SOFIA UNIVERSITY ”ST. KLIMENT OHRIDSKI”
FACULTY OF MATHEMATICS AND INFORMATICS

A CONTRIBUTION TO SUBDIFFERENTIAL CALCULUS

Mihail Atanasov Hamamdjiev

DISSERTATION FOR PHD DEGREE IN MATHEMATICS

S O F I A
2 0 1 8

Acknowledgements:

I would like to dedicate this work to the memory of my first teacher in mathematics, my father, Atanas Hamamdjiev and to the memory of my first university teacher in mathematics, Vasil Tsanov. Rest in peace! I want to thank my scientific advisor, Milen Ivanov, for his patience, support and dedication. Also, many thanks to all my co-workers and last, but not least, my family.

Contents

Preface **2**

1 Preliminaries **3**

2 Moreau-Yosida regularization **8**
2.1 Renormings in separable Banach spaces – rotund norms and connection to Gâteaux differentiabilty 8
2.2 Continuity property of the norm 11
2.3 Additional properties of the Moreau-Yosida regularization 17

3 Multidirectional Inequalities **21**
3.1 Motivation and former results 21
3.2 Axiomatic definition of an abstract subdifferential 22
3.3 Statement of the new multidirectional inequality and comparison with other results 25
3.4 Proof of the new multidirectional inequality 27

4 Conclusion **37**

Bibliography **38**

Preface

The modern calculus is developed in the 17th century by Sir Isaac Newton and Gottfried Leibniz. An important concept in these early days was the property of differentiablity of the functions considered. Newton formulated the laws of motion and universal gravitation and considered Mechanics in ideal conditions, like celestial bodies in motion. Naturally, in this context, all the functions are smooth and calculus has proved to be a superb tool.

On the other hand, by considering other fields of Mechanics, Moreau and many others faced the need of systematic study of non-smooth phenomena. In settings in which smoothness is lacking, one usually works with the concept of subgradient. The basic idea is simple: If we imagine the graph of the function $t \to |t|$ as approximated by smooth functions around zero, it might appear as if it has many tangents at 0. This work is focused on the subdifferential calculus.

The current dissertation is splitted into several chapters: Preface, Chapter 1 - Preliminaries, Chapter 2 - Moreau-Yosida regularization, Chapter 3 - Multidirectional Inequalities, Chapter 4 - Conclusion and Bibliography.

In Chapter 1 - Preliminaries are given some basic definitions and well known fundamental results, which are used later on.

In Chapter 2 - Moreau-Yosida regularization are discussed some nice properties that the regularization has when the underlying space is separable.

Chapter 3 - Multidirectional Inequalities is dedicated to generalizations of the classical Mean Value theorem.

Chapter 1

Preliminaries

The aim of the chapter is to give definitions of the objects we will deal with later on and some conventions. We will deal only with Banach spaces, in some cases separable. First we will ratify some basic notations. The underlying Banach space will be denoted by X with associated norm $\|\cdot\|$. The closed unit ball and the unit sphere, associated with this norm, will be denoted by

$$B_X := \{x \in X : \|x\| \leq 1\}$$

and

$$S_X := \{x \in X : \|x\| = 1\},$$

respectively. The topological dual space of X will be denoted X^* which is the space of all continuous linear functionals on X with a dual norm $\|\cdot\|_*$ defined as

$$\|p\|_* = \sup_{x \in B_X} |\langle p, x\rangle| .$$

We will often use again $\|\cdot\|$ instead, when the meaning is clear. The natural pairing $\langle p, x\rangle$ for $x \in X$ and $p \in X^*$ will often be denoted by $p(x)$.
Furthermore, let X be a vector space equipped with two norms $\|\cdot\|_1$ and $\|\cdot\|_2$. Then these norms are said to be equivalent, if there exist positive constants $c, C > 0$ such that

$$c\,\|x\|_2 \leq \|x\|_1 \leq C\,\|x\|_2$$

for every $x \in X$, see e.g. [2, p. 15].

Definition 1.0.1. A norm $\|\cdot\|$ on a Banach space X is rotund (or strictly convex), denoted (R), if for $x, y \in S_X$ it follows that $\|x + y\| = 2$ if and only if $x = y$, see e.g. [1, p. 42].

It is well-known that the above definition is equivalent to the following

Definition 1.0.2. A norm $\|\cdot\|$ on a Banach space X is rotund (or strictly convex), denoted (R), if for $x, y \in S_X$, $x \neq y$ we have that $\|tx + (1-t)y\| < 1$ for all $t \in (0,1)$.

Geometrically, this means that the unit sphere in X does not contain any line segments. As it is well-known, the space l_∞ has no equivalent rotund norm, see e.g. [1, p. 76]. Thus the existence of equivalent (R) norms is not guaranteed by the intrinsic properties of a general Banach space.

The next definition is classical, see e.g. [1, p. 2].

Definition 1.0.3. Let $(X, \|\cdot\|)$ be a Banach space and $f : X \to \mathbb{R}$ be a function.

(i) We say that f is Gâteaux differentiable at $x \in X$ if for every $h \in X$ the following limit

$$f'(x)(h) := \lim_{t \to 0} \frac{f(x+th) - f(x)}{t}$$

exists and is a linear continuous function in h, that is $f'(x) \in X^*$. Then the functional $f'(x)$ is called Gâteaux derivative or Gâteaux differential of f at x.

(ii) We say that $\|\cdot\|$ is Gâteaux smooth or a G-norm if $\|\cdot\|$ is Gâteaux differentiable at all $x \in S_X$ (by homogeneity this implies that $\|\cdot\|$ is Gâteaux differentiable for every $x \in X \setminus \{0\}$).

In addition, we will mention that if $\|\cdot\|$ is Gâteaux differentiable at $x \in S_X$ and f_x denotes its derivative, then $\|f_x\| = f_x(x) = 1$. This is true as the norm is 1-Lipschitz function, so $\|f_x\| \leq 1$. Furthermore,

$$1 - f_x(x) = \lim_{t \to 0} \frac{\|x + tx\| - \|x\| - f_x(tx)}{t} = 0,$$

that is, $\|f_x\| = 1$.

Next we give the standard in Convex analysis definition of a subdifferential, see e.g. [6, p. 6].

Definition 1.0.4. Let $f : X \to \mathbb{R}$ be a convex function. Then we define the subdifferential operator ∂^- of f at the point x to be a multi-valued map $\partial^- f : X \to 2^{X^*}$ defined as

$$\partial^- f(x) := \{p \in X^* : \langle p, y - x \rangle \leq f(y) - f(x) : \ \forall y \in X\}.$$

Note that this is the same as saying that the affine function $p(y) + \alpha$ where $\alpha = f(x) - \langle p, x \rangle$ is dominated by f and is equal to it at $y = x$. In a similar way one defines also the notions of superdifferential, ε-subdifferential and ε-superdifferential, respectively:

$$\partial^+ f(x) := \{p \in X^* : \langle p, y - x \rangle \geq f(y) - f(x) : \ \forall y \in X\},$$

$$\partial_\varepsilon^- f(x) := \{p \in X^* : \langle p, y - x\rangle \leq f(y) - f(x) + \varepsilon : \ \forall y \in X\},$$
$$\partial_\varepsilon^+ f(x) := \{p \in X^* : \langle p, y - x\rangle \geq f(y) - f(x) - \varepsilon : \ \forall y \in X\}.$$

In spite of defining these sets with upper ± indexes for consistency, we will denote from now on the subdifferentials with the $\partial f(x)$ and $\partial_\varepsilon f(x)$ signs for simplicity as they are more popular in the literature.

As the subdifferential is a multivalued map, $\partial f(x)$ consists of a (convex) subset of X^*. When the set $\partial f(x)$ is a sigleton, it coincides with the derivative of f at x, usually denoted $f'(x)$. The notion of the norm of this derivative $\|f'(x)\|$ is clear as it is element of X^*. In the general case $\partial f(x)$ contains more elements of X^*, thus we need to define what a "norm" of a subdifferential means here.

Definition 1.0.5. A "norm" of a subdifferential ∂ of a convex function f at the point x is a real multivalued map $\|\cdot\| : X \to \mathbb{R}$ defined as follows:

$$\|\partial f(x)\| := \{\|p\| : p \in \partial f(x)\}$$

Note that we don't pretend for accuracy of the name of the latter object – if similar objects are named else in the literature – as there is no further observations outside the main result of Section 2.2 on the properties of this so called "norm" in this dissertation. At first glimpse, however, it looks like a nice generalization of the notion norm for multivalued maps.

Next, we define the notions of infimal/supremal convolution.

Definition 1.0.6. Let X be a topological space and φ, ψ be two real valued proper functions (i.e. their domains are non-empty) defined on X. Then we define their infimal and supremal convolutions (or, inf-convolution and sup-convolution), respectively, as follows:

$$f \square g(x) := \inf\{f(x - y) + g(y) : y \in X\},$$
$$f \triangle g(x) := \sup\{f(x - y) + g(y) : y \in X\}.$$

Here we will show that the inf-convolution of two convex functions is again a convex function (which is well-known over $\mathbb{R}^n$, but we didn't find any explicit references for a Banach space setting). In fact, we will prove something, which is more general and well-known.

Proposition 1.0.7. Let X, Y be Banach spaces and the function $F : X \times Y \to \mathbb{R} \cup \{+\infty\}$ be proper and convex. Then the marginal function $f : X \to \mathbb{R} \cup \{+\infty\}$ defined by

$$f(x) = \inf_{y \in Y}\{F(x, y)\}$$

is convex.

Proof. We consider any $x_1, x_2 \in \mathrm{dom} f$ (as for $f(x_i) = +\infty$ the assertion is trivial) such that $x_1 \neq x_2$ and $\lambda \in (0,1)$. As by definition $f(x_i) = \inf_{y\in Y}\{F(x_i, y)\} < +\infty$, one can always find numbers $\xi_i \in (f(x_i), +\infty)$ and $y_i \in Y$ such that $F(x_i, y_i) < \xi_i$, $i = 1, 2$. By convexity of F one can observe that

$$\begin{aligned} f(\lambda x_1 + (1-\lambda)x_2) &= \inf_{y\in Y}\{F(\lambda x_1 + (1-\lambda)x_2, y)\} \\ &\leq F(\lambda x_1 + (1-\lambda)x_2, \lambda y_1 + (1-\lambda)y_2) \\ &\leq \lambda F(x_1, y_1) + (1-\lambda)F(x_2, y_2) \\ &< \lambda\xi_1 + (1-\lambda)\xi_2. \end{aligned}$$

Now we can take the inf of the right hand side when $\xi_i \downarrow f(x_i)$ to obtain the needed inequality $f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2)$. □

Corollary 1.0.8. Let X be a Banach space and $f, g : X \to \mathbb{R} \cup \{+\infty\}$ be proper and convex functions. Then $f \square g$ is a convex function.

Proof. We apply the above proposition to the function $F(x, y) = f(y - x) + g(y)$ which is convex as a sum of two convex functions. Now observe that

$$\inf_{y\in Y} F(x, y) \equiv f\square g(x) = \inf_{y\in Y}\{f(y - x) + g(y)\}.$$

□

Definition 1.0.9. Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be a proper function and let $g_\lambda(x) := \frac{1}{2\lambda}\|x\|^2$ for any $\lambda > 0$. The Moreau-Yosida regualrization f_λ of f is defined via the inf-convolution:

$$f_\lambda(x) := f\square g_\lambda(x).$$

Note that the last one implies

$$f_\lambda(x) = \inf_{y\in X}\{f(x - y) + \frac{1}{2\lambda}\|y\|^2)\} = \inf_{y\in X}\{f(y) + \frac{1}{2\lambda}\|x - y\|^2\}.$$

We will also need a classical result in Functional analysis – the prominent Ekeland variational principle (EVP), which is known to be equivalent to completeness of metric spaces. For reference, see e.g. [12, p. 62].

Definition 1.0.10. Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be a proper function. We say that $\bar{x} \in X$ is an ε-minimizer of f if $f(\bar{x}) \leq \inf f(X) + \varepsilon$.

