

We are now ready to apply the Aubin criterion to obtain a new necessary and sufficient condition for strong regularity of variational inequalities over polyhedral sets, which complements the criterion given in Dontchev and Rockafellar [64, Theorem 2]:

Theorem 3.2.7. The variational inequality mapping (3.71) is strongly regular at $\bar{x}$ for $\bar{y}$ if and only if for all choices of faces F_1 and F_2 of the critical cone $\bar{K}$ to the set C at $\bar{x}$ for $\bar{v} = \bar{y} - f(\bar{x})$, with $F_1 \supset F_2$, the following condition holds:

$$\forall\ v \in \boldsymbol{R}^n \quad \exists\ u \in F_1 - F_2 \text{ such that } (v - Au) \in (F_1 - F_2)^0 \text{ and } v - Au \perp u.$$

Proof. According to Aubin criterion given in Theorem 2.74, the mapping F_0 (and, hence F) is metrically regular, and hence strongly regular, if and only if the limsup of the inner norms of the graphical derivatives is finite. The form of the graphical derivative of F_0 is given in (3.72).

On the other hand, next result gives the form of the of critical cones in a neighborhood of a fixed reference point. It is extracted from the proof of Dontchev and Rockafellar [64, Theorem 2]: Let C be a convex polyhedral set, let $\bar{v} \in N_C(\bar{x})$ and let $\bar{K}$ be the critical cone to C at $\bar{x}$ for $\bar{v}$. Then there exists an open neighborhood O of $(\bar{x}, \bar{v})$ such that for every choice of $(x, v) \in \operatorname{gph} N_C \cap O$ the corresponding critical cone $K(x, v)$ has the form

$$K(x, v) = F_1 - F_2,$$

for some faces F_1, F_2 of $\bar{K}$ with $F_1 \supset F_2$. And conversely, for every two faces F_1, F_2 of $\bar{K}$ with $F_1 \supset F_2$ and every neighborhood O of $(\bar{x}, \bar{v})$ there exists $(x, v) \in O$ such that $K(x, v) = F_1 - F_2$.

So, there are finitely many critical cones near the reference point $(\bar{x}, \bar{v})$ that are to be taken into account, and these cones are given by faces of $\bar{K}$ in a way described in this result. Hence, for any choice of faces F_1 and F_2 of $\bar{K}$ with $F_1 \supset F_2$ it is enough to ensure that $\|A + N_{F_1 - F_2}\|^-$ is finite. For any cone K

$$v \in N_K(x) \iff x \in K, \quad v \in K^0, \quad x \perp v.$$

It remains to observe that the inner norm of the mapping $A + N_{F_1 - F_2}$ will be finite if and only if the condition claimed in the theorem holds. □

Our last application of Aubin criterion is a new proof of the radius theorem first proved by Dontchev, Lewis and Rockafellar in [62].

Theorem 3.2.8. Let X and Y be finite-dimensional linear normed spaces and let $F : X \rightrightarrows Y$ has closed graph locally near $(\bar{x}, \bar{y}) \in \operatorname{gph} F$. Then

$$\inf_{G \in L(X,Y)} \{\|G\| \mid F + G \text{ is not metrically regular at } \bar{x} \text{ for } \bar{y} + G(\bar{x})\} = \frac{1}{\operatorname{reg} F(\bar{x}|\bar{y})}.$$

Moreover, the infimum is unchanged if taken with respect to linear mappings G of rank 1, but also remains unchanged when the perturbations G are locally Lipschitz continuous functions with $\|G\|$ replaced by the Lipschitz modulus $\operatorname{lip} G(\bar{x})$ of G at $\bar{x}$.

Proof. The general perturbation inequality derived in Dontchev, Lewis and Rockafellar [62, Corollary 3.4] yields (also in infinite dimensions) the estimate

$$(3.73) \quad \inf_{G: X \to Y} \{\operatorname{lip} G(\bar{x}) \mid F + G \text{ is not metrically regular at } \bar{x} \text{ for } \bar{y} + G(\bar{x})\} \geq \frac{1}{\operatorname{reg} F(\bar{x}|\bar{y})}.$$

It remains to show the opposite inequality. The limit cases are easy to handle, since if $\operatorname{reg} F(\bar{x}|\bar{y}) = \infty$ we have nothing to prove, and if $\operatorname{reg} F(\bar{x}|\bar{y}) = 0$, then by the general perturbation inequality (3.73), which also holds in this case, we obtain the claimed equality.

Let now $0 < \operatorname{reg} F(\bar{x}|\bar{y}) < \infty$. By Theorem 3.2.1 we have $\operatorname{reg} F(\bar{x}|\bar{y}) = d^-_{DF}(\bar{x}|\bar{y}) = d^-_{D^{\star\star}F}(\bar{x}|\bar{y})$ where $d^-_{DF}(\bar{x}|\bar{y})$ is defined in the beginning of Subsection 3.2.2 while $d^-_{D^{\star\star}F}(\bar{x}|\bar{y})$ is defined in the same way with DF replaced by $D^{\star\star}F$.

Take a sequence of positive reals $\varepsilon_k \to 0$. Then for any k there exists $(x_k, y_k) \in \operatorname{gph} F$ with $(x_k, y_k) \to (\bar{x}, \bar{y})$ and

$$d^-_{D^{\star\star}F}(\bar{x}|\bar{y}) + \varepsilon_k \geq \|D^{\star\star}F(x_k|y_k)^{-1}\|^- \geq d^-_{D^{\star\star}F}(\bar{x}|\bar{y}) - \varepsilon_k > 0.$$

For short, set $H_k := D^{\star\star}F(x_k|y_k)$; then H_k is a sublinear mapping with closed graph. For $S_k := H_k^{*+}$ the norm duality gives us $\|H_k^{-1}\|^- = \|S_k^{-1}\|^+$.

For each k choose a positive real r_k which satisfies $\|S_k^{-1}\|^+ - \varepsilon_k < 1/r_k < \|S_k^{-1}\|^+$. From the last inequality there must exist $(\hat{y}_k, \hat{x}_k) \in \operatorname{gph} S_k$ with $\|\hat{x}_k\| = 1$ and $\|S_k^{-1}\|^+ \geq \|\hat{y}_k\| > 1/r_k$. Pick $y_k^* \in Y$ with $\langle \hat{y}_k, y_k^* \rangle = \|\hat{y}_k\|$ and $\|y_k^*\| = 1$ and define the rank-one mapping $\hat{G}_k \in L(Y, X)$ as

$$\hat{G}_k(y) := -\frac{\langle y, y_k^* \rangle}{\|\hat{y}_k\|}\hat{x}_k.$$

Then $\hat{G}_k(\hat{y}_k) = -\hat{x}_k$ and hence $(S_k + \hat{G}_k)(\hat{y}_k) = S_k(\hat{y}_k) + \hat{G}_k(\hat{y}_k) = S_k(\hat{y}_k) - \hat{x}_k \ni 0$. Therefore, $\hat{y}_k \in (S_k + \hat{G}_k)^{-1}(0)$ and since $\hat{y}_k \neq 0$, by Dontchev, Lewis and Rockafellar [62, Proposition 2.5],

$$\|(S_k + \hat{G}_k)^{-1}\|^+ = \infty. \tag{3.74}$$

Note that $\|\hat{G}_k\| = \|\hat{x}_k\|/\|\hat{y}_k\| = 1/\|\hat{y}_k\| < r_k$.

Since the sequences $\hat{y}_k$, $\hat{x}_k$ and y_k^* are bounded, we can extract from them subsequences converging respectively to $\hat{y}$, $\hat{x}$ and y^*; the limits then satisfy $\|\hat{y}\| = d^-_{D^{\star\star}F}(\bar{x}|\bar{y})$, $\|\hat{x}\| = 1$ and $\|y^*\| = 1$. Define the rank-one mapping $\hat{G} \in L(Y, X)$ as

$$\hat{G}(y) := -\frac{\langle y, y^* \rangle}{\|\hat{y}\|}\hat{x}.$$

Then we have $\|\hat{G}\| \leq 1/d^-_{D^{\star\star}F}(\bar{x}|\bar{y})$ and $\|\hat{G}_k - \hat{G}\| \to 0$.