Theorem 1.0.11 (Ekeland variational principle). Let (X, d) be a complete metric space and let $f : X \to \mathbb{R} \cup \{+\infty\}$ be a bounded below lower semicontinuous proper function. Given an ε-minimizer $\bar{x}$ of f and given $\gamma, \rho > 0$ satisfying $\gamma\rho \geq \varepsilon$, one can find $u \in X$ such that the following inequalities hold:

(i) $d(u, \bar{x}) \leq \rho$,

(ii) $f(u) + \gamma d(u, \bar{x}) \leq f(\bar{x})$,

(iii) $f(u) < f(x) + \gamma d(u, x)$ for all $x \in X \setminus \{u\}$.

Next, we will need the notion of weak-* topology. First we consider the linear functionals $\varphi_x : X^* \to \mathbb{R}$ defined by $p \to \varphi_x(p) = \langle p, x\rangle$. As x runs through X we obtain a collection $(\varphi_x)_{x\in X}$ of maps from X^* to $\mathbb{R}$.

Definition 1.0.12. The weak-* topology on X^* is the coarsest topology on X^* associated to the collection $(\varphi_x)_{x\in X}$. We denote the convergence in this topology by $p_n \to^{w^*} p$.

This association is done in the following way (see e.g. [14, Chapter 3]): For all open sets ω in $\mathbb{R}$ the preimages $\varphi_p^{-1}(\omega)$ are necessarily open in X^*. Thus we obtain a family of subsets of X^* for which we first take finite intersections and for the new family we then take arbitrary unions. The resulting family of subsets in X^* is closed under finite intersections and arbitrary unions, thus they form the coarsest topology which is called the weak-* topology on X^*. A fundamental characteristic of the w^*-convergence is given by the following proposition.

Proposition 1.0.13. For $p_n, p \in X^*$ we have that $p_n \to^{w^*} p$ if and only if

$$\langle p_n, x\rangle \to \langle p, x\rangle \ \ \forall x \in X.$$

We will also present the fundamental notions of topological boundary and interior (see e.g. [15, 44-45]).

Definition 1.0.14. A point x of a subset A of a topological space X is an interior point of A if and only if A is a neighbourhood of x. The set of all interior points of A is the topological interior of A, denoted by $\mathrm{Int}(A)$.

Definition 1.0.15. The topological boundary of a subset A of a topological space X is the set of all points x, which are interior to neither A, nor $X \setminus A$ and it is denoted $\mathrm{Bnd}(A)$. Equivalently, x is a point of the boundary if and only if each neighbourhood of x intersects both A and $X \setminus A$. It is clear that the boundary of A is identical with the boundary of $X \setminus A$.

Chapter 2

Moreau-Yosida regularization

2.1 Renormings in separable Banach spaces – rotund norms and connection to Gâteaux differentiabilty

The aim of this section is to recall the construction of an equivalent norm on a separable Banach space whose dual norm is rotund. The existence of such for separable X is well-known. Furthermore, we will connect the rotundness of a norm with the Gâteaux differentiabilty.

First we will give a characterization of the (R) norms.

Lemma 2.1.1. Let $(X, \|\cdot\|)$ be a Banach space. Then $\|\cdot\|$ is (R) if and only if $\|\cdot\|^2$ is a strictly convex function.

Proof. First we note that for non-decreasing and convex function $f : \mathbb{R} \to \mathbb{R}$ and a convex function $g : X \to \mathbb{R}$ the composition $f \circ g : X \to \mathbb{R}$ is a convex function. Indeed, for every $x, y \in X$ and $t \in [0, 1]$ we have

$$\begin{aligned} f \circ g(tx + (1-t)y) &\leq f(tg(x) + (1-t)g(y)) \\ &\leq tf \circ g(x) + (1-t)f \circ g(y). \end{aligned}$$

Then, as taking second power of a non-negative number is convex and non-decreasing function and the norm is a convex function, we have that $\|\cdot\|^2 : X \to \mathbb{R}$ is a convex function.

We assume that $\|\cdot\|^2$ is strictly convex function, that is for every $x, y \in X$, $x \neq y$ and $t \in (0, 1)$ we have

$$\|tx + (1-t)y\|^2 < t\|x\|^2 + (1-t)\|y\|^2 .$$

In particular, for $x, y \in S_X$ we obtain $\|tx + (1-t)y\| < 1$ and this implies that $\|\cdot\|$ is (R).

For the other direction we assume the contrary. Let $\|\cdot\|$ be (R) and assume that $\|\cdot\|^2$ is not strictly convex function. Then we can find $x \neq y$ and $t \in (0,1)$ such that

$$t\|x\|^2 + (1-t)\|y\|^2 = \|tx + (1-t)y\|^2 . \tag{2.1}$$

By the triangle inequality

$$\begin{aligned} t\|x\|^2 + (1-t)\|y\|^2 &= \|tx + (1-t)y\|^2 \\ &\leq [t\|x\| + (1-t)\|y\|]^2 \\ &= t^2\|x\|^2 + 2t(1-t)\|x\|\,\|y\| + (1-t)^2\|y\|^2 . \end{aligned}$$

Then, after rearranging, we obtain

$$(t-t^2)\|x\|^2 - 2t(1-t)\|x\|\,\|y\| + (1-t)(1-(1-t))\|y\|^2 \leq 0$$

which is equivalent to

$$t(1-t)[\|x\| - \|y\|]^2 \leq 0.$$

This implies that $\|x\| = \|y\|$ as $t \in (0,1)$ and without loss of generality we can take $x, y \in S_X$, which means exactly that

$$t\|x\|^2 + (1-t)\|y\|^2 = 1.$$

Now by (2.1) we obtain

$$\|tx + (1-t)y\| = 1.$$

The last one is a contradiction with the (R) property of $\|\cdot\|$ as $x \neq y$ by assumption. □

Let $(X, \|\cdot\|)$ be a separable Banach space. As X is separable, we can fix a sequence $(x_n) \subseteq S_X$, $n = 1, 2, \ldots$ which is dense in S_X. Then we consider the mapping $[\cdot,\cdot] : X^* \times X^* \to \mathbb{R}$ defined by:

$$[f,g] := \sum_{n=1}^{+\infty} \frac{1}{2^n} f(x_n)g(x_n).$$

The following assertion is true:

Proposition 2.1.2. $[\cdot,\cdot]$ is an inner product on X^*, i.e. we have:
1) $[f,g] = [g,f]$ for every $f, g \in X^*$,
2) $[f_1 + f_2, g] = [f_1, g] + [f_2, g]$ for every $f_1, f_2, g \in X^*$,
3) $[\lambda f, g] = \lambda[f,g]$ for every $f, g \in X^*$ and $\lambda \in \mathbb{R}$,
4) $[f,f] \geq 0$ for every $f \in X$ and if $[f,f] = 0$, then $f = 0$.

Proof. Only the last assertion is not trivial, namely, if $[f, f] = 0$, then $f = 0$. By the definition we have

$$0 = [f, f] = \sum_{n=1}^{+\infty} \frac{1}{2^n} f^2(x_n),$$

thus $f(x_n) = 0$ for all n. Assume that $f \neq 0$. Then $\exists \delta > 0$ such that

$$\|f\| := \sup_{S_X} f = \delta > 0$$

and we can fix $\bar{x} \in S_X$ such that $|f(\bar{x})| > \delta/2$. By density, there is $n_0 \in \mathbb{N}$ such that $\left\|\bar{x} - x_{n_0}\right\| < 1/2$ and $f(x_{n_0}) = 0$. Now we have

$$|f(\bar{x})| = \left|f(\bar{x} - x_{n_0}) + f(x_{n_0})\right| = \left|f(\bar{x} - x_{n_0})\right| \leq \|f\| \left\|\bar{x} - x_{n_0}\right\| < \delta/2,$$

which is a contradiction with the assumption. □

As a consequence, this inner product induces a natural norm

$$|\!|\!| f |\!|\!| = \sqrt{[f, f]} = (\sum (1/2^n) f^2(x_n))^{1/2}$$

on X^* that satisfies the parallelogram rule:

$$|\!|\!| f + g |\!|\!|^2 + |\!|\!| f - g |\!|\!|^2 = 2|\!|\!| f |\!|\!|^2 + 2|\!|\!| g |\!|\!|^2.$$

The parallelogram rule implies rotundness of the norm trivially: let $f, g \in X^*$ be such that $|\!|\!| f |\!|\!| = |\!|\!| g |\!|\!| = 1$ and $|\!|\!| f + g |\!|\!| = 2$ and after inserting these in the parallelogram rule gives $|\!|\!| f - g |\!|\!| = 0$, or $f = g$. Then, by Definition 1.0.3 we have that $|\!|\!|\cdot|\!|\!|$ is (R) on X^*. Furthermore, note that as $f \to f^2(x_n)$ is strictly convex function, the mapping $f \to \sum (1/2^n) f^2(x_n)$ is also a strictly convex function. In view of Lemma 2.1.1 this means that $|\!|\!|\cdot|\!|\!|$ is also (R).

Now we define a norm with the wanted properties on X^* by

$$|\cdot|_* := \sqrt{\|\cdot\|_*^2 + |\!|\!|\cdot|\!|\!|^2}.$$

We have to check only equivalence and rotundness of this norm. Obviously,

$$|\!|\!| f |\!|\!|^2 = \sum_{n=1}^{+\infty} \frac{1}{2^n} f^2(x_n) \leq \|f\|_*^2 \sum_{n=1}^{+\infty} \frac{1}{2^n} = \|f\|_*^2,$$

thus $\|f\|_* \leq |f|_* \leq \sqrt{2}\,\|f\|_*$. The rotundness follows directly from Lemma 2.1.1 and the observation that $|\cdot|_*^2 := \|\cdot\|_*^2 + |\!|\!|\cdot|\!|\!|^2$ is a strictly convex function as $|\!|\!|\cdot|\!|\!|^2$ is a strictly convex function.

Now we will show that the rotund norm $|\cdot|_*$ defined above, is a dual norm of some equivalent norm $|\cdot|$ on X. It is well known that it is enough to show that $\|\cdot\|_*$ is a w^*-lower semi-continuous. For $f \in X^*$ let $f_k \to^{w^*} f$. Then by the uniform boundedness principle we have $\|f\|_* \leq \liminf \|f_k\|_*$. On the other hand, as $(x_n) \subseteq S_X$ is fixed, again by the uniform boundedness principle (f_k) is bounded and then $(f_k(x_n))$ is bounded. Then the series, which defines the norm $|\!|\!|\cdot|\!|\!|$ is uniformly convergent on (x_n), thus we have

$$\lim_{k\to\infty} |\!|\!| f_k |\!|\!|^2 = \sum_{n=1}^{\infty} \frac{1}{2^n} \lim_{k\to\infty} f_k^2(x_n) = \lim_{k\to\infty} \sum_{n=1}^{\infty} \frac{1}{2^n} f_k^2(x_n) = \sum_{n=1}^{\infty} \frac{1}{2^n} f^2(x_n) = |\!|\!| f |\!|\!|^2$$

Then $|\cdot|_*$ is indeed w^* lower semicontinuous. Then, by the bipolar theorem, $|\cdot|_*$ is dual to an (equivalent) norm $|\cdot|$ on X as it can be seen e.g. in [2, p.126, Lemma 3.97].

Finally, we state the connection between (R) norms and G-norms, see e.g. [1, p.43]. In fact, Proposition 2.3.6 contains this as a partial case.

Proposition 2.1.3. If the dual norm $\|\cdot\|_*$ on X^* of $\|\cdot\|$ is rotund, then the norm $\|\cdot\|$ on X is Gâteaux differentiable.

In this way we constructed explicitly an equivalent norm $|\cdot|$ on any separable space, whose dual is (R), thus the norm $|\cdot|$ is an equivalent G-norm on any separable space. We note once again that the existence of such a norm is well known fact.

2.2 Continuity property of the norm

This section is dedicated to the central result of the chapter. We establish a continuity property of the norm of the subdifferential of the Moreau-Yosida regularization of a convex function on a Banach space. The results from this section are is published in [8].

Theorem 2.2.1. Let $(X, \|\cdot\|)$ be a Banach Space and let $f : X \to \mathbb{R} \cup \{\infty\}$ be a proper, convex, and bounded below function. Then its Moreau-Yosida regularization f_λ, where $\lambda > 0$, has the property that the multi-valued mapping

$$x \to \|\partial f_\lambda(x)\|,$$

defined by Definition 2.2, is continuous in the following sense: $\forall x \in X$ and $\forall \varepsilon > 0$ $\exists \delta > 0$ such that if $\|y - x\| < \delta$ and $p \in \partial f_\lambda(x)$, $q \in \partial f_\lambda(y)$, then

$$\Big| \|p\| - \|q\| \Big| < \varepsilon.$$

Note that the boundedness of f is used only to ensure that Moreau-Yosida regularization is well defined, thus can be replaced with any condition ensuring the latter, for example $f(x) > -c(1 + \|x\|)$. In order to verify this result we will use the auxiliary function

$$h(x) := f_{\frac{1}{2}}(x) = \inf_{y \in X}\{f(y) + \|x - y\|^2\} \tag{2.2}$$

and will prove the theorem for $\lambda = \frac{1}{2}$. Note that the proof works in a similar way for every $\lambda > 0$ as $f_\lambda(x) = \frac{1}{2\lambda} \inf_{y \in X}\{2\lambda f(y) + \|x - y\|^2\}$ and the function $2\lambda f$ is proper, convex and bounded below if f is such.

Definition 2.2.2. We define the ε-argmin set for the function h defined by (2.2) and any $\varepsilon > 0$ as

$$j_\varepsilon(x) := \{y \in X : f(y) + \|x - y\|^2 < h(x) + \varepsilon\}.$$

It is obvious that the set $j_\varepsilon(x)$ is non-empty and convex $\forall x \in X$. Also, $j_{\varepsilon_1}(x) \subseteq j_{\varepsilon_2}(x)$ for $\varepsilon_1 \leq \varepsilon_2$.

Now we need several lemmas concerning the properties of the ε-argmin.

Lemma 2.2.3. If $y, z \in j_\varepsilon(x)$, then $\Big|\|x - y\| - \|x - z\|\Big| < 2\sqrt{\varepsilon}$. Thus $j_\varepsilon(x)$ is bounded $\forall x \in X$, $\forall \varepsilon > 0$.

Proof. We assume that for $\varepsilon > 0$ the set $j_\varepsilon(x)$ is not a singleton as otherwise there is nothing to prove. Without loss of generality we can fix $x = 0$, $h(x) = 0$. Then pick $y, z \in j_\varepsilon(0)$, $y \neq z$. Let $2w = y + z$, then $w \in j_\varepsilon(0)$ as $j_\varepsilon(0)$ is convex. Then we have

$$f(y) < -\|y\|^2 + \varepsilon,$$

$$f(z) < -\|z\|^2 + \varepsilon$$

and, on the other hand, $0 = h(0) \leq f(w) + \|w\|^2$ so

$$-\|w\|^2 \leq f(w).$$

By convexity of f we obtain

$$-2\left\|\frac{y+z}{2}\right\|^2 = -2\|w\|^2 \leq 2f(w) \leq f(y) + f(z) < 2\varepsilon - \|y\|^2 - \|z\|^2,$$

which yields

$$2\|y\|^2 + 2\|z\|^2 < \|y + z\|^2 + 4\varepsilon.$$

But as

$$\|y + z\|^2 \leq [\|y\| + \|z\|]^2 = \|y\|^2 + \|z\|^2 + 2\|y\|\,\|z\|$$

it follows that

$$\|y\|^2 + \|z\|^2 - 2\|y\|\|z\| = [\|y\| - \|z\|]^2 < 4\varepsilon.$$

Thus

$$\Big|\|y\| - \|z\|\Big| < 2\sqrt{\varepsilon}.$$

Now, we will show that $j_\varepsilon(x)$ is bounded. Fix some $z \in j_\varepsilon(x)$. Choose any $y \in j_\varepsilon(x)$. Then we have

$$\|y - x\| \le \|y - z\| + \|z - x\| \le \|y - z\| - \|z - x\| + 2\|z - x\| \le 2\sqrt{\varepsilon} + 2\|z - x\|.$$

Then $\|y\| \le \|x\| + \|y - x\|$, so

$$\|y\| \le \|x\| + 2\sqrt{\varepsilon} + 2\|z - x\|$$

and the right hand side does not depend on y. □

Note that Lemma 2.2.3 implies that for any finite set of points $\{x_i\}_{i=1}^n$ one can find a constant K such that $\cup_{i=1}^n j_\varepsilon(x_i) \subseteq KB_X$. Thus we are able to prove

Lemma 2.2.4. Let $j_\varepsilon(x) \subseteq KB_X$ and $j_\varepsilon(y) \subseteq KB_X$ for some constant $K > 0$. Then h is $4K$-Lipschitz on KB_X.

Proof. Let $x, y \in KB_X$. Let $z_n \in j_{\varepsilon_n}(x)$ for some $\varepsilon_n \downarrow 0$. Then we have

$$-h(x) < -f(z_n) - \|z_n - x\|^2 + \varepsilon_n,$$

$$h(y) \le f(z_n) + \|z_n - y\|^2.$$

After summing these inequalities we obtain

$$h(y) - h(x) < \|z_n - y\|^2 - \|z_n - x\|^2 + \varepsilon_n.$$

Now we note that for n large enough $z_n \in j_\varepsilon(x) \subseteq KB_X$, thus $z_n - y$ and $z_n - x$ are in $2KB_X$. Then we have

$$\begin{aligned}\|z_n - y\|^2 - \|z_n - x\|^2 &\le [\|z_n - y\| - \|z_n - x\|][\|z_n - y\| + \|z_n - x\|] \\ &\le 4K\|x - y\|.\end{aligned}$$

This means exactly $h(y) - h(x) \le 4K\|y - x\| + \varepsilon_n \to 4K\|y - x\|$ as $\varepsilon_n \downarrow 0$. □

Lemma 2.2.5. For every $x \in X$ and every $\varepsilon > 0$ there exists $c(\varepsilon) > 1$ such that for every $\delta : 0 < \delta < \varepsilon$ and every $y \in X$ with $\|x - y\| < \delta$ we have

$$j_\delta(x) \subseteq j_{c(\varepsilon)\delta}(y).$$

Proof. We fix some $x \in X$ and $\varepsilon > 0$. Then we fix $\delta : 0 < \delta < \varepsilon$ and $y \in X$ such that $\|x - y\| < \delta$. We note that

$$h(x) = \inf_{z \in X}\{f(z) + \|z - x\|^2\} \leq f(x). \tag{2.3}$$

Denote $U_K(x) = \{z \in X : \|z - x\| \geq K\}$ for $K \geq 0$. It is clear that $X \setminus U_K(x)$ is bounded. Then for every $z \in U_K(x)$ we have

$$f(z) + \|z - x\|^2 \geq \inf f + K^2. \tag{2.4}$$

Then the function $F(K) := \inf f + K^2$ is strictly increasing and unbounded above and we can find some $K_1 = K_1(\varepsilon) > 0$ such that

$$f(x) + \delta < f(x) + \varepsilon < \inf f + K_1^2.$$

The last inequality, (2.3) and (2.4) yield

$$\forall z \in U_{K_1}(x) \implies h(x) + \delta < f(z) + \|z - x\|^2.$$

Then $j_\delta(x) \subseteq X \setminus U_{K_1}(x)$. Similarly, we can find $K_2 = K_2(\varepsilon) > 0$ with $j_\delta(y) \subseteq X \setminus U_{K_2}(y)$. As $X \setminus U_{K_1}(x)$ and $X \setminus U_{K_2}(y)$ are both bounded, we can find some $K = K(\varepsilon) > 0$ such that $(X \setminus U_{K_1}(x)) \cup (X \setminus U_{K_2}(y)) \subseteq KB_X$. In particular, $j_\delta(x) \cup j_\delta(y) \subseteq KB_X$. Furthermore, we note that $x \notin U_{K_1}(x)$ and $y \notin U_{K_2}(y)$, thus $x, y \in KB_X$. For every $z \in j_\delta(x)$ we have $z \in KB_X$, so $y - z \in 2KB_X$ and $x - z \in 2KB_X$. Then we have

$$\begin{aligned} f(z) + \|z - y\|^2 &= f(z) + \|z - x\|^2 + \|z - y\|^2 - \|z - x\|^2 \\ &< h(x) + \delta + \|x - y\|[\|z - x\| + \|z - y\|] \\ &\leq h(x) + \delta + \delta 4K \\ &= h(y) + h(x) - h(y) + \delta[1 + 4K]. \end{aligned}$$

Then, by Lemma 2.2.4 it follows that h is $4K$-Lipschitz on KB_X, so we finally obtain

$$f(z) + \|z - y\|^2 < h(y) + \delta 4K + \delta[1 + 4K] = h(y) + \delta[1 + 8K].$$

Now we set $c(\varepsilon) := 1 + 8K(\varepsilon) > 1$ to finish the proof. □

Lemma 2.2.6. Let $p \in \partial h(x)$ and $y \in j_\varepsilon(x)$. Then

$$-p \subseteq \partial_\varepsilon g(x - y).$$

Proof. As $p \in \partial h(x)$, for every $z \in X$ we have

$$p(-z) = p(x-z) - p(x) \leq h(x-z) - h(x). \tag{2.5}$$

As $y \in j_\varepsilon(x)$, we have

$$h(x) \leq f(y) + g(y-x) < h(x) + \varepsilon,$$

thus

$$-h(x) < -f(y) - g(y-x) + \varepsilon. \tag{2.6}$$

Furthermore, by the definition of h for every $z \in X$ we have

$$h(x-z) \leq f(y) + g(y-(x-z)). \tag{2.7}$$

Inserting (2.6) and (2.7) into (2.5) yields

$$p(-z) < f(y) + g(y-x+z) - f(y) - g(y-x) + \varepsilon,$$

which is equivalent to

$$g(y-x) - [-p(y-x)] < g(y-x+z) - [-p(y-x+z)] + \varepsilon.$$

Thus, $-p \in \partial_\varepsilon g(y-x)$. □

For the next lemma we will need the Brøndsted-Rockafellar theorem, see e.g. [4].

Theorem 2.2.7 (Brøndsted-Rockafellar). Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be a proper, convex and lower semicontinuous function and $x_0 \in \mathrm{dom} f$. Take $\varepsilon > 0$ and $x_0^* \in \partial_\varepsilon f(x_0)$. Then for every $\lambda > 0$ one can find $x \in X$, $x^* \in X^*$ such that

$$x^* \in \partial f(x),\ \|x - x_0\| \leq \frac{\varepsilon}{\lambda},\ \left\|x^* - x_0^*\right\| \leq \varepsilon.$$

Lemma 2.2.8. Let $p \in \partial h(x)$ and $y \in j_{\varepsilon^2}(x)$. Then

$$\Big|\|p\| - 2\|x-y\|\Big| \leq 3\varepsilon.$$

Proof. As $p \in \partial h(x)$ and $y \in j_{\varepsilon^2}(x)$, Lemma 2.2.6 yields that $-p \in \partial_{\varepsilon^2} g(x-y)$. Set $q := -p$, $u = x - y$, i.e. $q \in \partial_{\varepsilon^2} g(u)$. Applying Brøndsted-Rockafellar theorem to g one can find $z \in X$, $z^* \in X^*$ such that for $\lambda = \varepsilon$ we have

$$z^* \in \partial \|\cdot\|^2 (z),\ \|z^* - q\| \leq \varepsilon,\ \|z - u\| \leq \varepsilon.$$

Next, note that as $z^* \in \partial \|\cdot\|^2 (z)$, we have that $\|z^*\| = 2\|z\|$. Then

$$\begin{aligned}
\Big|\|p\| - 2\|x-y\|\Big| &= \Big|\|q\| - 2\|u\|\Big| \\
&= \Big|\, \|q\| - \|z^*\| + \|z^*\| - 2\|z\| + 2\|z\| - 2\|u\| \,\Big| \\
&= \Big|\, \|q\| - \|z^*\| + 2\|z\| - 2\|u\| \,\Big| \\
&\leq \|q - z^*\| + 2\|z-u\| \\
&\leq 3\varepsilon.
\end{aligned}$$

□

Now we are ready to prove Theorem 2.2.1.