Denote $G := (\hat{G})^*$ and suppose that $F + G$ is metrically regular at $\bar{x}$ for $\bar{y} + G(\bar{x})$. Then Theorem 3.2.1 yields that for some finite positive constant c and for k sufficiently large we have

$$c > \|D^{\star\star}(F + G)(x_k|y_k + G(x_k))^{-1}\|^- = \|(D^{\star\star}F(x_k|y_k) + G)^{-1}\|^-,$$

which, by norm duality and the equality $G^* = ((\hat{G})^*)^* = \hat{G}$, is equivalent to

$$c > \|([D^{\star\star}F(x_k|y_k) + G]^{*+})^{-1}\|^+ = \|[D^{\star\star}F(x_k|y_k)^{*+} + G^*]^{-1}\|^+ = \|(S_k + \hat{G})^{-1}\|^+. \tag{3.75}$$

We apply the following lemma, which is a reformulation of a result by Robinson [141], see also Dontchev and Lewis [61]:

Lemma 3.2.9. For a sublinear mapping $H : X \to Y$ with closed graph and for $B \in L(X, Y)$, if $[\|H^{-1}\|^+]^{-1} \geq \|B\|$, then

$$\|(H + B)^{-1}\|^+ \leq [[\|H^{-1}\|^+]^{-1} - \|B\|]^{-1}.$$

Now we are ready to complete the proof of Theorem 3.2.8. Take k sufficiently large such that $\|\hat{G} - \hat{G}_k\| \leq 1/(2c)$ and (x_k, y_k) that satisfies (3.75). Setting $P_k := S_k + \hat{G}$ and $B_k := \hat{G}_k - \hat{G}$ we have that $[\|P_k^{-1}\|^+]^{-1} \geq 1/c > 1/(2c) \geq \|B_k\|$. By Lemma 3.2.9 we obtain

$$\|(S_k + \hat{G}_k)^{-1}\|^+ = \|(P_k + B_k)^{-1}\|^+ \leq [[\|P_k^{-1}\|^+]^{-1} - \|B_k\|]^{-1} \leq 2c < \infty,$$

which contradicts (3.74). Hence, $F + G$ is not metrically regular at $\bar{x}$ for $\bar{y} + G(\bar{x})$. Remembering that $\|G\| = \|\hat{G}\| \leq 1/\operatorname{reg} F(\bar{x}|\bar{y})$ we complete the proof. □

3.3 Long orbit or empty value principle, fixed point and surjectivity theorems

It is well known that there exist close relations between iteration schemes related to dissipative mechanical systems, Ekeland perturbed minimization principle, fixed point theorems and inverse and implicit function theorems of all kinds, see for example Aubin and Ekeland [8].

In the present section we explore some of these relations in a new light. Generalizing from Ivanov [92] (see also e.g. Fabian and Priess [72], Penôt [132, p.62]), we obtain a flexible Long Orbit or Empty Value (LOEV) principle, see Theorem 3.3.2 and Corollary 3.3.3.

The reader may notice certain similarity between the conditions imposed on the map S in LOEV principle and that in Caristi-Kirk Fixed Point Theorem. Indeed, the latter readily follows from the former, see Theorem 3.3.4.

The rest of this section is devoted to surjectivity results. These are derived from a novel Theorem 3.3.8 which may be regarded as a kind of interpolation between classical Graves Theorem and a recent result of Ekeland [69]. However, we do not require differentiability in any usual sense, exploring instead a generalization of so called contingent derivative, see e.g. Aubin and Frankowska [10].

Theorem 3.3.8 can also be considered as generalization of Aubin derivative criterion for metric regularity of a single-valued map.

3.3.1 Long Orbit or Empty Value (LOEV) principle

Everywhere in this subsection (M, ρ) denotes complete metric space.

Definition 3.3.1. Let $S : M \rightrightarrows M$ be a multivalued map. We say that S satisfies the condition $(*)$ if $x \notin S(x)$, $\forall x \in M$, and whenever $y \in S(x)$ and $\lim_n x_n = x$, there are infinitely many x_n's such that

$$y \in S(x_n).$$

The following results, which constitute LOEV principle, suggest why maps satisfying $(*)$ could be useful.

Theorem 3.3.2 (LOEV principle). Let $S : M \rightrightarrows M$ satisfy $(*)$ and let $x_0 \in M$ be arbitrary.

Then at least one of (a) and (b) below is true:

(a) There are $x_i \in M$, $i = 1, 2, \ldots$, such that

$$x_{i+1} \in S(x_i), \quad i = 0, 1, \ldots,$$

and

$$\sum_{i=0}^{\infty} \rho(x_i, x_{i+1}) = \infty;$$

(b) There is $\bar{x} \in M$ such that

$$S(\bar{x}) = \varnothing.$$

Proof. Assume (a) was not true, that is, each S-orbit, starting at x_0, is of finite length.

We can construct finite or infinite orbit $(x_i)_{i \geq 0} \subset M$ by the following procedure.

If $x_0, x_1, \ldots, x_i$ are already chosen, then

either $S(x_i) = \varnothing$ and we are done;

or $s_i := \min\{1, \sup\{\rho(x_i, y) : \ y \in S(x_i)\}\} > 0$. Take $x_{i+1} \in S(x_i)$ such that

$$\rho(x_i, x_{i+1}) > \frac{s_i}{2}. \tag{3.76}$$

If we end up with infinite orbit then, since (a) was assumed false,

$$\sum_{i=0}^{\infty} \rho(x_i, x_{i+1}) < \infty \Rightarrow \lim_{n \to \infty} x_n =: \bar{x}.$$

In particular, $\lim_i \rho(x_i, x_{i+1}) = 0$ and from (3.76) it follows that $s_i \to 0$.

Assume that $S(\bar{x}) \neq \varnothing$.

Take $\bar{y} \in S(\bar{x})$. By $(*)$ we have $\rho(\bar{y}, \bar{x}) > 0$ and $\bar{y} \in S(x_i)$ for infinitely many i's. By the definition of s_i we have that $s_i \geq \rho(\bar{y}, x_i)$ for infinitely many i's. Passing to limit over the latter subsequence we get $0 \geq \rho(\bar{y}, \bar{x}) > 0$. Contradiction. □

Corollary 3.3.3. Let $S : M \rightrightarrows M$ satisfy $(*)$ and let $x_0 \in M$ and $K > 0$ be arbitrary.
Then at least one of the two conditions below is true:
(a) There are $x_i \in M$, $i = 1, 2, \dots, n+1$, such that

$$x_{i+1} \in S(x_i), \quad i = 0, 1, \dots, n,$$

and

$$\sum_{i=0}^{n} \rho(x_i, x_{i+1}) > K;$$

(b) There is $\bar{x} \in M$ such that $\rho(x_0, \bar{x}) \le K$ and

$$S(\bar{x}) = \varnothing.$$

Proof. Assume that (b) is not true, that is $S(x) \neq \varnothing$ for all $x \in M$ such that $\rho(x_0, x) \le K$.
If (a) from Theorem 3.3.2 is true, we have nothing to prove.
If not, from the proof of Theorem 3.3.2 we get a $\bar{x} \in M$ such that $S(\bar{x}) = \varnothing$ and $\bar{x}$ is either the end point of finite orbit or the limit of infinite orbit $(x_i)_{i \ge 0}$. Since $\rho(x_0, \bar{x}) > K$, by triangle inequality we have that

$$\sum_{i \ge 0} \rho(x_i, x_{i+1}) > K. \quad \square$$

3.3.2 Caristi-Kirk fixed point theorem

As first application of LOEV principle we prove the following Theorem due to Caristi and Kirk (see Kirk [100] and Caristi [42]).

Theorem 3.3.4 (Caristi-Kirk Fixed Point Theorem). Let (M, ρ) be complete metric space. Let $T : M \rightrightarrows M$ satisfy $T(x) \neq \varnothing$ for all $x \in M$. Let, moreover, the function $f : M \to [0, \infty)$ be lower semicontinuous. Assume that for any $x \in M$ there exist $y \in T(x)$ such that

$$\rho(x, y) \le f(x) - f(y). \tag{3.77}$$

Then T has a fixed point, that is, there exists $\bar{x}$ such that $\bar{x} \in T(\bar{x})$.

Proof. Assume that $x \notin T(x)$ for all $x \in M$.
Define for $x \in M$

$$S(x) := \{y \in M : \ f(y) < f(x) - 2^{-1}\rho(x, y)\}.$$

We are given that for each $x \in M$ there is a $y \in T(x)$ satisfying (3.77). By assumption $\rho(x, y) > 0$ and then (3.77) ensures that $S(x) \neq \varnothing$.

The lower semicontinuity of f implies that S satisfies $(*)$.

From Theorem 3.3.2 we get an orbit $(x_n)_0^\infty$, $x_{i+1} \in S(x_i)$, with

$$\sum_{i=0}^{\infty} \rho(x_i, x_{i+1}) = \infty.$$

But the definition of S implies $f(x_{i+1}) - f(x_i) < -\rho(x_{i+1}, x_i)/2$ and summing the latter we get $f(x_i) \to -\infty$. Contradiction with $f \geq 0$. □

3.3.3 Surjectivity theorems

Our second application of LOEV principle are the following two Lemmas, which can be regarded as a kind of quantitative version of Theorem 3.3.4. Later on we will apply these results for obtaining several surjectivity results for functions.