Proof. (of Theorem 2.2.1*)* Let $p \in \partial h(x)$ and $q \in \partial h(y)$. We will prove that $\forall \varepsilon > 0$ $\exists \delta \in (0, \varepsilon)$ such that if

$$\|x-y\| < \delta, \text{ then } \Big|\|p\| - \|q\|\Big| < 8\varepsilon.$$

First we fix $x \in X$ and $\varepsilon > 0$. The first step is to show that we can find $\delta \in (0, \varepsilon)$ such that if $\|x-y\| < \delta$, then $j_{\varepsilon^2}(x) \cap j_{\varepsilon^2}(y) \neq \emptyset$. By Lemma 2.2.5 we have that for ε^2 we can find $c := c(\varepsilon^2) > 1$ such that for every $\delta : 0 < \delta < \varepsilon^2$ with $\|x-y\| < \delta$ we have $j_\delta(x) \subseteq j_{c(\varepsilon^2)\delta}(y)$. Then we can find $\delta < \min\{\varepsilon, \varepsilon^2, \frac{\varepsilon^2}{c(\varepsilon^2)}\}$, which means that $\delta < \varepsilon$, $c(\varepsilon^2)\delta < \varepsilon^2$ and $\delta < \varepsilon^2$. Thus we have

$$j_\delta(x) \subseteq j_{c(\varepsilon^2)\delta}(y) \subseteq j_{\varepsilon^2}(y)$$

and

$$j_\delta(x) \subseteq j_{\varepsilon^2}(x).$$

As $j_\delta(x)$ is never empty for $\delta > 0$, it follows that

$$\emptyset \neq j_\delta(x) \subseteq j_{\varepsilon^2}(x) \cap j_{\varepsilon^2}(y).$$

For the next step we choose some $z \in j_{\varepsilon^2}(x) \cap j_{\varepsilon^2}(y)$. Then Lemma 2.2.8 yields that

$$\Big|\, \|p\| - 2\|z-x\| \,\Big| < 3\varepsilon,$$

$$\Big|\, \|q\| - 2\|z-y\| \,\Big| < 3\varepsilon$$

and we have

$$\begin{aligned}\Big|\,\|p\| - \|q\|\,\Big| &= \Big|\,\|p\| - 2\,\|z-x\| + 2\,\|z-x\| - 2\,\|z-y\| + 2\,\|z-y\| - \|q\|\,\Big| \\ &\le \Big|\,\|p\| - 2\,\|z-x\|\,\Big| + 2\Big|\,\|z-x\| - \|z-y\|\,\Big| + \Big|2\,\|z-y\| - \|q\|\,\Big| \\ &\le 3\varepsilon + 2\,\|x-y\| + 3\varepsilon \\ &\le 8\varepsilon\end{aligned}$$

as δ is fixed in $(0,\varepsilon)$, which gives the continuity of the mapping $x \to \|\partial h(x)\|$. □

2.3 Additional properties of the Moreau-Yosida regularization

The aim of this section is to present some well-known properties of the Moreau-Yosida regularization and an important consequence (Theorem 2.3.1) of the main Theorem 2.2.1 of Section 2.2. A brief version of what follows is again covered by [8].

Theorem 2.3.1. If $(X, \|\cdot\|)$ is a Banach space such that $\|\cdot\|_*$ is strictly convex, then the Moreau-Yosida regularization f_λ for any $\lambda > 0$ of a proper, convex, lower semicontinuos and bounded below $f : X \to \mathbb{R} \cup \{+\infty\}$ is Gâteaux differentiable at every point $x \in X$ and the mapping

$$x \to \|f'_\lambda(x)\|$$

is continuous.

In view of Section 2.1, it is clear that for separable Banach spaces the existence of an equivalent norm, whose dual is rotund, makes possible the definition of a Moreau-Yosida regularization with this property. Thus in this section we will assume that the former norm $\|\cdot\|$ of X has a rotund dual. Again we can assume that $\lambda = 1/2$ and consider the function $g(x) := g_{1/2}(x) = \|x\|^2$, and the auxiliary function

$$h(x) := f_{1/2}(x) = \inf_{y \in X}\{f(y) + g(x-y)\}.$$

Note that the choice of λ is not restrictive for what follows as one can always again consider the function $\frac{1}{2\lambda}\inf\{2\lambda f(y) + g(x-y)\}$ instead, as it has the same properties as $h(x)$. First we recall the notion of a convex conjugate of a proper function $\varphi : X \to \mathbb{R} \cup \{+\infty\}$. This is a function $\varphi^* : X^* \to \mathbb{R} \cup \{+\infty\}$ defined as

$$\varphi^*(p) := \sup_{x\in X}\{p(x) - \varphi(x)\}. \tag{2.8}$$

Note that φ^* is always convex and lower semicontinuous (closed) on X^*. The function $g = \|\cdot\|^2$ has some nice (well-known) properties concerning its convex conjugate g^* which we state in the next proposition.

Proposition 2.3.2 (Properties of g^*)**.**

(i) $g^* = \frac{1}{4}\|\cdot\|_*^2$
(ii) g^* is strictly convex.

Proof. (i) First note that $p(x) - \|x\|^2 \le \|p\|\,\|x\| - \|x\|^2 \le \frac{1}{4}\|p\|^2 - (\frac{1}{2}\|p\| - \|x\|)^2 \le \frac{1}{4}\|p\|^2$. Then we have $g^*(p) \le \frac{1}{4}\|p\|^2$. Next we show that this is in fact an equality. Let us take a maximizing sequence $(x_n) \subseteq S_X$ such that $\lim_{n\to\infty} p(x_n) = \|p\|$. It follows that

$$p\Big(\frac{\|p\|}{2}x_n\Big) - \Big\|\frac{\|p\|}{2}x_n\Big\|^2 = \frac{\|p\|}{2}p(x_n) - \frac{\|p\|^2}{4}\|x_n\|^2 \to \frac{\|p\|^2}{4}.$$

(ii) For this one we note that as $\|\cdot\|_*$ is (R), Lemma 2.1.1 implies the strict convexity of g^* directly. □

The next proposition is also a well-known fact which makes clear the geometry of the convex conjugate h^* of the Moreau-Yosida regularization h.

Proposition 2.3.3. $h^*(p) = (f\square g)^*(p) = f^*(p) + g^*(p)$ for every $p \in X^*$.

Proof. First we show that $(f\square g)^* \le f^* + g^*$. For every $\varepsilon > 0$ we can find $\bar{x} \in X$ such that

$$(f\square g)^*(p) \le p(\bar{x}) - (f\square g)(\bar{x}) + \varepsilon = p(\bar{x}) + \sup_{y\in X}\{-f(\bar{x}-y) - g(y)\} + \varepsilon.$$

Then we can find $\bar{y} \in X$ with

$$\sup_{y\in X}\{-f(\bar{x}-y) - g(y)\} \le -f(\bar{x}-\bar{y}) - g(\bar{y}) + \varepsilon.$$

Thus, we obtain

$$\begin{aligned}(f\square g)^*(p) &\le p(\bar{x}-\bar{y}) + p(\bar{y}) - f(\bar{x}-\bar{y}) - g(\bar{y}) + 2\varepsilon\\ &\le \sup_{z\in X}\{p(z) - f(z)\} + \sup_{z\in X}\{p(z) - g(z)\} + 2\varepsilon\\ &= f^*(p) + g^*(p) + 2\varepsilon\end{aligned}$$

and the inequality is proved as the latter holds true for every $\varepsilon > 0$.

Next we will show that $f^* + g^* \geq (f \square g)^*$. For every $\varepsilon > 0$ we can find $\bar{x}, \bar{y} \in X$ such that

$$f^*(p) \leq p(\bar{x}) - f(\bar{x}) + \varepsilon,$$

$$g^*(p) \leq p(\bar{y}) - g(\bar{y}) + \varepsilon.$$

Then

$$f^*(p) + g^*(p) \leq p(\bar{x} + \bar{y}) - f(\bar{x}) - g(\bar{y}) + 2\varepsilon.$$

Set $\bar{z} := \bar{x} + \bar{y}$, then $\bar{x} = \bar{z} - \bar{y}$. Thus, we have

$$\begin{aligned} f^*(p) + g^*(p) &\leq p(\bar{z}) - f(\bar{z} - \bar{y}) - g(\bar{y}) + 2\varepsilon \\ &\leq \sup_{z \in X}\{p(z) - f(z - \bar{y}) - g(\bar{y})\} + 2\varepsilon \\ &\leq \sup_{z \in X}\{p(z) - \inf_{u \in X}\{f(z - u) + g(u)\}\} + 2\varepsilon \\ &= \sup_{z \in X}\{p(z) - f \square g(z)\} + 2\varepsilon \\ &= (f \square g)^*(p) + 2\varepsilon \end{aligned}$$

and the inequality is proved as the latter holds true for every $\varepsilon > 0$. □

As f^* is convex, from Proposition 2.3.2 and Proposition 2.3.3 directly follows the next result.

Corollary 2.3.4. h^* is a strictly convex function.

The aim of the following is to show that h is a differentiable function. In [2, p.340] we can find Theorem 7.17 which states:

Theorem 2.3.5. Let φ be a convex continuous function defined on a nonempty open convex subset U of a Banach space X and let $x_0 \in U$. The following are equivalent:
(i) φ is Gâteaux differentiable at x_0.
(ii) $x_n^* \to_{w^*} x_0^*$ whenever $x_0^* \in \partial\varphi(x_0)$ and $x_n^* \in \partial_{\varepsilon_n}\varphi(x_0)$, $\varepsilon_n \downarrow 0$.
(iii) $\partial\varphi(x_0)$ is a singleton.

We need only one more fact to conclude. It is well-known and can be considered as a generalization of Šmulyan's test, see e.g. [1, p.4].

Proposition 2.3.6. Let $\varphi : X \to \mathbb{R}$ be a continuous and convex function such that $\varphi^* : X^* \to \mathbb{R}$ is a strictly convex function. Then φ is Gâteaux differentiable at every point $x \in X$.

Proof. First we claim that $p \in \partial\varphi(x)$ if and only if $\varphi^*(p) + \varphi(x) = p(x)$. Let $\varphi^*(p) + \varphi(x) = p(x)$ be true. Then we have for every $\bar{y} \in X$

$$p(\bar{y}) - \varphi(\bar{y}) \leq \sup\{p(y) - \varphi(y)\} = p(x) - \varphi(x)$$

or, $p \in \partial\varphi(x)$. Now, let $p \in \partial\varphi(x)$.

By the definition of the convex conjugate (see (2.8)) it trivially follows that

$$\varphi^*(p) + \varphi(x) \geq p(x).$$

On the other hand, as $p \in \partial\varphi(x)$, then $p(y) - \varphi(y) \leq p(x) - \varphi(x)$ for all $y \in X$, thus $\varphi^*(p) \leq p(x) - \varphi(x)$ and the assertion is proved.

Next, we assume that $\partial\varphi(x)$ is not a singleton, thus there are $p_i \in \partial\varphi(x)$, $i = 1, 2$ and $p_1 \neq p_2$. As the subdifferential set is convex, it follows for any $\lambda \in (0, 1)$ that

$$\lambda p_1 + (1 - \lambda)p_2 \in \partial\varphi(x).$$

This means that $\varphi^*(\lambda p_1 + (1 - \lambda)p_2) + \varphi(x) = \langle \lambda p_1 + (1 - \lambda)p_2, x \rangle$.

Also, we have

$$\lambda[\varphi^*(p_1) + \varphi(x)] = \lambda p_1(x),$$

$$(1 - \lambda)[\varphi^*(p_2) + \varphi(x)] = (1 - \lambda)p_2(x),$$

and after summing the last two equations we obtain

$$\lambda\varphi^*(p_1) + (1 - \lambda)\varphi^*(p_2) + \varphi(x) = \langle \lambda p_1 + (1 - \lambda)p_2, x \rangle .$$

This means exactly that $\varphi^*(\lambda p_1 + (1 - \lambda)p_2) = \lambda\varphi^*(p_1) + (1 - \lambda)\varphi^*(p_2)$ which contradicts with the strict convexity of φ^*. As $\partial\varphi(x)$ is never empty, $\partial\varphi(x)$ is a singleton and by Theorem 2.3.5 it follows that φ is Gâteaux differentiable everywhere. □

Now we will present the proof of Theorem 2.3.1.