Our notation is standard. $(X, \|\cdot\|)$ and $(Y, \|\cdot\|)$ are Banach spaces with dual spaces X^* and Y^*, respectively. B_X is the closed unit ball of the Banach space X, respectively B_X° is the open unit ball. The set U is a neighbourhood of the point $x_0 \in X$ if $x_0 + \varepsilon B_X \subset U$ for some $\varepsilon > 0$, that is, x_0 is interior point to U, and so on.

Lemma 3.3.5. Let $\alpha, \beta, \gamma > 0$. Let X and Y be Banach spaces. Let $f : X \to Y$ be a continuous on βB_X function such that $f(0) = 0$.

Suppose that for all $y \in Y$ such that $\|y\| < \alpha$ and all $x \in X$ such that $\|x\| < \beta$ and $f(x) \neq y$, there is $z \in X$ with $\|z\| < \beta$ and

$$\|f(z) - y\| < \|f(x) - y\| - \gamma\|z - x\|. \tag{3.78}$$

Then for every $\bar{y} \in Y$ satisfying $\|\bar{y}\| < \min\{\alpha, \beta\gamma\}$ there exists $\bar{x} \in X$ with $f(\bar{x}) = \bar{y}$ and

$$\|\bar{x}\| \leq \frac{\|\bar{y}\|}{\gamma}.$$

Proof. Obviously, if $\bar{y} = 0$ we can take $\bar{x} = 0$.

Fix $\bar{y} \in Y$ such that $\bar{y} \neq 0$ and $\|\bar{y}\| < \alpha$.

For $x \in \beta B_X$ define

$$S(x) := \{z \in \beta B_X : \ \|f(z) - \bar{y}\| < \|f(x) - \bar{y}\| - \gamma\|z - x\|\}$$

and note that S satisfies $(*)$ on βB_X due to continuity of f.

It is given that

$$S(x) \neq \varnothing, \quad \forall x \in X \text{ such that } \|x\| < \beta \text{ and } f(x) \neq \bar{y}. \tag{3.79}$$

Now we apply Corollary 3.3.3 for the space $M = \beta B_X$, the map S, the initial point $x_0 = 0$ and the constant $K = \dfrac{\|\bar{y}\|}{\gamma}$.

So, if we assume that there is finite orbit $(x_i)_{i=0}^{n+1}$ such that $x_{i+1} \in S(x_i)$ for $i = 0, \ldots, n$, and

$$\sum_{i=0}^{n} \|x_{i+1} - x_i\| > \frac{\|\bar{y}\|}{\gamma} \iff \gamma \sum_{i=0}^{n} \|x_{i+1} - x_i\| > \|\bar{y}\|,$$

then by the definition of S and $f(0) = 0$

$$\|f(x_{n+1}) - \bar{y}\| - \|\bar{y}\| < -\gamma \sum_{i=0}^{n} \|x_{i+1} - x_i\| < -\|\bar{y}\| \Rightarrow \|f(x_{n+1}) - \bar{y}\| < 0,$$

contradiction. Thus from Corollary 3.3.3 it follows that there is $\bar{x} \in KB_X$ such that $S(\bar{x}) = \varnothing$. Since $K < \beta$, (3.79) implies that $f(\bar{x}) = \bar{y}$. □

Let us recall that the function $f : X \to Y$ is *metrically regular* at $\bar{x} \in X$ if there exists $\kappa > 0$ together with neighbourhoods $U \ni \bar{x}$ and $V \ni f(\bar{x})$ such that

$$d(x, f^{-1}(y)) \le \kappa d(y, f(x)), \quad \forall x \in U, \ \forall y \in V,$$

see Dontchev and Rockafellar [67, p.253 (2)]. The infimum of κ over all such combinations of κ, U and V is called the *regularity modulus* for f at $\bar{x}$ for $f(\bar{x})$ and denoted by $\operatorname{reg}(f; \bar{x})$. The absence of metric regularity is signaled by $\operatorname{reg}(f; \bar{x}) = \infty$.

Recall also that metric regularity is equivalent to openness at linear rate, see Dontchev and Rockafellar [67, p.254 (5)].

Lemma 3.3.6. Let the assumptions of Lemma 3.3.5 be satisfied. Then there exist a neighbourhood $U \ni 0$ in X and a neighbourhood $V \ni 0$ in Y such that for every $x' \in U$ and every $\bar{y} \in V$ there exists $\bar{x} \in X$ with $f(\bar{x}) = \bar{y}$ and

$$\|x' - \bar{x}\| \le \frac{\|f(x') - \bar{y}\|}{\gamma}.$$

That is, f is metrically regular at 0 with $\operatorname{reg}(f; 0) \le \gamma^{-1}$.

Proof. From continuity of f at 0 there exist $\delta \le \beta/4$ and $\eta < \alpha$ such that for all $x' \in \delta B_X$ and all $\bar{y} \in \eta B_Y$ it holds $\|f(x') - \bar{y}\| \le \beta\gamma/4$. Set $U = \delta B_X$, $V = \eta B_Y$.

Fix $x' \in U$ and $\bar{y} \in V$. If $f(x') = \bar{y}$, take $\bar{x} = x'$. Let now $f(x') \neq \bar{y}$.

For $x \in \beta B_X$ define

$$S(x) := \{z \in \beta B_X : \ \|f(z) - \bar{y}\| < \|f(x) - \bar{y}\| - \gamma\|z - x\|\}$$

and note that S satisfies $(*)$ on βB_X due to continuity of f.

It is given that

$$S(x) \neq \emptyset, \quad \forall x \in X \text{ such that } \|x\| < \beta \text{ and } f(x) \neq \bar{y}. \tag{3.80}$$

Now we apply Corollary 3.3.3 for the space $M = \beta B_X$, the map S, the initial point $x_0 = x'$ and the constant $K = \dfrac{\|f(x') - \bar{y}\|}{\gamma}$.

So, if we assume that there is finite orbit $(x_i)_{i=0}^{n+1}$ such that $x_{i+1} \in S(x_i)$ for $i = 0, \ldots, n$, and

$$\sum_{i=0}^{n} \|x_{i+1} - x_i\| > \frac{\|f(x') - \bar{y}\|}{\gamma} \iff \gamma \sum_{i=0}^{n} \|x_{i+1} - x_i\| > \|f(x') - \bar{y}\|,$$

then by the definition of S

$$\|f(x_{n+1}) - \bar{y}\| - \|f(x') - \bar{y}\| < -\gamma \sum_{i=0}^{n} \|x_{i+1} - x_i\| < -\|f(x') - \bar{y}\|,$$

hence $\|f(x_{n+1}) - \bar{y}\| < 0$, contradiction. Thus from Corollary 3.3.3 it follows that there is $\bar{x} \in x' + KB_X$ such that $S(\bar{x}) = \emptyset$. Since $\|\bar{x}\| \leq \|\bar{x} - x'\| + \|x'\| \leq K + \beta/4 \leq \beta/4 + \beta/4 < \beta$, (3.80) implies that $f(\bar{x}) = \bar{y}$. □

Before proceeding further we have to recall several notions.

The multivalued map $H : X \rightrightarrows Y$ is called *positively homogeneous* if its graph

$$\operatorname{gph} H := \{(x, y) \in X \times Y : \ y \in H(x)\}$$

is a cone. That is, $y \in H(x) \iff ty \in H(tx)$ for all $t \geq 0$.

The *inner norm* of such H, see Dontchev and Rockafellar [67, p.256], is

$$\|H\|^- = \sup_{\|x\| \leq 1} \inf_{y \in H(x)} \|y\|.$$

The *inverse* of H is the positively homogeneous map $H^{-1} : Y \rightrightarrows X$ such that $H^{-1}(y)$ is the set of solutions of $y \in H(x)$.

Obviously, if $\|H\|^- < \infty$ then H is surjective and H^{-1} is everywhere defined. Moreover, if $\|H^{-1}\|^- < \kappa$ then

$$\forall v \in Y \setminus \{0\} \ \exists u \in H^{-1}(v) : \ \|u\| < \kappa \|v\|. \tag{3.81}$$

Below we will generalize the *contingent cone* for the case of single valued map, that is *function*. But in order to make more sense, we will first recall the notion of *contingent derivative* of a function, see Aubin and Ekeland [8].

If $f : X \to Y$ is a function then the *contingent derivative*

$$df(x) : X \rightrightarrows Y$$

of f at x is the contingent cone to $\operatorname{gph} f = \{(x,y) : \ y = f(x)\}$ at $(x, f(x))$. That is, $v \in df(x)(u)$ if and only if there are $u_n \to u$, $v_n \to v$ and $t_n > 0$ with $t_n \to 0$, such that

$$f(x + t_n u_n) = f(x) + t_n v_n.$$

The graph of $df(x)$ is closed cone. So, $df(x)$ is positively homogeneous map.

Definition 3.3.7. Let X and Y be Banach spaces. Let $\theta > 0$ and $\mu \geq 0$. We consider the following equivalent norm on $X \times Y$:

$$\|(x,y)\|_\theta := \theta\|x\| + \|y\|.$$

The *approximate contingent derivative* $d^\mu_\theta f(x)$ of f at x is defined by: $v \in d^\mu_\theta f(x)(u)$ if there are $(u_n, v_n) \to (u, v)$ and $t_n > 0$ with $t_n \to 0$ such that

$$\operatorname{dist}_\theta\left((x + t_n u_n, f(x) + t_n v_n), \operatorname{gph} f\right) \leq \mu t_n \|u_n\|, \tag{3.82}$$

where $\operatorname{dist}_\theta$ means the distance measured in norm $\|(\cdot,\cdot)\|_\theta$.

It is clear that $d^\mu_\theta f \supset df$. Obviously, $d^0_\theta f = df$ for all $\theta > 0$.

Roughly speaking, the following general result states that if the approximate contingent derivative is linearly open at *uniform* rate in a neighbourhood, then so is the function itself. Several known results follow from this. They are presented at the end of the section.

Theorem 3.3.8. Let X and Y be Banach spaces. Let $f : X \to Y$ be continuous on $x_0 + \varepsilon B_X$ for some $\varepsilon > 0$. Suppose that $\theta > 0$, and $\mu \geq 0$, $m > 0$ with $m\mu < 1$ are such that

$$\|d^\mu_\theta f(x)^{-1}\|^- \leq m, \quad \forall x \in x_0 + \varepsilon B^\circ_X.$$

Then for any $\kappa > m$ for $c := \kappa^{-1} - \mu$ it holds that for any $y \in Y$ such that $\|y - f(x_0)\| < \varepsilon \min\{c, \theta\}$ there exists $x \in X$ such that

$$\|x - x_0\| \leq \frac{\|y - f(x_0)\|}{\min\{c, \theta\}} < \varepsilon,$$

and $f(x) = y$.

Moreover, f is metrically regular at x_0 with $\operatorname{reg}(f; x_0) \leq \max\left\{\dfrac{m}{1 - m\mu}, \theta^{-1}\right\}$.

Proof. We assume without loss of generality that $x_0 = 0$ and $f(0) = 0$.

Take $\kappa > m$.

Let $x \in \varepsilon B_X^\circ$ and $y \neq f(x)$.

Apply (3.81) to $d_\theta^\mu f(x)$ and $v = y - f(x) \neq 0$ to get $u \in X$ such that

$$\|u\| < \kappa\|v\| \text{ s.t. } d_\theta^\mu f(x)(u) \ni v.$$

By Definition 3.3.7 here are $(u_n, v_n) \to (u, v)$ and $t_n > 0$ with $t_n \to 0$ such that (3.82) holds. Thus, there are $x_n \in X$ with

$$\theta\|x + t_n u_n - x_n\| + \|f(x) + t_n v_n - f(x_n)\| \le \mu t_n\|u_n\| + t_n^2.$$

If we omit the second addend at left hand side, we will get $\theta\|x + t_n u_n - x_n\| \le O(t_n)$ and, therefore, $x_n \to x$. So, we may assume that $\|x_n\| < \varepsilon$.

We now add and subtract y within the second addend at the left hand side above, use $v = y - f(x)$ and rearrange to get

$$\| - v + t_n v_n + y - f(x_n)\| \le \mu t_n\|u_n\| - \theta\|x + t_n u_n - x_n\| + t_n^2.$$

We add and subtract $t_n v$ within left hand side and use triangle inequality to get

$$\|v - t_n v + f(x_n) - y\| - t_n\|v - v_n\| \le \mu t_n\|u_n\| - \theta\|x + t_n u_n - x_n\| + t_n^2.$$

Applying once again triangle inequality:

$$\|f(x_n) - y\| - (1 - t_n)\|v\| \le t_n\|v - v_n\| + \mu t_n\|u_n\| - \theta\|x + t_n u_n - x_n\| + t_n^2.$$

That is,

$$\|f(x_n) - y\| - \|v\| \le -t_n\|v\| + \mu t_n\|u_n\| - \theta\|x + t_n u_n - x_n\| + \delta_n t_n, \tag{3.83}$$

where $\delta_n \to 0$. We consider two cases.

Case 1. $u = 0$. Then $u_n \to 0$ and

$$\begin{aligned} \|f(x_n) - y\| - \|v\| &\le -t_n\|v\| - \theta\|x - x_n\| + \delta_n' t_n, \text{ where } \delta_n' \to 0 \\ &< -\theta\|x - x_n\|, \end{aligned}$$

for all n large enough, since $\|v\| > 0$. Finally, in this case we get

$$\|f(x_n) - y\| - \|f(x) - y\| < -\theta\|x - x_n\|, \quad n > N_1. \tag{3.84}$$

Case 2. $u \neq 0$. Since $t_n\|u_n\| = t_n\|u\| + o(t_n)$, (3.83) becomes

$$\|f(x_n) - y\| - \|v\| \le -t_n(\|v\| - \mu\|u\|) - \theta\|x + t_n u - x_n\| + \delta_n'' t_n,$$

where $\delta_n'' \to 0$. Since $\|u\| < \kappa\|v\|$, we have $\|v\| - \mu\|u\| > (\kappa^{-1} - \mu)\|u\|$. Recall that $c = \kappa^{-1} - \mu > 0$ to get from the above:

$$\|f(x_n) - y\| - \|v\| < -ct_n\|u\| - \theta\|x + t_n u - x_n\|$$

for n large enough. Obviously, $-ct_n\|u\| - \theta\|x + t_n u - x_n\| \leq -\min\{c, \theta\}(t_n\|u\| + \|x + t_n u - x_n\|)$ and by triangle inequality

$$\|f(x_n) - y\| - \|f(x) - y\| < -\min\{c, \theta\}\|x - x_n\|, \quad n > N_2. \tag{3.85}$$

From (3.84) and (3.85) we conclude that there is $z \in X$ with $\|z\| < \varepsilon$ such that

$$\|f(z) - y\| - \|f(x) - y\| < -\min\{c, \theta\}\|x - z\|.$$

With $\alpha = \infty$, $\beta = \varepsilon$ and $\gamma = \min\{c, \theta\}$ we apply Lemma 3.3.5 to get the first conclusion and Lemma 3.3.6 to get that

$$\operatorname{reg}(f; x_0) \leq (\min\{c, \theta\})^{-1} = \max\{c^{-1}, \theta^{-1}\} = \max\left\{\frac{\kappa}{1 - \kappa\mu}, \frac{1}{\theta}\right\}$$

and since the latter holds for all $\kappa > m$, we finally obtain that

$$\operatorname{reg}(f; x_0) \leq \max\left\{\frac{m}{1 - m\mu}, \frac{1}{\theta}\right\}. \quad \square$$

With the help of Theorem 3.3.8 we can proof with an unified approach a lot of surjectivity results. First of them will be the classical Graves theorem.

Let X and Y be Banach spaces and A be bounded linear surjection from X onto Y. Recall, e.g. Dontchev and Rockafellar [67, p.254], that

$$\operatorname{reg} A = \sup_{\|y\| \leq 1} d(0, A^{-1}(y)),$$

where $A^{-1}(y) = \{x \in X : \; Ax = y\}$. Note that $\operatorname{reg} A < \infty$ by Banach Open Mapping Theorem. In the sequel we use that if $\kappa > \operatorname{reg} A$ then for any $y \in Y$, $y \neq 0$ there is $x \in X$ with $\|x\| < \kappa\|y\|$ and $Ax = y$.