Proof. (*of Theorem* 2.3.1) The proof is now straightforward: we apply Proposition 2.3.6 to the function h (which is convex by Corollary1.0.8 and obviously continuous) and to the strictly convex h^* (by Corollary 2.3.4) to obtain that the Moreau-Yosida regularization h is Gâteaux differentiable at every point $x \in X$. Finally, as in this case $\partial h(x) = \{h'(x)\}$ we have by Theorem 2.2.1 that the mapping

$$x \to \|h'(x)\|$$

is continuous. □

Chapter 3

Multidirectional Inequalities

The section below deals with inequalities, which are multidirectional in their nature and represent a natural generalization of the one dimensional Mean Value Theorem of Lagrange. Instead of dealing only with differentiable functions, the aim here is to widen the class of functions to which such inequalities are applicable by replacing the smoothness condition with subdifferentiablity. On the other hand, as there are many different types of subdifferentials, our aim is to extend their set to very general axiomatically defined subdifferential. Throughout this chapter, we will use the notation $[A, B]$ for the interval defined by two subsets in the underlying space X :

$$[A, B] := \{z \in X : z = \lambda x + (1 - \lambda)y \ \forall x \in X, \forall y \in Y, \forall \lambda \in [0, 1]\}.$$

Note that when A and B are convex and closed, $[A, B]$ coincides with the convex hull of $A \cup B$. Furthermore, we will use also a notation $A_\delta := A + \delta B_X$, i.e. a lower positive real index δ of a set A will denote the points in the set A extended with a closed neighbourhood of thickness δ.

3.1 Motivation and former results

In this section we discuss the first in historical sense multidirectional mean value inequality which is due to Clarke and Ledyaev, see [9, p.340]. The classical Mean Value Theorem in its inequality form asserts that if $f : \mathbb{R} \to \mathbb{R}$ is a smooth function on the interval $[x, y]$ one can always find a point $\xi \in (x, y)$ such that

$$\langle f'(\xi), y - x\rangle \geq f(y) - f(x).$$

This result remains true when the underlying space is of higher dimension and $[x, y]$ is a line segment in it, but remains purely one-dimensional in its nature. Basically, the original result (Theorem 2.1. found in [9]) extends this classical Mean

Value Theorem to many directions (as multidirectional inequality suggests). Its statement follows here.

Theorem 3.1.1. Let X be a Banach space and $A, B \subseteq X$ be nonempty, closed, convex and bounded sets. Furthermore, $f : X \to \mathbb{R}$ is supposed to be continuously Gâteaux differentiable in a neighbourhood of $[A, B]$. In addition, let at least one of A and B be compact and let $\varepsilon > 0$ be given. Then one can find $\xi \in [A, B]$ such that

$$\inf_B f - \sup_A f < \langle f'(\xi), y - x \rangle + \varepsilon \ \forall x \in A, \forall y \in B.$$

In the same work [9] one can find also a relaxation of the smooth requirement, which involves the Clarke subdifferential. Recall that the generalized directional derivative of a locally Lipschitz function $f : X \to \mathbb{R}$ at the point x in direction v is defined as

$$f^\circ(x, v) := \limsup_{x' \to x, \lambda \downarrow 0} \frac{f(x' + \lambda v) - f(x')}{\lambda}$$

and then the Clarke subdifferential of f at x is defined as follows

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

Now, for the following corollary one only needs f to be locally Lipschitz on $[A, B]$. It corresponds to Theorem 4.1. from [9].

Corollary 3.1.2. Let X be a Banach space and $A, B \subseteq X$ be nonempty, closed, convex and bounded sets. Furthermore, $f : X \to \mathbb{R}$ is supposed to be locally Lipschitz on $[A, B]$. In addition, let at least one of A and B be compact and let $\varepsilon > 0$ be given. Then one can find $\xi \in [A, B]$ and $p \in \partial_C f(\xi)$ such that

$$\inf_B f - \sup_A f < \langle p, y - x \rangle + \varepsilon \ \forall x \in A, \forall y \in B.$$

We would like to stress the fact that the result presented here relies on setting in which at least one of the sets A and B is compact (which doesn't matter when the underlying space X is finite dimensional, but is essential in general Banach space) and that the subdifferential here is concretely the one of Clarke. Later on we will discuss some similarities and differences of our main chapter result and the original. In some sense, our result is closer to the second Clarke-Ledyiaev inequality, see [10].

3.2 Axiomatic definition of an abstract subdifferential

In Section 3.1 we recalled a result, concerning a concrete subdifferential the Clarke one. As there are many different kind of subdifferentials, it is often hard to say if

a certain result holds for a particular subdifferential. A way around this problem is defining a subdifferential by its properties rather than by a direct construction – via a set of axioms, a way proposed by Ioffe and established by Thibault and others. The aim of this section is to define a general type of subdifferential in an axiomatic way which we will call *feasible*.

Let X be a Banach space. First we recall that a subdifferential operator ∂ applied to a lower semicontinuous function $f : X \to \mathbb{R} \cup \{+\infty\}$ generates a multi-valued map

$$\partial f : X \to 2^{X^*}.$$

Definition 3.2.1. Let X be a Banach space and $f :\to \mathbb{R} \cup \{+\infty\}$ be a proper lower semicontinuous function. We say that the subdifferential ∂ is *feasible* if the following properties hold:

(P1) $\partial f(x) = \varnothing$ if $x \notin \mathrm{dom} f$;

(P2) $\partial f(x) = \partial g(x)$ whenever f and g coincide on a neighbourhood of x;

(P3) If f is a convex and continuous function in a neighbourhood of x then $\partial f(x)$ coincides with the standard subdifferential in Convex analysis (see Definition 1.0.4);

(P4) If g is a convex and continuous in a neighbourhood of z and $f + g$ has a local minimum at z then for each $\varepsilon > 0$ there are $p \in \partial f(x)$ and $q \in \partial g(y)$ such that

$$\|x - z\| < \varepsilon,\ \|y - z\| < \varepsilon,\ |f(x) - f(z)| < \varepsilon \text{ and } \|p + q\| < \varepsilon.$$

Axioms (P1), (P2) and (P3) are principal, while (P4) is technical - this is the so called *fuzzy* sum rule. For our purposes it will be more suitable to use another axiom (P4') – which is equivalent to (P4) – as it encapsulates the standard application of the Ekeland variational principle. This is regulated in the following

Proposition 3.2.2. Let X be a Banach space and let $f : X \to \mathbb{R} \cup \{+\infty\}$ be a proper and lower semicontinuous function. If the subdifferential ∂ is feasible, then (P4) is equivalent to

(P4') If g is convex and continuous and if $f + g$ is bounded below, then there are sequences $p_n \in \partial f(x_n)$ and $q_n \in \partial g(y_n)$ such that

$$\|x_n - y_n\| \to 0,\ (f(x_n) + g(y_n)) \to \inf(f + g) \text{ and } \|p_n + q_n\| \to 0.$$

Proof. First we prove that (P4) $\to$ (P4'). Let $(z_n)_1^\infty$ be a minimizing sequence for $f+g$, that is $f(z_n) + g(z_n) < \inf(f + g) + \varepsilon_n$ for some $\varepsilon_n \downarrow 0$. By Ekeland variational principle (Theorem 1.0.11) one can find $u_n \in X$ such that $\|u_n - z_n\| \le \varepsilon_n$ and, in addition,

$$f(x) + g(x) + \varepsilon_n \|x - u_n\| \ge f(u_n) + g(u_n).$$

Set $\hat{g}_n(x) := g(x) + \varepsilon_n \|x - u_n\|$. Since $f + \hat{g}_n$ has a minimum at u_n and $\hat{g}_n$ is convex and continuous, by (P4) there are $p_n \in \partial f(x_n)$ and $\hat{q}_n \in \partial \hat{g}_n(y_n)$ such that $\|x_n - u_n\| < \varepsilon_n$,$\|y_n - u_n\| < \varepsilon_n$,$|f(x_n) - f(u_n)| < \varepsilon_n$, and $\|p_n + \hat{q}_n\| < \varepsilon_n$.

All conclusions of (P4') except the last one follow from the triangle inequality. For the last one we note that any subgradient of $\varepsilon_n \|\cdot - u_n\|$ is of norm less or equal to ε_n. By the Sum Theorem of Convex Analysis, (see e.g. Penôt, p.206), there is $q_n \in \partial g(y_n)$ such that $\|q_n - \hat{q}_n\| \leq \varepsilon_n$. Then we have $\|p_n + q_n\| \leq 2\varepsilon_n \to 0$.

Next we prove the (P4') $\to$ (P4) direction. Let the bounded below $f + g$ attains a local minimum. Without loss of generality we can assume that this minimum is global - we only need to cut the functions in a proper way, for example, if $\bar{z}$ is the local minimum, we will consider the functions

$$f_1(x) := \begin{cases} f(x), & x \in \bar{z} + \delta_1 B_X \\ +\infty, & \text{else,} \end{cases} \text{ and } g_1(x) := \begin{cases} g(x), & x \in \bar{z} + \delta_1 B_X \\ +\infty, & \text{else,} \end{cases}$$

where $\delta_1 > 0$ is small enough. Thus, z would be a global minimum for the function $f_1 + g_1$ and we will proceed without considering neighbourhoods. Then, by (P4') we find x_n, y_n, p_n and q_n such that

$$\|x_n - y_n\| \to 0, \ (f(x_n) + g(y_n)) \to \min(f + g) \text{ and } \|p_n + q_n\| \to 0.$$

We will fix some $\varepsilon > 0$ and by continuity of the function g we can find N large enough such that

$$|g(x_N) - g(y_N)| \leq \frac{\varepsilon}{4}$$

and

$$\|x_N - y_N\| \leq \frac{\varepsilon}{2}, \ f(x_N) + g(y_N) < \min(f + g) + \frac{\varepsilon}{4} \text{ and } \|p_n + q_n\| < \varepsilon.$$

Set $x := x_N$, $y := y_N$, $p := p_N$ and $q := q_N$. Then, from the continuity of g and the estimate for the minimum of $f + g$ at x we obtain

$$f(x) + g(x) - g(x) + g(y) = f(x) + g(y) < \min(f + g) + \frac{\varepsilon}{2}.$$

After rearranging, we obtain

$$f(x) + g(x) < \min(f + g) + \frac{\varepsilon}{2}.$$

The latter means that x is a $\varepsilon/2$-minimizer of $f + g$. Now we apply the EVP (see Theorem 1.0.11) at the point x. Let us choose $0 < \delta < \varepsilon/2$ such that if $\|x - z\| < \delta$, then $|g(x) - g(z)| < \varepsilon/4$. Set $\rho := \delta$, then set $\gamma = \varepsilon/(2\rho)$, thus $\gamma\rho = \varepsilon/2$. The we find $z \in X$ such that

$$\begin{aligned} &\|z - x\| \leq \delta, \\ &f(z) + g(z) < f(x) + g(x) + \gamma \|z - x\|. \end{aligned}$$

The latter estimate gives us

$$f(z) - f(x) < g(x) - g(z) + \gamma \|z - x\| \leq \delta < \frac{\varepsilon}{4} + \frac{\varepsilon}{2} = \frac{3\varepsilon}{4}.$$

On the other hand, as x is $\varepsilon/2$-minimizer of $f+g$ we have at the point z the following estimate

$$f(x) - f(z) < g(z) - g(x) + \frac{\varepsilon}{2} < \frac{\varepsilon}{4} + \frac{\varepsilon}{2} = \frac{3\varepsilon}{4}.$$

It follows that $|f(x) - f(z)| < \varepsilon$. The other estimates of (P4) are clear by the choice of δ.