Theorem 3.3.9 (Graves, e.g. p. 276–278 in [67]). Let X and Y be Banach spaces. Consider a function $f : X \to Y$ continuous on $\bar{x} + \varepsilon B_X$ for some $\bar{x} \in X$ and $\varepsilon > 0$. Let A be bounded linear and surjective from X onto Y and let $m = \operatorname{reg} A$. Suppose that there is $\mu \geq 0$ such that $m\mu < 1$ and

$$\|f(x) - f(x') - A(x - x')\| \leq \mu\|x - x'\|, \quad \forall x, x' \in \bar{x} + \varepsilon B_X. \tag{3.86}$$

Set $\bar{y} := f(\bar{x})$. For any $\kappa > m$ and $c := \kappa^{-1} - \mu$, if y is such that $\|y - \bar{y}\| < c\varepsilon$, then the equation $y = f(x)$ has a solution x such that

$$\|x - \bar{x}\| \le \frac{\|y - \bar{y}\|}{c},$$

in particular $\|x - \bar{x}\| < \varepsilon$.

Moreover, f is metrically regular at $\bar{x}$ with $\operatorname{reg}(f;\bar{x}) \le \dfrac{m}{1 - m\mu}$.

Proof. We will show that (3.86) implies that

$$\left\|d^{\mu}_{\theta} f(x)^{-1}\right\|^{-} \le m, \quad \forall x \in \bar{x} + \varepsilon B^{\circ}_X,\ \forall \theta > 0. \tag{3.87}$$

Fix arbitrary $x \in \bar{x} + \varepsilon B^{\circ}_X$.

Take $\kappa > m$. Let $v \in Y \setminus \{0\}$ be arbitrary. Since $\kappa > \operatorname{reg} A$, there is $u \in X$ such that $Au = v$ (thus $u \neq 0$) and

$$\|u\| < \kappa \|v\|. \tag{3.88}$$

For all $t \in (0, 1)$ such that $\|x + tu\| < \varepsilon$ from (3.86) it follows that

$$\|f(x + tu) - f(x) - A(tu)\| \le \mu \|tu\|.$$

Hence for $d := \operatorname{dist}_\theta((x + tu, f(x) + tv), \operatorname{gph} f)$ we have

$$\begin{aligned} d &\le \theta\|x + tu - (x + tu)\| + \|f(x) + tv - f(x + tu)\| \\ &= \|f(x) + tv - f(x + tu)\| \le \mu t \|u\|. \end{aligned}$$

By Definition 3.3.7 we have $v \in d^{\mu}_{\theta} f(x)u$ for all $\theta > 0$.

From this and (3.88) it follows that $\|d^{\mu}_{\theta} f(x)^{-1}\|^{-} \le \kappa$. Since this holds for all $\kappa > m$, (3.87) is verified.

Finally, we apply Theorem 3.3.8 with κ and $\theta = \dfrac{1 - m\mu}{m}$ to conclude. □

We can also derive Aubin derivative criterion for metric regularity, see Section 3.2, in the partial case of continuous function.

Theorem 3.3.10. Let X and Y be Banach spaces. Let $f : X \to Y$ be continuous on $\bar{x} + RB_X$ for some $R > 0$. Suppose that $m > 0$ is such that

$$\|df(x)^{-1}\|^{-} \le m, \quad \forall x \in \bar{x} + RB_X.$$

Then for every $\kappa > m$ and any $y \in Y$ such that $\|y - f(\bar{x})\| < R/\kappa$ there exists $x \in X$ such that $\|x - \bar{x}\| \le \kappa\|y - f(\bar{x})\|$ and $f(x) = y$.

Moreover, f is metrically regular at $\bar{x}$ with $\operatorname{reg}(f;\bar{x}) \le m$.

Proof. Fix $\kappa > m$.

We apply Theorem 3.3.8 with κ, $\varepsilon = R$, $\mu = 0$ and $\theta = 1/m$. □

By a similar way we can obtain a slight generalization of a result proved by Ekeland in [69]:

Theorem 3.3.11. Let X and Y be Banach spaces. Let $f : X \to Y$ be continuous and Gâteaux-differentiable with $f(0) = 0$. Assume that the derivative $Df(x)$ has a right-inverse $L(x)$, linear and uniformly bounded in a neighbourhood of 0. That is, there are $R, m > 0$ such that

$$\forall x \in RB_X,\ \forall v \in Y \Rightarrow Df(x)L(x)v = v \text{ and } \|L(x)\| \leq m.$$

Then for each $\bar{y}$ such that $\|\bar{y}\| < R/m$ and each $\nu > m$ there is some $\bar{x}$ such that

$$\|\bar{x}\| < R,\ \|\bar{x}\| \leq \nu\|\bar{y}\|,\ \text{and } f(\bar{x}) = \bar{y}.$$

Moreover, f is metrically regular at 0 with $\operatorname{reg}(f; 0) \leq m$.

Proof. Let $x \in RB_X$ and $v \in Y$. Since $Df(x) \in df(x)$, the contingent derivative $df(x)$ is surjective. If $u = L(x)v$, that is $Df(x)u = v$, then $v \in df(x)u$. Since

$$\|u\| \leq \|L(x)\|\|v\| \leq m\|v\|,$$

we have that

$$\|df(x)^{-1}\|^{-} \leq m. \tag{3.89}$$

Using (3.89), we apply Theorem 3.3.10 with R, m and $\nu > m$ to conclude. □

Let us note that the results from this section obtained for single-valued map can be easily extended to set-valued maps using an approach of A. Ioffe.

Also, Long Orbit or Empty Value principle could be applied for obtaining Nash-Mozer-Ekeland type surjectivity result in Fréchet spaces but this will be a subject of future research.

Bibliography

[1] E. Asplund, Čebyšev sets in Hilbert spaces, Trans. Amer. Math. Soc., **144**, 1969, 235–240.

[2] H. Attouch and R. J.-B. Wets, Isometries for the Legendre-Fenchel transform, Trans. Amer. Math. Soc., **296**, 1986, 33–60.

[3] J.-P. Aubin, Contingent derivatives of set-valued maps and existence of solutions to nonlinear inclusions and differential inclusions, Adv. Math., Suppl. Stud., **7A**, 1981, 159–229.

[4] J.-P. Aubin, Comportement lipschitzien des solutions de problémes de minimisation convexes, Compt. rend. Acad. Sci. Paris, Sér. I, **295**, 1982, 235–238.

[5] J.-P. Aubin, Lipschitz behavior of solutions to convex minimization problems, Math. Oper. Res., **9**, 1984, 87–111.

[6] J.-P. Aubin, ***Viability theory***, Birkhäuser, Berlin, 1991.

[7] J.-P. Aubin, A. Bayen, N. Bonneuil, and P. Saint-Pierre, ***Viability, control and game theories: Regulation of complex evolutionary systems under uncertainity***, Springer-Verlag, 2006.

[8] J.-P. Aubin and I. Ekeland, ***Applied nonlinear analysis***, John Wiley & Sons, New York, 1984 and Courier Corporation, 2006.

[9] J.-P. Aubin and H. Frankowska, On the inverse function theorem for set-valued maps, J. Math. Pures Appliquées, **66**, 1987, 71–89.

[10] J.-P. Aubin and H. Frankowska, ***Set-valued analysis***, Birkhäuser, New York, 1990 and Boston, MA: Birkhäuser, reprint of the 1990 original edition, 2009.

[11] D. Aussel, A. Daniilidis and L. Thibault, Subsmooth sets: Functional characterizations and related concepts, Trans. Amer. Math. Soc., **357**, 2004, 1275–1301.

[12] B. Beauzamy, ***Introduction to Banach spaces and their geometry***, 2nd edition, North-Holland, Amsterdam, 1985.

[13] G. Beer, ***Topologies on closed and closed convex sets***, Kluwer Academic Publishers, North-Holland, Dordrecht, 1993.

[14] F. Bernard and L. Thibault, Prox-regular functions and sets in Banach spaces, Set-Valued Analysis, **12**, 2004, 25–47.

[15] F. Bernard and L. Thibault, Prox-regular functions in Hilbert spaces, J. Math. Anal. Appl., **303**, 2005, 1–14.

[16] F. Bernard, L. Thibault and D. Zagrodny, Integration of primal lower nice functions in Hilbert spaces, J. Optimization Theory Appl., **124**, 2005, Issue 3, 561–579.

[17] F. Bernard and L. Thibault, Uniform prox-regularity of functions and epigraphs in Hilbert spaces, Nonlinear Anal. Theory Meth. Appl., **60**, 2005, 187–207.

[18] F. Bernard, L. Thibault and N. Zlateva, Characterizations of prox-regular sets in uniformly convex Banach spaces, J. Convex Anal., **13**, 2006, 525–559.

[19] F. Bernard, L. Thibault and N. Zlateva, Prox-regular sets and epigraphs in uniformly convex Banach spaces: Various regularities and other properties, Trans. Amer. Math. Soc., **363**, 2010, No 4, 2211–2247.

[20] J. Bonnans and A. Ioffe, Quadratic growth and stability in convex programming problems with multiple solutions, J. Convex Anal., **2**, 1995, 41–57.

[21] J. Bonnans and A. Shapiro, Optimization problems with perturbations: A guided tour, SIAM Rev., **40**, 1998, 228–264.

[22] J. Bonnans and A. Shapiro, ***Perturbation analysis of optimization problems***, Springer Series in Operations Research, Springer, New York, 2000.

[23] J. M. Borwein, Adjoint process duality, Math. Oper. Res., **8**, 1983, 403–434.

[24] J. M. Borwein, S. P. Fitzpatrick and J. R. Giles, The differentiability of real functions on normed linear space using generalized subgradients, J. Math. Anal. Appl., **128**, 1987, 512–534.

[25] J. M. Borwein and J. R. Giles, The proximal normal formula in Banach space, Trans. Amer. Math. Soc., **302**, 1987, 371–381.

[26] J. Borwein and A. Ioffe, Proximal analysis in smooth spaces, Set-Valued Anal., **4**, 1996, No 1, 1–24.

[27] J. Borwein and W. Moors, Essentially smooth Lipschitz functions, J. Funct. Anal., **149**, 1997, No. 2, 305–351.

[28] J. Borwein, W. Moors and X. Wang, Generalized subdifferentials: a Baire categorical approach, Trans. Amer. Math. Soc., **353**, 2001, No. 10, 3875–3893.

[29] J. Borwein, B. Mordukhovich and Y. Shao, On the equivalence of some basic principles in variational analysis, J. Math. Anal. Appl., **229**, 1999, No 1, 228–257.

[30] J. M. Borwein and H. M. Strójwas, Tangential approximations, Nonlinear Anal. Theory Meth. Appl., **9**, 1985, 1347–1366.

[31] J. M. Borwein and H. M. Strójwas, Proximal analysis and boundaries of closed sets in Banach space. I. Theory, Canad. J. Math., **38**, 1986, 431–452.

[32] J. M. Borwein and H. M. Strójwas, Proximal analysis and boundaries of closed sets in Banach space. II. Applications, Canad. J. Math., **39**, 1987, No 2, 428–472.

[33] M. Bounkhel and L. Thibault, On various notions of regularity of sets in nonsmooth analysis, Nonlinear Anal. Theory Meth. Appl., **48**, 2002, 223–246.

[34] H. Brezis, ***Analyse fonctionnelle***, Masson, Paris, 1983.

[35] A. Brøndsted, Conjugate convex functions in topological vector spaces, Mat.-Fys. Medd., Danske Vid. Selsk., **34**, 1964, No. 2, p. 27.

[36] A. Brøndsted and R. T. Rockafellar, On the subdifferentiability of convex functions, Proc. Amer. Math. Soc., **16**, 1965, 605–611.

[37] A. Canino, On p-convex sets and geodesics, J. Differential Equations, **75**, 1988, 118–157.

[38] P. Cardaliaguet, A differential game with two players and one target, SIAM J. Control and Optim., **34**, 1996, 1441–1460.

[39] P. Cardaliaguet, M. Quincampoix and P. Saint-Pierre, Numerical methods for differential games, In: Stochastic and differential games: Theory and numerical methods, Annals of the International Society of Dynamic Games (M. Bardi, T. E. S. Raghavan, and T. Parthasarathy eds.), Birkhäuser, 1999, 177–247.

[40] P. Cardaliaguet, M. Quincampoix and P. Saint-Pierre, Pursuit differential games with state constraints, SIAM J. Control and Optim., **39**, 2001, 1615–1632.

[41] P. Cardaliaguet, M. Quincampoix and P. Saint-Pierre, Differential games through viability theory: Old and recent results, Annals of the International Society of Dynamic Games, Birkhäuser, **9**, 2007, 2–23.

[42] J. Caristi, Fixed point theorems for mappings satisfying inwardness conditions, Trans. Amer. Math. Soc., **215**, 1976, 241–251.

[43] I. Cioranescu, ***Geometry of Banach spaces, duality mappings and nonlinear problems***, Mathematics and its Applications, **62**, Kluwer, Dordrecht, 1990.

[44] F. H. Clarke, ***Optimization and nonsmooth analysis***, John Wiley & Sons, New York, 1983.

[45] F. H. Clarke, Y. Ledyaev, J. R. Stern and R. P. Wolenski, ***Nonsmooth analysis and control theory***, Graduate Texts in Mathematics, Springer Verlag, New York, **178**, 1998.

[46] F. H. Clarke, J. R. Stern and R. P. Wolenski, Proximal smoothness and the lower-C^2 property, J. Convex Anal., **2**, 1995, 117–144.

[47] C. Combary, A. Elhilali, A. Levi, R. Poliquin and L. Thibault, Convex composite functions in Banach spaces and the primal lower-nice property, Proc. Amer. Math. Soc., **126**, 1998, No 12, 3701–3708.

[48] G. Colombo and V. V. Goncharov, Variational inequalities and regularity properties of sets in Hilbert spaces, J. Convex Anal., **8**, 2001, 197–221.

[49] R. Correa, P. Gajardo and L. Thibault, Subdifferential representation formula and subdifferential criteria for the behavior of nonsmooth functions, Nonlinear Anal. Theory Meth. Appl., **65**, 2006, Issue 4, 864—891.

[50] R. Correa and A. Jofre, Tangentially continuous directional derivatives in nonsmooth analysis, J. Optim. Theory Appl., **61**, 1989, 1–21.

[51] R. Correa, A. Jofre and L. Thibault, Characterization of lower semicontinuous convex functions, Proc. Amer. Math. Soc., **116**, 1992, 61–72.

[52] R. Correa, A. Jofre and L. Thibault, Subdifferential monotonicity as characterization of convex functions, Numer. Funct. Anal. Optim., **15**, 1994, No 5-6, 531–536.

[53] R. Correa, A. Jofre and L. Thibault, Subdifferential characterization of convexity, In: Recent advances in nonsmooth optimization (Eds. Ding-Zhu Du et al.), Singapore, World Scientific, 1995, 18–23.

[54] R. Correa and L. Thibault, Subdifferential analysis of bivariate separate regular functions, J. Math. Anal. Appl., **148**, 1990, 157–174.

[55] M. G. Crandal and P.-L. Lions, Viscosity solutions of Hamilton-Jacobi equations, Trans. Amer. Math. Soc., **227**, 1983, 1–24.

[56] M. Degiovanni, A. Marino and M. Toques, General properties of (p, q)-convex functions and operators, Ricerche Mat., **32**, 1983, 285–319.

[57] R. Deville, G. Godefroy and V. Zizler, ***Smoothness and renormings in Banach spaces***, Pitman Monographs and Surveys in Pure and Applied Mathematics, **64**, John Wiley & Sons Inc., New York, 1993.

[58] J. Diestel, ***Geometry of Banach spaces – selected topics***, Lecture Notes in Mathematics, **485**, Springer-Verlag, Berlin-New York, 1975.

[59] J. Diestel and J. J. Uhl, ***Vector measures***, Mathematical Surveys, **15**, AMS, Providence, RI, 1977.

[60] J. Diestel and J. J. Uhl, The Radon-Nikodym theorem for Banach space valued measures, Rocky Mt. J. Math., **6**, 1976, 1–46.

[61] A. L. Dontchev and A. S. Lewis, Perturbations and metric regularity, Set-Valued Anal., **13**, 2005, Issue 4, 417–438.

[62] A. L. Dontchev, A. S. Lewis and R. T. Rockafellar, The radius of metric regularity, Trans. Amer. Math. Soc., **355**, 2002, 493–517.

[63] A. L. Dontchev, M. Quincampoix and N. Zlateva, Aubin criterion for metric regularity, Journal of Convex Anal., **13**, 2006, No 2, 281–297.

[64] A. L. Dontchev and R. T. Rockafellar, Characterization of strong regularity of variational inequalities over polyhedral convex sets, SIAM J. Optim., **6**, 1996, 1087–1105.

[65] A. L. Dontchev and R. T. Rockafellar, Ample parameterization of variational inclusions, SIAM J. Optim., **12**, 2001, 170–187.

[66] A. L. Dontchev and R. T. Rockafellar, Regularity and conditioning of solution mappings in variational analysis, Set-Valued Anal., **12**, 2004, 79–109.

[67] A. L. Dontchev and R. T. Rockafellar, ***Implicit functions and solution mappings***, Springer, 2009.

[68] J. F. Edmond and L. Thibault, BV solutions of nonconvex sweeping process differential inclusion with perturbation, J. Differential Equations, **226**, 2006, Issue 1, 135–179.