□

Here we would like to state several remarks about the feasible subdifferential. The axioms (P1), (P2) and (P3) are very common. The form here is essentially the one found in [13] and the fuzzy sum rule (P4) here is formally stronger than the corresponding one in [13], but we are not aware of an example showing that it is actually stronger. This can be an interesting open question.

It immediately follows from Corollary 4.64 [12, p.305] that the smooth subdifferential in a smooth Banach space (the types of smoothness synchronized, of course) is feasible. Note that (P4) is modelled after the corresponding property for the smooth subdifferential.

Corollary 5.52 [12, p.385] and Theorem 7.23 [12, p. 475] imply that the Clarke subdifferential and the G-subdifferential of Ioffe satisfy even more than (P4), that is if z is a local minimum of $f + g$, then

$$0 \in \partial f(z) + \partial g(z).$$

Similarly, it is known that the limiting subdifferential on Asplund space, considered by Morduchovich and others, is feasible.

All that was said above implies that most of the subdifferentials are feasible under natural assumptions on their underlying spaces.

3.3 Statement of the new multidirectional inequality and comparison with other results

This section is dedicated to the main result of the chapter, a multidirectional inequality of new type, stated in Theorem 3.3.1. This result can be found in [7]. Then we make a comparison with other similar results.

Theorem 3.3.1. Let X be a Banach space and let ∂ be a feasible subdifferential. Let A and B be non-empty, closed, bounded and convex subsets of X. Let $f : X \to \mathbb{R} \cup \{+\infty\}$

be a proper lower semicontinuous function such that $A \cap \mathrm{dom} f \neq \emptyset$. Let f be bounded below on the set $C := \overline{[A,B]_\delta}$ for some $\delta > 0$. Let

$$\mu < \inf_C f. \tag{3.1}$$

Let $r, s \in \mathbb{R}$ be such that

$$r = \inf_A f, \;\; s < \inf_{B_\delta} f. \tag{3.2}$$

Then, for every $\varepsilon > 0$ there are $\xi \in [A,B]_\delta$ and $p \in \partial f(\xi)$ such that

$$f(\xi) < \inf_{[A,B]} f + |r - s| + \varepsilon, \tag{3.3}$$

$$\|p\| < \frac{\max\{r,s\} - \mu}{r} + \varepsilon, \tag{3.4}$$

and

$$\inf_B p - \inf_A p > s - r. \tag{3.5}$$

First we will make a comparison to the original result, stated in Section 3.1. Our multidirectional inequality (3.5) is of a different nature, as it involves difference of infimums instead of difference of infimum and supremum. On the other hand, if A is replaced with a singleton, i.e. $A = \{a\}$ (this is what we call a *point-set* setting to distinguish it from the general case, which is a *set-set* setting) our result implies something very close to the original one. For f bounded below on $C = \overline{[A,B]_\delta}$ one can always find $N \in \mathbb{N}$ such that $\forall n > N$ we have that f is bounded below on $\overline{[A,B]_{1/n}}$, thus we can choose a (monotonically increasing) sequence $s_n < \inf_{B_{1/n}} f \leq \inf_B f$. As $n \to +\infty$, we have $\inf_{B_{1/n}} f \to \inf_B f$, then for every $\varepsilon > 0$ we can find a number n large enough such that $|s_n - \inf_B f| < \varepsilon$. By Theorem 3.3.1 we can find ξ_n arbitrarily close to $[A,B]$ and $p_n \in \partial f(\xi_n)$ such that they satisfy (3.5), i.e.

$$\inf_B p_n - p_n(a) > s_n - f(a) > \inf_B f - f(a) - \varepsilon,$$

which is similar to the original result in the setting point-set for a more general type of subdifferential, but however, this ξ is not in $[A,B]$ as the limit $\lim \xi_n$ does not exist in general.

In [10] the autors proved another multidirectional inequality. The setting there is point-set for $A = \{a\}$ and B - closed, convex and bounded and the underlying space X is Hilbert. The function f is proper, lower semicontinuous and bounded below on

the set $[a,B]_\delta$ for some $\delta > 0$. The subdifferential is proximal, i.e. $p \in \partial^\pi f(x)$ if $f(x) \neq +\infty$ and $\exists \sigma > 0$ such that for every y in a neighbourhood of x we have

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

Then we can find for every $r < \lim_{\delta \downarrow 0} \inf_{B_\delta} f - f(a)$ and every $\varepsilon > 0$ a point $\xi \in [a,B]_\varepsilon$ and $p \in \partial^\pi f(\xi)$ such that

$$r < \inf_B p - \langle p, a \rangle \text{ and } f(\xi) < \inf_{[a,B]} f + |r| + \varepsilon.$$

It is obvious that Theorem 3.3.1 extends this result for $A = \{a\}$ in multiple directions - underlying space and type of the subdifferential.

We will include here one more comparison – to the result, found in [11], Theorem 3.5. The result there examines point-set setting and is very similar to this in [10], but is also more general – for Banach space with β-smooth equivalent norm and the corresponding β-subdifferential. Again, Theorem 3.3.1 covers the latter result.

Before we present the proof of Theorem 3.3.1, we will stress the fact that our result is of type set-set, thus it is more general than any similar result of point-set type.

3.4 Proof of the new multidirectional inequality

Here we will prove the main result of the chapter. For the proof several propositions will be needed. We will deal mainly with the supremum convolution in what follows and first we state a lemma, very simmilar to Lemma 2.2.6.

Lemma 3.4.1. Let X be a Banach space and $f, g : X \to \mathbb{R} \cup \{+\infty\}$ be functions, $p \in \partial^+(f \Delta g)(x)$ and $y \in X$ is such that

$$f \Delta g(x) - \varepsilon \leq f(x-y) + g(y), \tag{3.6}$$

then $p \in \partial^+_\varepsilon g(y)$.

Proof. By the definition of the sup-convolution for any $z \in X$

$$\begin{aligned} f \Delta g(x+z) &\geq f((x+z)-(y+z)) + g(y+z) \\ &= f(x-y) + g(y+z). \end{aligned}$$

From this, $p \in \partial^+(f \Delta g)(x)$ and (3.6) it follows that

$$\begin{aligned} p(z) &\geq f \Delta g(x+z) - f \Delta g(x) \\ &\geq (f(x-y) + g(y+z)) - (f(x-y) + g(y) + \varepsilon) \\ &\geq g(y+z) - g(y) - \varepsilon. \end{aligned}$$

That is, $p \in \partial^+_\varepsilon g(y)$. □

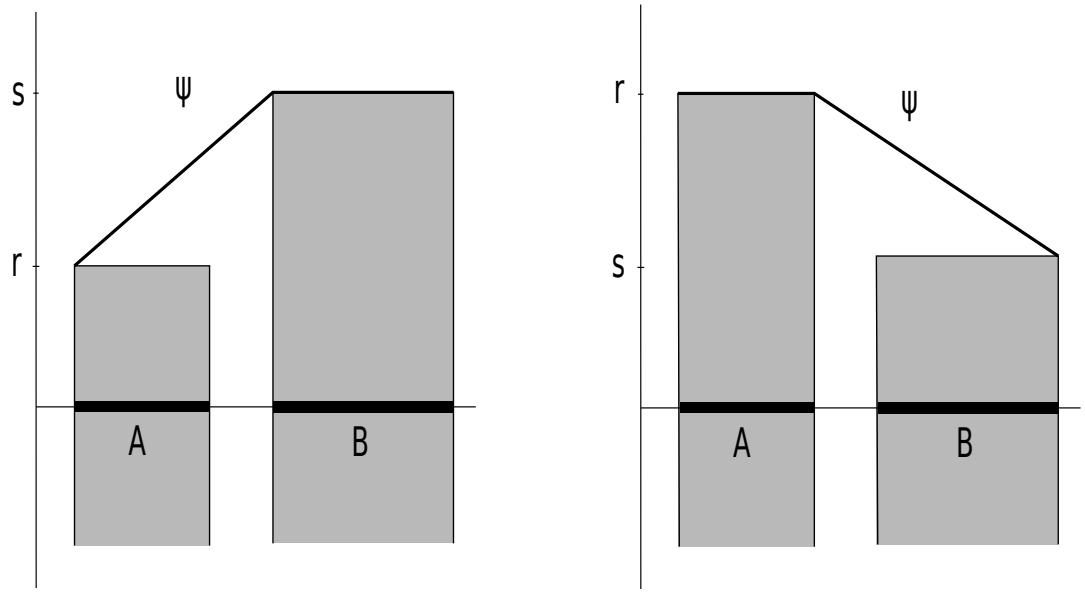

Figure 3.1: *A simple sketch of $\psi(x)$ in the plane – the two possible cases $r < s$ and $s < r$.*

We will denote the hypograph of a function $f : X \to \mathbb{R}$ by

$$\text{hyp } f := \{(x,t) : t \le f(x)\} \subseteq X \times \mathbb{R}.$$

The aim of the next proposition is to provide an auxiliary estimate for a special function ψ.

Proposition 3.4.2. Let A and B be non-empty, convex subsets of the Banach space X. Let $r, s \in \mathbb{R}$ be such that $r \neq s$. We define the concave function ψ (see Figure 3.1.) by its hypograph:

$$\text{hyp } \psi := \text{co}\{A \times [-\infty, r], B \times [-\infty, s]\}, \tag{3.7}$$

where co denotes the convex envelope of the sets. Let $x_0 \in \text{dom } \psi = [A, B]$ be such that $\psi(x_0) \neq s$. Let $p \in \partial_\varepsilon^+ \psi(x_0)$. Then

$$\inf_A p - \inf_B p \le r - s + \frac{r-s}{\psi(x_0) - s}\varepsilon. \tag{3.8}$$

Proof. Let $p \in \partial_\varepsilon^+ \psi(x_0)$. Set $l(x) := \psi(x_0) + p(x - x_0)$. Note that by definition we have $l \ge \psi - \varepsilon$. First we note that as $(x_0, \psi(x_0)) \in \text{hyp}\psi$, we can find points $(u, \bar{r}) \in A \times [-\infty, r]$, $(v, \bar{s}) \in B \times [-\infty, s]$ such that

$$(x_0, \psi(x_0)) = \lambda(u, \bar{r}) + (1 - \lambda)(v, \bar{s}),$$

for some $\lambda \in [0,1]$. Suppose that $\lambda = 0$. Then $(x_0, \psi(x_0)) = (v, \bar{s})$ or $\psi(x_0) = \psi(v) = \bar{s} \leq s$. On the other hand, $\psi(v) \geq s$. So, $\psi(x_0) = s$, a contradiction to $\psi(x_0) \neq s$. Thus $\lambda \in (0,1]$. Furthermore, as

$$x_0 = \lambda u + (1-\lambda)v;\ \psi(x_0) = \lambda \bar{r} + (1-\lambda)\bar{s}$$

and $\bar{r} \leq r$ and $\bar{s} \leq s$, we have that

$$\psi(x_0) \leq \lambda r + (1-\lambda)s$$

and at the same time $(x_0, \lambda r + (1-\lambda)s) = \lambda(u,r) + (1-\lambda)(v,s) \in \mathrm{hyp}\psi$. It follows that

$$\lambda \bar{r} + (1-\lambda)\bar{s} = \lambda r + (1-\lambda)s, \text{ or equivalently}$$

$$\lambda(r - \bar{r}) + (1-\lambda)(s - \bar{s}) = 0.$$

There are two possible cases:

Case 1.