[69] I. Ekeland, An inverse function theorem in Fréchet spaces, Ann. Inst. H. Poincaré Anal. Non Linéaire, **28**, 2011, No 1, 91–105.

[70] I. Ekeland and G. Lebourg, Generic Fréchet differentiability and perturbed optimization problems in Banach spaces, Trans. Amer. Math. Soc., **224**, 1976, 193–216.

[71] M. Fabian, P. Habala, P. Hájek, V. Montesinos Santalucía, J. Pelant and V. Zizler, ***Functional analysis and infinite-dimensional geometry***, CMS Books in Mathematics, Springer-Verlag, New York, 2001.

[72] M. Fabian and D. Priess, A generalization of the interior mapping theorem of Clarke and Pourciau, Comment. Math. Univ. Carolinae, **28**, 1987, 311–324.

[73] H. Federer, Curvature measures, Trans. Amer. Math. Soc., **93**, 1959, 418–491.

[74] W. Fenchel, On conjugate convex functions, Canad. J. Math., **1**, 1949, 73–77.

[75] S. Fitzpatrick, Differentiation of real-valued functions and continuity of metric projections, Proc. Amer. Math. Soc., **91**, 1984, 544–548.

[76] S. Fitzpatrick and R. R. Phelps, Bounded approximants to monotone operators on Banach spaces, Ann. Inst. Henri Poincaré, Anal. Non Linéaire, **9**, 1992, No. 5, 573–595.

[77] H. Frankowska, Some inverse mapping theorems, Ann. Inst. Henri Poincaré, Anal. Non Linéaire, **7**, 1990, 183–234.

[78] P. G. Georgiev and N. P. Zlateva, Second-order subdifferentials of $C^{1,1}$ functions and optimality conditions, Set-Valued Anal., **4**, 1996, 101–117.

[79] J.-P. Gossez, On the subdifferential of a saddle function, J. Funct. Anal., **11**, 1972, 220–230.

[80] E. El Haddad and R. Deville, The viscosity subdifferential of the sum of two functions in Banach space. I: First order case, J. Convex Anal., **3**, 1996, No 2, 295–308.

[81] J. Hartung, An extension of Sion's minimax theorem with an application to a method for constrained games, Pacific J. Math., **103**, 1982, 401–408.

[82] F. Hausdorff, Über halbstetige Funcktionen und deren Verallgemeinerung, Math. Zeit., **5**, 1919, 292–309.

[83] R. Haydon, ***Trees and renormings***, Publ. Math. Univ. Pierre Marie Curie 104, Semin. Initiation Anal., 30me Annee: 1990/91, No 8, 1991.

[84] J.-B. Hiriart-Urruty, Lipschitz r-continuity of the approximate subdifferential of a convex function, Math. Scand., **47**, 1980, 123–134.

[85] J.-B. Hiriart-Urruty, J.-J. Strodiot and V. H. Nguyen, Generalized Hessian matrix and second-order optimality conditions for problems with $C^{1,1}$ data, Appl. Math. Optim., **11**, 1984, 43–56.

[86] R. B. Holmes, ***Geometric functional analysis and its applications***, Graduate Texts in Mathematics, **24**, New York-Heidelberg-Berlin, Springer-Verlag, 1975.

[87] A. Ioffe, Approximate subdifferentials and applications 2, Mathematica, **33**, 1986, 111–128.

[88] A. Ioffe, Approximate subdifferentials and applications 3, Mathematica, **36**, 1989, 1–36.

[89] A. Ioffe, Proximal analysis and approximate subdifferentials, J. Lond. Math. Soc., **II 41**, 1990, No 1, 175–192.

[90] A. D. Ioffe, Metric regularity and subdifferential calculus, Uspekhi Mat. Nauk, **55**, 2000, No 3(333), 103–162; English translation Math. Surveys, **55**, 2000, 501–558.

[91] M. Ivanov, Sequential representation formulae for G-subdifferential and Clarke subdifferential in smooth Banach spaces, Journal of Convex Anal., **11**, 2004, No 1, 179–196.

[92] M. Ivanov, New proof of Ekeland perturbed minimisation method. Compt. rend. Acad. bulg. Sci., **68**, 2015, No 11, 1353–1356.

[93] M. Ivanov and N. Zlateva, Abstract subdifferential calculus and semi-convex functions, Serdica Math. J., **23**, 1997, 35–58.

[94] M. Ivanov and N. Zlateva, On primal lower-nice property, Compt. rend. Acad. bulg. Sci, **54**, 2001, No 11, 5–10.

[95] M. Ivanov and N. Zlateva, Subdifferential characterization of primal lower-nice functions on smooth Banach spaces, Compt. rend. Acad. bulg. Sci, **57**, 2004, No 5, 13–18.

[96] M. Ivanov and N. Zlateva, Long orbit or empty value principle, fixed point and surjectivity theorems, Compt. rend. Acad. bulg. Sci, **69**, 2016, No 5, 553–562.

[97] A. Jofre and L. Thibault, D-representation of subdifferentials of directionally Lipschitz functions, Proc. Amer. Math. Soc., **110**, 1990, No. 1, 117–123.

[98] A. Jourani and L. Thibault, Metric regularity for strongly compactly Lipschitzian mappings, Nonlinear Anal. Theory Meth. Appl., **24**, 1995, 229–240.

[99] A. Jourani and L. Thibault, Coderivatives of multivalued mappings, locally compact cones and metric regularity, Nonlinear Anal. Theory Meth. Appl., **35**, 1999, 925–945.

[100] W. A. Kirk, A Fixed point theorem for mappings which do not increase distances, Amer. Math. Monthly, **72**, 1965, No 9, 1004–1006.

[101] D. Klatte and R. Henrion, Regularity and stability in nonlinear semi-infinite optimization, In: Semi-infinite Programming (Reemtsen et al. eds.), Workshop, Cottbus, Germany, September 1996, Kluwer Academic Publishers, Boston; Nonconvex Optim. Appl., **25**, 1998, 69–102.

[102] D. Klatte and B. Kummer, ***Nonsmooth equations in optimization. Regularity, calculus, methods and applications***, Kluwer Academic Publishers, Dordrecht, 2002.

[103] V. Klee, Convexity of Chebyshev sets, Math. Annalen, **142**, 1961, 292–304.

[104] N. N. Krasovskii and A. I. Subbotin, ***Game-theorical control problems***, Springer-Verlag, New York, 1988.

[105] E. Krauss, A representation of maximal monotone operators by saddle functions, Rev. Roumaine Math. Pures Appl., **30**, 1985, 823–837.

[106] A. Kruger, Properties of generalized differentials, Siberian Math. Journal, 1983, 822–832.

[107] A. Kruger and B. Mordukhovich, Extreme points and the Euler equation in nondifferentiable optimization problems, Dokl. Acad. Nauk BSSR, **24**, 1982, 684–687.

[108] K. S. Lau, Almost Chebyshev subsets in reflexive Banach spaces, Indiana Univ. Math. J., **27**, 1978, No 5, 791–795.

[109] P.-J. Laurent, Approximation et optimisation, Enseignement des sciences, **13**, Paris, Hermann, 1972.

[110] Y. S. Ledyaev and Q. J. Zhu, Implicit multifunction theorems, Set-Valued Anal., **7**, 1999, 209–238.

[111] A. Levi, R. Poliquin and L. Thibault, Partial extensions of Attouch's theorem with applications to proto-derivatives of subgradient mappings, Trans. Amer. Math. Soc., **347**, 1995, 1269–1294.

[112] J. Lindenstrauss and L. Tzafriri, ***Classical Banach spaces. I. Sequence spaces***, Springer-Verlag, Berlin-New York, 1977.

[113] J. Lindenstrauss and L. Tzafriri, ***Classical Banach spaces. II. Function spaces***, Springer-Verlag, Berlin-New York, 1979.

[114] P. D. Loewen, The proximal subgradient formula in Banach space, Canad. Math. Bull., **31**, 1988, 353–361.

[115] S. Marcellin and L. Thibault, Evolution problems associated with primal lower nice functions, J. Convex Anal., **13**, 2006, No 2, 385–421.

[116] B. Maury and J. Venel, Un modèle de mouvements de foule, ESAIM Proceedings, **18**, 2007, 143–152.

[117] L. McLinden, Dual operations on saddle functions, Trans. Amer. Math. Soc., **179**, 1973, 363–381.

[118] L. McLinden, An extension of Fenchel's duality theorem to saddle functions and dual minimax problems, Pacific J. Math., **50**, 1974, 135–158.

[119] L. McLinden, Conjugacy correspondences: a unified view, Trans. Amer. Math. Soc., **203**, 1975, 257–274.

[120] Ph. Michel and J.-P. Penôt, A generalized derivative for calm and stable functions, Differental Integral Equations, **5**, 1992, 433–454.

[121] B. Mordukhovich, Metric approximations and necessary optimality conditions for general classes of nonsmooth extremal problems, Soviet Math. Dokl., **22**, 1980, 526–530.

[122] B. Mordukhovich, Complete characterization of openness, metric regularity, and Lipschitzian properties of multifunctions, Trans. Amer. Math. Soc., **340**, 1993, 1–35.

[123] B. Mordukhovich, Coderivatives of set-valued mappings: calculus and applications, Nonlinear Anal. Theory Meth. Appl., **30**, 1997, 3059–3070.

[124] B. S. Mordukhovich, ***Variational analysis and generalized differentiation I and II***, Springer, New York, Comprehensive Studies in Mathematics, **330** and **331**, 2005.

[125] B. Mordukhovich and Y. Shao, Differential characterizations of covering, metric regularity, and Lipschitzian properties of multifunctions between Banach spaces, Nonlinear Anal. Theory Meth. Appl., **25**, 1995, 1401–1424.

[126] B. S. Mordukhovich and Y. Shao, Nonsmooth sequential analysis in Asplund spaces, Trans. Amer. Math. Soc., **348**, 1996, 1235–1280.

[127] J.-J. Moreau, Sur la fonction polaire d'une fonction semicontinue supérieurement, Compt. rend. Acad. Sci. Paris, **258**, 1964, 1128–1130.

[128] J.-J. Moreau, Proximité et dualité dans un espace hilbertien, Bull. Soc. Math. France, **93**, 1965, 273–299.

[129] J.-J. Moreau, ***Fonctionnelles convexes***, Séminaire sur les équations aux dérivées partielles, Collège de France (1966–1967).

[130] T. S. Motzkin, Sur quelques propriétés caractéristiques des ensembles convexes, Att. R. Acad. Lincei, Rend. **21**, 1935, 562–567.

[131] O. I. Pak, On properties of the subdifferential of a convex-concave function, Kibernetika, **3**, 1982, 127–129.

[132] J.-P. Penôt, ***Calculus without derivatives***, Springer, 2013.

[133] R. R. Phelps, ***Convex functions, monotone operators and differentiability***, 2nd ed., Lecture Notes in Mathematics, **1364**, Springer-Verlag, Berlin, 1993.

[134] S. Plaskacz and M. Quincampoix, Value-functions for differential games and control systems with discontinuous terminal cost, SIAM J. Control and Optim., **39**, 2001, 1485–1498.

[135] R. Poliquin, Subgradient monotonicity and convex functions, Nonlinear Anal. Theory Meth. Appl., **17**, 1990, 305–317.

[136] R. Poliquin, Integration of subdifferentials of nonconvex functions, Nonlinear Anal. Theory Meth. Appl., **17**, 1991, 358–398.

[137] R. A. Poliquin and R. T. Rockafellar, Prox-regular functions in variational analysis, Trans. Amer. Math. Soc., **348**, 1996, 1805–1838.

[138] R. A. Poliquin, R. T. Rockafellar and L. Thibault, Local differentiability of distance functions, Trans. Amer. Math. Soc., **352**, 2000, 5231–5249.

[139] D. Preiss, Differentiability of Lipschitz functions on Banach spaces, J. Funct. Anal., **91**, 1990, 312–345.

[140] M. Quincampoix and N. Zlateva, Parameterized minimax problem: on Lipschitz-like dependence of the solution with respect to the parameter, SIAM J. Optim., **19**, 2008, No 3, 1250–1269.

[141] S. M. Robinson, Normed convex processes, Trans. Amer. Math. Soc., **174**, 1972, 127–140.

[142] R. T. Rockafellar, Minimax theorems and conjugate saddle functions, Math. Scand., **14**, 1964, 151–173.

[143] R. T. Rockafellar, Extension of Fenchel's duality theorem for convex functions, Duke Math. J., **33**, 1966, 81–90.

[144] R. T. Rockafellar, Characterization of the subdifferentials of convex functions, Pacific J. Math., **17**, 1966, 497–510.

[145] R. T. Rockafellar, A general correspondence between dual minimax problems and convex programs, Pacific J. Math., **25**, 1968, No. 3, 597–611.

[146] R. T. Rockafellar, Monotone operators associated with saddle functions and minimax problems, In: Proc. Symp. Pure Math., Amer. Math. Soc., **18**, 1970, 241–250.

[147] R. T. Rockafellar, On the maximal monotonicity of subdifferential mappings, Pacific J. Math., **33**, 1970, No. 1, 209–216.

[148] R. T. Rockafellar, ***Convex analysis***, Prinston, New Jersey, Prinston University Press, 1970.

[149] R. T. Rockafellar, Saddle-points and convex analysis, In: Differential games related topics, Proc. Intern. Summer School, Varenna 1970, Amsterdam, North-Holland, 1971, 109–127.

[150] R. T. Rockafellar, Generalized directional derivatives and subgradients of non-convex functions, Canad. J. Math., **32**, 1980, No. 2, 257–280.

[151] R. T. Rockafellar, Proximal subgradients, marginal values, and augmented Lagrangians in nonconvex optimization, Math. Oper. Res., **6**, 1981, No 3, 424–436.

[152] R. T. Rockafellar and R. J.-B. Wets, ***Variational analysis***, Springer-Verlag, Berlin, Comprehensive Studies in Mathematics, **317**, 1998.

[153] A. S. Shapiro, Existence and differentiability of metric projections in Hilbert spaces, SIAM J. Optimization, **4**, 1994, 130–141.

[154] A. Shapiro, On Lipschitzian stability of optimal solutions of parametrized semi-infinite programs, Math. Oper. Res., **19**, 1994, 743–752.

[155] A. Shapiro, Sensitivity analysis of parameterized variational inequalities, Math. Oper. Res., **30**, 2005, 109–126.

[156] S. Simons, ***Minimax monotonicity***, Lecture Notes in Mathematics, **1693**, Berlin, Springer, 1998.

[157] C. Stegall, Optimization of functions on certain subsets of Banach spaces, Math. Ann., **236**, 1978, 171–176.

[158] L. Thibault, Sweeping process with regular and non regular sets, J. Differential Equations, **193**, 2003, 1–26.

[159] L. Thibault, Regularization of nonconvex sweeping process in Hilbert space, Set-Valued Anal., **16**, 2008, Issue 2, 319–333.

[160] L. Thibault and D. Zagrodny, Integration of subdifferentials of lower semicontinuous functions on Banach spaces, J. Math. Anal. Appl., **189**, 1995, 33–58.

[161] L. Thibault and D. Zagrodny, Enlarged inclusion of subdifferentials, Canad. Math. Bull., **48**, 2005, No 2, 283–301.

[162] L. Thibault and N. Zlateva, Integrability of subdifferentials of certain bivariate functions, Nonlinear Anal. Theory Meth. Appl., **54**, 2003, 1251–1269.

[163] L. Thibault and N. Zlateva, Integrability of subdifferentials of directionally Lipschitz functions, Proc. Amer. Math. Soc., **133**, 2005, No 10, 2939–2948.

[164] L. Thibault and N. Zlateva, Partially ball weakly inf-compact saddle functions, Math. Oper. Res., **30**, 2005, No 2, 404–419.

[165] J. Treiman, Generalized gradients, Lipschitz behaviour and directional derivatives, Canad. J. Math., **37**, 1985, No. 6, 1074–1084.

[166] L. P. Vlasov, On Chebyshev sets, Soviet Math. Dokl., **8**, 1967, 401–404.

[167] L. P. Vlasov, Almost convex and Chebyshev sets, Math. Notes Acad. Sci. USSR, **8**, 1970, 776–779.

[168] Z. Wu and J. Ye, Some results on integration of subdifferentials, Nonlinear Anal. Theory Meth. Appl., **39**, 2000, 955–976.

[169] Z. B. Xu and G. F. Roach, Characteristic inequalities of uniformly convex and uniformly smooth Banach spaces, J. Math. Anal. Appl., **157**, 1991, 189–210.

[170] D. Zagrodny, Approximate mean value theorem for upper subderivatives, Nonlinear Anal. Theory Meth. Appl., **12**, 1988, 1413–1428.

[171] E. H. Zarantonello, Projections on convex sets in Hilbert space and spectral theory I and II. In: Contributions to Nonlinear Functional Analysis (Ed. E. H. Zarantonello), Academic Press, New York, 1971, 237–424.

[172] N. Zlateva, Integrability through infimal regularization, Compt. rend. Acad. bulg. Sci., **68**, 2015, No 5, 551–560.