$\lambda \in (0,1)$. In this case we have that $r = \bar{r}$ and $s = \bar{s}$. Then

$$(x_0, \psi(x_0)) = (\lambda u + (1-\lambda)v, \lambda r + (1-\lambda)s) = \lambda(u,r) + (1-\lambda)(v,s).$$

It follows that

$$p(v - x_0) \geq \psi(v) - \psi(x_0) - \varepsilon \geq s - \psi(x_0) - \varepsilon = \lambda(s-r) - \varepsilon.$$

As we have that

$$u - x_0 = -\frac{1-\lambda}{\lambda}(v - x_0),$$

we get

$$p(u - x_0) \leq \frac{1-\lambda}{\lambda}[-\lambda(s-r) + \varepsilon] = (\lambda - 1)(s-r) + \frac{1-\lambda}{\lambda}\varepsilon.$$

Thus, we obtain

$$\begin{aligned} \inf_A l \leq l(u) &= \psi(x_0) + p(u - x_0) \\ &\leq \psi(x_0) + (\lambda - 1)(s-r) + \frac{1-\lambda}{\lambda}\varepsilon = r + \frac{1-\lambda}{\lambda}\varepsilon. \end{aligned}$$

On the other hand, as $l \geq \psi - \varepsilon$ we have

$$\inf_B l \geq \inf_B(\psi - \varepsilon) = \inf_B \psi - \varepsilon \geq s - \varepsilon$$

and after combining the last two inequalities, we obtain

$$\inf_B p - \inf_A p = \inf_B l - \inf_A l \geq s - \varepsilon - r - \frac{1-\lambda}{\lambda}\varepsilon = s - r - \frac{\varepsilon}{\lambda}.$$

Finally, $s - \psi(x_0) = \lambda(s - r)$, so $1/\lambda = (s - r)/(s - \psi(x_0))$ we get (3.8).
Case 2.
$\lambda = 1$. Here $r = \bar{r}$. Then $(x_0, \psi(x_0)) = (u, r)$. We have

$$\inf_A l \leq l(u) = \psi(x_0) + p(u - x_0) = r + p(0) = r,$$

$$\inf_B l \geq \inf_B \psi - \varepsilon \geq s - \varepsilon.$$

Thus we get

$$\inf_B p - \inf_A p = \inf_B l - \inf_A l \geq s - r - \varepsilon.$$

Since $s - \psi(x_0) = s - \psi(u) = s - r$, this is equivalent to (3.8). □

In our proof the function φ_K defined below plays the role of the linear function in the standard proof of the Lagrange Mean Value Theorem. Furthermore, the next proposition will be the main tool for proving our main result – Theorem 3.3.1.

Proposition 3.4.3. Let A and B be non-empty, convex subsets of the Banach space X and let $r, s \in \mathbb{R}$ be such that $r \neq s$. We consider the function ψ as defined in Proposition 3.4.2. Let $K > 0$ and

$$\varphi_K(x) := (-K\|\cdot\|)\Delta\psi(x) = \sup\{\psi(y) - K\|x - y\| : y \in X\}.$$

Then φ_K is K-Lipschitz and concave.
Let $\bar{x}$ be such that there exists $c > 0$ for which the sets

$$U := \{z \in [A, B] : \psi(z) - K\|z - \bar{x}\| > \varphi_K(\bar{x}) - c\}$$

and

$$V := \{z \in [A, B] : |s - \psi(z)| < c\}$$

do not intersect, i.e. $U \cap V = \varnothing$.
If $p \in \partial^+\varphi_K(\bar{x})$, then

$$\inf_B p - \inf_A p \geq s - r.$$

Proof. Since ψ is bounded from above, φ_K is well defined. As a sup-convolution of two concave functions, φ_K is itself concave (see e.g. [12, p.41]).

As $-K\|\cdot\|$ is a K-Lipschitz function, it easily follows that φ_K is also K-Lipschitz. Let $x, y \in X$ and let $(z_n)_1^\infty$ be a maximizing sequence for $\varphi_K(x)$. Then there is a sequence $\varepsilon_n \downarrow 0$ such that

$$\varphi_K(x) > \psi(z_n) - K\|x - z_n\| - \varepsilon_n$$

On the other hand,

$$\varphi_K(y) \leq \psi(z_n) - K\|y - z_n\|$$

After substracting the latter inequalities, we obtain

$$\begin{aligned}\varphi_K(y) - \varphi_K(x) &\le K\|x - z_n\| + \varepsilon_n - K\|y - z_n\| \\ &\le K\|x - z_n - y + z_n\| + \varepsilon_n \\ &= K\|x - y\| + \varepsilon_n.\end{aligned}$$

As $\varepsilon_n \to 0$, we obtain $\varphi_K(y) - \varphi_K(x) \le K\|x - y\|$. The other direction is obtained similarly by considering a maximizing sequence for $\varphi_K(y)$, thus φ_K is K-Lipschitz function.

For the rest of the proof note that from $U \cap V = \emptyset$ it easily follows that for any sequence $(x_n)_1^\infty \subseteq [A, B]$ such that

$$\varphi_K(\bar{x}) = \lim_{n\to\infty}(\psi(x_n) - K\|\bar{x} - x_n\|) \tag{3.9}$$

it holds that

$$|s - \psi(x_n)| \ge c, \ \forall n \text{ large enough.} \tag{3.10}$$

We can assume that the latter is fulfilled for all n.

The function φ_K can be expressed as

$$\varphi_K(x) = \sup\{\psi(y) - K\|x - y\| : y \in [A, B]\},$$

since $\psi = -\infty$ outside $[A, B]$. From (3.9) it follows that we can find $\varepsilon_n \downarrow 0$ with

$$\psi(x_n) - K\|\bar{x} - x_n\| \le \varphi_K(\bar{x}) < \psi(x_n) - K\|\bar{x} - x_n\| + \varepsilon_n.$$

In particular, we have

$$(-K\|\cdot\|)\Delta\psi(\bar{x}) - \varepsilon_n < -K\|\bar{x} - x_n\| + \psi(x_n).$$

Then, for $p \in \partial^+\varphi_K(\bar{x})$ we apply Lemma 3.4.1. It follows that $p \in \partial^+_{\varepsilon_n}\psi(x_n)$. So, from Proposition 3.4.2 and (3.10) we get

$$\inf_A p - \inf_B p \le r - s + \frac{r - s}{\psi(x_n) - s}\varepsilon_n \le r - s + \frac{|s - r|}{c}\varepsilon_n.$$

Since $\varepsilon_n \downarrow 0$, we are done. □

Besides the last statement, we will need several more preparatory claims.

Lemma 3.4.4. Let A, B be non-empty, convex and bounded subsets of the Banach space X. Let $r, s \in \mathbb{R}$, $r \ne s$ and let ψ be constructed as in Proposition 3.4.2. If the sequence $(x_n)_1^\infty$ is such that $\psi(x_n) \to s$ as $n \to \infty$, then $\operatorname{dist}(x_n, B) \to 0$.

Proof. Let $(x_n, \psi(x_n)) = \lambda_n(u_n, r) + (1 - \lambda_n)(v_n, s)$ for some $u_n \in A$, $v_n \in B$ and $\lambda_n \in [0, 1]$. From $r \neq s$ and $\lambda_n r + (1 - \lambda_n)s \to s$ it immediately follows that $\lambda_n \to 0$. But $\operatorname{dist}(x_n, B) \leq \|x_n - v_n\| = \lambda_n \|u_n - v_n\| \to 0$ since A and B are bounded. □

Lemma 3.4.5. Let A and B be non-empty, convex and bounded subsets of the Banach space X. Let $r, s \in \mathbb{R}$, $r \neq s$ and let φ_K be constructed as in Proposition 3.4.3 for some $K > 0$. Then

$$\min\{r, s\} \leq \varphi_K(x) \leq \max\{r, s\}, \ \forall x \in [A, B]. \tag{3.11}$$

Proof. Since

$$\begin{aligned}(A \cup B) \times (-\infty, \min\{r, s\}] &\subseteq A \times (-\infty, r] \cup B \times (-\infty, s] \\ &\subseteq (A \cup B) \times (-\infty, \max\{r, s\}],\end{aligned}$$

after taking convex envelopes we get

$$[A, B] \times (-\infty, \min\{r, s\}] \subseteq \operatorname{hyp}\psi \subseteq [A, B] \times (-\infty, \max\{r, s\}],$$

or, in other words, $\min\{r, s\} \leq \psi \leq \max\{r, s\}$ on $[A, B]$.

Since $\psi \leq \max\{r, s\}$, from the definition of the sup-convolution it readily follows that $\varphi_K \leq \max\{r, s\}$.

On the other hand, if $x \in [A, B]$, then $\varphi_K(x) \geq \psi(x) \geq \min\{r, s\}$. □

Lemma 3.4.6. Let A and B be non-empty and convex subsets of the Banach space X and $\delta > 0$. For the set $C = \overline{[A, B]_\delta}$ we have that its topological boundary $\operatorname{Bnd}(C)$ (see Definition 1.0.15) satisfies

$$\operatorname{Bnd}(C) \subseteq \{x \in C : \operatorname{dist}(x, [A, B]) = \delta\}.$$

Proof. First note that since C is closed, we have that $\operatorname{Bnd}(C) = C \setminus \operatorname{Int}(C)$ (Recall Definition 1.0.14). Observe that

$$\forall x \in X : \ \operatorname{dist}(x, [A, B]) > \delta \implies x \notin C.$$

Indeed, if $\operatorname{dist}(x, [A, B]) > \delta$ for some $x \in X$, then $\exists \varepsilon > 0$ with $\operatorname{dist}(x, [A, B]) > \delta + \varepsilon$. Then for every $y \in B(x, \varepsilon) := \{y \in X : \|y - x\| < \varepsilon\}$ we have

$$\|y - z\| \geq \|z - x\| - \|y - x\| > \|z - x\| - \varepsilon, \ \forall z \in X.$$

This implies that

$$\inf_{[A,B]} \|y - \cdot\| \geq \inf_{[A,B]} \|x - \cdot\| - \varepsilon > \delta + \varepsilon - \varepsilon = \delta,$$

which means that $[A, B]_\delta \cap B(x, \varepsilon) = \varnothing$. Thus $x \notin C$. Next, it is obvious that

$$\forall x \in X : \ \operatorname{dist}(x, [A, B]) < \delta \implies x \in \operatorname{Int}(C).$$

□

Now we are ready to prove the main result of the section – Theorem 3.3.1.

Proof. (*of Theorem* 3.3.1) First we fix $\varepsilon > 0$.

Then we fix s_1 such that $s < s_1 < s + \min\{\varepsilon, \varepsilon\delta\}$, $s_1 < \inf_{B_\delta} f$ and $s_1 \neq r$. Note that

$$|r - s_1| < |r - s| + \varepsilon. \tag{3.12}$$

Also,

$$\frac{\max\{r, s_1\} - \mu}{\delta} < \frac{\max\{r, s\} - \mu}{\delta} + \varepsilon.$$

Now take $\delta_1 \in (0, \delta)$ such that

$$K := \frac{\max\{r, s_1\} - \mu}{\delta_1} < \frac{\max\{r, s\} - \mu}{\delta} + \varepsilon. \tag{3.13}$$

Then let φ_K be the function, constructed in Proposition 3.4.3 with these r, s_1 and K. By Lemma 3.4.6 we have $\mathrm{Bnd}(C) \subseteq \{x \in C : \mathrm{dist}(x, [A, B]) = \delta\}$. If $x \in \mathrm{Bnd}(C)$ then

$$\begin{aligned}
\varphi_K(x) &\leq \sup\{\psi(y) : y \in [A, B]\} + \sup\{-K\,\|y - x\| : y \in [A, B]\} \\
&\leq \max\{r, s_1\} - K \inf\{\|y - x\| : y \in [A, B]\} \\
&= \max\{r, s_1\} - K\delta \\
&< \mu.
\end{aligned}$$

That is,

$$\forall x \in \mathrm{Bnd}(C) \implies \varphi_K(x) < \mu. \tag{3.14}$$

Now we set

$$f_1(x) := \begin{cases} f(x), & x \in C \\ +\infty & x \notin C. \end{cases}$$

Since C is closed, f_1 is lower semicontinuous. Also, $\inf f_1 > \mu$ from (3.1).

From (P2) of Definition 3.2.1 we have that $\partial f_1(x) = \partial f(x)$ for $x \in C \setminus \mathrm{Bnd}(C)$.

Now we consider the function

$$g(x) := f_1(x) - \varphi_K(x)$$

and note that $\mathrm{dom} f_1 = \mathrm{dom} g \subseteq C$. From the above and (3.14) we have that the lower semicontinuous function g (recall that φ_K is K-Lipschitz, due to Proposition 3.4.3) is bounded below and, moreover,

$$\inf\{g(x) : x \in \mathrm{Bnd}(C)\} > 0. \tag{3.15}$$

We claim that

$$\inf g \leq 0. \tag{3.16}$$

Indeed, from (3.2) for any $t > 0$ there is $x \in A$ such that $f_1(x) = f(x) < r + t$. On the other hand, $\varphi_K(x) \geq \psi(x) \geq r$ by the very construction of ψ, see (3.7). Therefore, $g(x) < t$.

Since $-\varphi_K$ is convex and continuous, we can apply (P4') from Proposition 3.2.2 to f_1, $-\varphi_K$ and the feasible subdifferential ∂ to get

$$p_n \in \partial f_1(x_n) \text{ and } q_n \in \partial(-\varphi_K)(y_n) = -\partial^+\varphi_K(y_n)$$

such that

$$\|x_n - y_n\| \to 0,\ (f_1(x_n) - \varphi_K(y_n)) \to \inf(f_1 - \varphi_K),\ \|p_n + q_n\| \to 0. \tag{3.17}$$

Next we will show that for all $n \in \mathbb{N}$ large enough the pair $(\xi, p) = (x_n, p_n)$ satisfies the conclusions of Theorem 3.3.1.

Lemma 3.4.7. $x_n \in \mathrm{Int}(C)$ for all $n \in \mathbb{N}$ large enough and, therefore, $p_n \in \partial f(x_n)$.

Proof. Note that $\inf(f_1 - \varphi_K) = \inf g \leq 0$ by (3.16), so $\lim(f_1(x_n) - \varphi_K(y_n)) \leq 0$ by (3.17).

Assume that there exists subsequence $(x_{n_i})_{i=1}^{\infty} \subseteq \mathrm{Bnd}(C)$.

Then

$$(f_1(x_{n_i}) - \varphi_K(y_{n_i})) = g(x_{n_i}) - (\varphi_K(x_{n_i}) - \varphi_K(y_{n_i})) \geq \inf_{\mathrm{Bnd}(C)} g - K\left\|x_{n_i} - y_{n_i}\right\|,$$

since φ_K is K-Lipschitz (see Proposition 3.4.3). Then, from (3.15) and (3.17) it follows that the latter tends to strictly positive limit, which is a contradiction. □

The estimate (3.4) is easy to check: From (3.17) and the K-Lipschitz continuity of φ_K, which implies $\|q_n\| \leq K$, it follows that $\limsup \|p_n\| \leq K$ and we need only to recall (3.13).

Lemma 3.4.8. For all $n \in \mathbb{N}$ large enough

$$f(x_n) < \inf_{[A,B]} f + |r - s| + \varepsilon.$$

Proof. Let $\nu := |r - s| + \varepsilon - |r - s_1|$. From (3.12) we have $\nu > 0$.

As in the proof of Lemma 3.4.7, we use the Lipschitz continuity of φ_K to see that for all $n \in \mathbb{N}$ large enough

$$f(x_n) - \varphi_K(y_n) < \inf\{f(x) - \varphi_K(x) : x \in [A, B]\} + \nu.$$

But from (3.11) we have $\min\{r, s_1\} \le \varphi_K \le \max\{r, s_1\}$ on $[A, B]$. Therefore,

$$f(x_n) - \max\{r, s_1\} < \inf_{[A,B]} f - \min\{r, s_1\} + \nu.$$

Obviously, $\max\{r, s_1\} - \min\{r, s_1\} = |r - s_1|$ and thus

$$f(x_n) < \inf_{[A,B]} f + |r - s_1| + \nu = \inf_{[A,B]} f + |r - s| + \varepsilon.$$

□

Lemma 3.4.9. For all $n \in \mathbb{N}$ large enough there exist numbers $c_n > 0$ such that the sets

$$U_{n,c_n} := \{z \in [A, B] : \psi(z) - K\|z - y_n\| > \varphi_K(y_n) - c_n\}$$

and

$$V_{c_n} := \{z \in [A, B] : |s_1 - \psi(z)| < c_n\}$$

do not intersect, that is $U_{n,c_n} \cap V_{c_n} = \varnothing$.

Proof. We fix $\bar{\varepsilon} > 0$ such that

$$\bar{\varepsilon} < \inf_C f - \mu,\ \bar{\varepsilon} < \inf_{B_\delta} f - s_1,\ \bar{\varepsilon} < \frac{\delta - \delta_1}{1 + 1/K}. \tag{3.18}$$

Let n be so large that for $\bar{x} = x_n$ and $\bar{y} = y_n$ it is fulfilled that

$$\|\bar{x} - \bar{y}\| < \bar{\varepsilon},\ f_1(\bar{x}) - \varphi_K(\bar{y}) < \bar{\varepsilon}, \tag{3.19}$$

see (3.17).

For this fixed n we assume the contrary, that is, for any positive $c_n > 0$ the sets U_{n,c_n} and V_{c_n}, defined by this c_n, intersect.

For any $m \in \mathbb{N}$ choose $z_m \in U_{n,1/m} \cap V_{1/m}$.

Then the sequence $(z_m) \subseteq [A, B]$ satisfies

$$\varphi_K(\bar{y}) = \lim_{m\to\infty} (\psi(z_m) - K\|z_m - \bar{y}\|),\ \lim_{m\to\infty} \psi(z_m) = s_1. \tag{3.20}$$

It follows that

$$\varphi_K(\bar{y}) = s_1 - K \lim_{m\to\infty} \|z_m - \bar{y}\|.$$

But by (3.19) and (3.1) we have

$$\varphi_K(\bar{y}) > f_1(\bar{x}) - \bar{\varepsilon} > \mu - \bar{\varepsilon}.$$

So,

$$K \lim_{m\to\infty} \|z_m - \bar{y}\| < s_1 - \mu + \bar{\varepsilon} \le K\delta_1 + \bar{\varepsilon}$$

from (3.13). On the other hand, Lemma 3.4.4 implies that $\operatorname{dist}(z_m, B) \to 0$, thus $\operatorname{dist}(\bar{y}, B) \leq \delta_1 + \bar{\varepsilon}/K < \delta - \bar{\varepsilon}$ from (3.18).

From this, (3.19) and the triangle inequality it follows that $\operatorname{dist}(\bar{x}, B) < \delta$. Therefore, $\bar{x} \in B_\delta$ and $f_1(\bar{x}) \geq \inf_{B_\delta} f$. But from (3.20) it is obvious that $\varphi_K(\bar{y}) \leq s_1$ and then $f_1(\bar{x}) - \varphi_K(\bar{y}) \geq \inf_{B_\delta} f - s_1 > \bar{\varepsilon}$ from (3.18). The last one, however, contradicts (3.19). □

Finally, from Lemma 3.4.9 and Proposition 3.4.3 it follows that

$$\inf_B(-q_n) - \inf_A(-q_n) \geq s_1 - r$$

for all $n \in \mathbb{N}$ large enough. Since A and B are both bounded, from (3.17) we get

$$\left|\inf_A p_n - \inf_A(-q_n)\right| \leq \|p_n + q_n\| \sup_{x \in A} \|x\| \to 0,$$

$$\left|\inf_B p_n - \inf_B(-q_n)\right| \leq \|p_n + q_n\| \sup_{x \in B} \|x\| \to 0.$$

From these three above and $s_1 > s$ it follows that

$$\inf_B p_n - \inf_A p_n > s - r$$

for all $n \in \mathbb{N}$ large enough. This completes the proof of Theorem 3.3.1. □

Chapter 4

Conclusion

Main contributions

1. In Chapter 2 – Moreau-Yosida regularization – we showed that the set-valued map, defined as the subdifferential of the Moreau-Yosida regularization of a proper, convex and bounded below function, has a norm, which is continuous in a certain meaningful sense.

Furthermore, if there exists an equivalent Gâteaux differentaible norm in the underlying space, the Moreau-Yosida regularization, defined via the latter norm, is differentiable and the norm of its derivative is continuous. This is naturally true in separable Banach spaces.

2. In Chapter 3 – Multidirectional Inequalities – is proved a new type of multidirectional mean value inequality of type set-set, which generalizes the classical Mean Value theorem. The result involves very general type of subdifferential, defined by axioms.

Furthermore, considered in the point-set setting the result is stronger than the previously known results of this type.

Publications, related to the dissertation:

1. Mihail Hamamdjiev, A Property Of The Moreau-Yosida Regularization, Comptes rendus de l'Académie bulgare des Sciences, Vol 71, No2, 161–168, 2018;

2. Mihail Hamamdjiev and Milen Ivanov, New Multirectional Mean Value Inequality, Journal of Convex Analysis, Volume 25, No. 4, 1279–1290, 2018.

Dissemination of the results, connected to the dissertation:

Some of the results in the dissertation have been presented on several conferences:

1. ”Selections of the Gâteaux subdifferential”, Spring Scientific Session of FMI 2015, 28 March, 2015, https://www-old.fmi.uni-sofia.bg/en/events-1;

2. ”Selections of the Gâteaux subdifferential”, MIDOC 2015, 15.10.2015-18.10.2015, https://math.bas.bg/midoc2015/;

3. ”A generalization of the Clarke-Ledyaev inequality”, Spring Scientific Session of FMI 2017, 25.03.2017, https://www.fmi.uni-sofia.bg/en/node/7095;

4. ”New multidirectional mean value inequality”, MDS2017, 10.07.2017-14.07.2017, https://mds2017.math.bas.bg/;

5. ”On Clarke-Ledyaev inequality”, Spring Scientific Session of FMI 2018, 31.03.2018, https://www.fmi.uni-sofia.bg/en/node/7555.

Declaration of originality

The author declares that the dissertaion contains original results obtained by him or in joint work with his scientific advisor. Results of other people were properly cited.

Bibliography

[1] Deville, Robert and Godefroy, Gilles and Zizler, Václav, Smoothness and renormings in Banach spaces, Pitman Monographs and Surveys in Pure and Applied Mathematics, 64, Longman Scientific & Technical, Harlow; copublished in the United States with John Wiley & Sons, Inc., New York, 1993.

[2] Fabian, Marián and Habala, Petr and Hájek, Petr and Montesinos, Vicente and Zizler, Václav, Banach space theory, CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, The basis for linear and nonlinear analysis, Springer, New York, 2011.

[3] Gilles Godefroy, Renormings of Banach spaces, Handbook of the geometry of Banach spaces. Volume 1, 781–835, 2001.

[4] A. Brondsted, Ralph Tyrrell Rockafellar, On the subdifferentiality of convex functions, In: Proceedings of the American Mathematical Society, 1965.

[5] Milen Ivanov, Sequential representation formulae for G-subdifferential and Clarke subdifferential in smooth Banach spaces, Journal of Convex Analysis, Volume 11, No 1, 179–196, 2004.

[6] Robert R. Phelps, Convex Functions, Monotone Operators and Differentiability, Springer - Verlag, 1989.

[7] Mihail Hamamdjiev and Milen Ivanov, New Multirectional Mean Value Inequality, Journal of Convex Analysis, Volume 25, No. 4, 1279–1290, 2018.

[8] Mihail Hamamdjiev, A Property Of The Moreau-Yosida Regularization, Comptes rendus de l'Académie bulgare des Sciences, Vol 71, No2, 161–168, 2018.

[9] Francis Clarke and Yuri Ledyaev, Mean Value Inequalities, Proceedings of the American Mathematical Society, Volume 122, Number 4, 1994.

[10] Francis Clarke and Yuri Ledyaev, Mean Value Inequalities in Hilbert Space, Transactions of the American Mathematical Society, Volume 344, Number 1, 1994.

[11] Qiji Zhu, Clarke-Ledyaev Mean Value Inequalities in Smooth Banach Spaces, Nonlinear Analysis: Theory, Methods & Applications, Volume 32, Issue 3, 315-324, May 1998.

[12] Jean-Paul Penôt, Calculus Without Derivatives, Springer, 2013.

[13] Lionel Thibault and Nadia Zlateva, Integrability of Subdifferentials of Directionally Lipschitz Functions, Proceedings of the American Mathematical Society, Vol. 133, No. 10, 2939-2948, 2005.

[14] Haim Brezis, Functional Analysis, Sobolev Spaces and Partial Differential Equations, Springer, 2011

[15] John L. Kelly General Topology, D. Van Nostrand Company, Inc., 1955