Recent Developments in Optimality Notions, Scalarizations and Optimality Conditions in Vector Optimization

LEE, Hon Leung

A Thesis Submitted in Partial Fulfilment
of the Requirements for the Degree of
Master of Philosophy
in
Mathematics

The Chinese University of Hong Kong
August 2011

Thesis/Assessment Committee

Professor LEUNG Chi Wai (Chair)

Professor NG Kung Fu (Thesis Supervisor)

Professor FENG Dejun (Committee Member)

Professor LI Chong (External Examiner)

Abstract

This thesis aims at concluding the latest results concerning the necessary conditions for optimality in vector optimization. We consider constrained optimization problems of which the functions concerned can be vector-valued or set-valued. First, we establish a notion of optimality which unifies several known notions. Next, generalized separation theorems are discussed. Through these separation theorems and the tools from variational analysis we formulate the necessary optimality conditions. In the last chapter we study approximate optimality with the related scalarization and variational results.

摘要

本文總結有關向量優化論最優必要條件的最近結果。我們研究向量值或集值函數的約束優化問題。我們先統一數個已知最優概念，然後討論幾個分離定理的推廣，並透過這些分離定理及變分分析闡述最優必要條件。最後一章我們會探討近似最優性及相關的純量化和變分的結果。

ACKNOWLEDGMENTS

I wish to express my gratitude to my supervisor Professor Ng Kung Fu for his excellent instruction and guidance on academics and my life, and valuable suggestions on my thesis.

In addition, I would like to thank my mathematics teachers who helped me acquire the foundation knowledge of mathematics.

Also, I would like to express my appreciation to my friends and schoolfellows, especially Mr. Chow Yat Tin, Ho Tak Chi, Lai Chun Kit, Lau Fu Man, Lee King Leung, Leung Cheung Yu, Ma Man Shun, Kwong Kwok Kun, Ng Chit Yu, and Wong Ting Kam, for their support and encouragement.

Finally, thanks to my parents who prepared delicious meals to me when I was working hard on this thesis.

Contents

Chapter 1

Introduction

The main goal of the thesis is to study the following constrained vector optimization problem:

$$\begin{aligned} &\text{minimize } F(x) \\ &\text{subject to } G_i(x) \cap (-\Lambda_i) \neq \emptyset, \quad i = 1, \cdots, m \\ &\qquad x \in \Omega. \end{aligned} \tag{1.1}$$

where $X, Y, Y_1, \cdots, Y_m$ are real topological vector spaces, $F : X \to 2^Y$, $G_i : X \to 2^{Y_i}$ for $i = 1, \cdots, m$ are set-valued mappings, $\emptyset \neq \Omega \subset X$, $\emptyset \neq \Lambda_i \subset Y_i$ for all $i = 1, \cdots, m$, and a set Θ in Y containing the origin. Here the minimization is with respect to the relation $\leq_\Theta$ which is defined by

$$y_1 \leq_\Theta y_2 \text{ if and only if } y_2 - y_1 \in \Theta. \tag{1.2}$$

This kind of problems is closely related to problems in navigation of robots, stochastic programming, optimal control and welfare economics etc. (see [3, 21, 26] and the references therein)

In the thesis we survey the results in recent papers (mainly [3], [16] and [36]) related to this.

A (single-valued) mapping $f : X \to Y$ can be viewed as a set-valued mapping from X into 2^Y where the value is $\{f(x)\}$ for each x in the domain X. Thus if the above F, G_i (for all i) are single-valued, then the problem (1.1) becomes:

$$\begin{aligned} &\text{minimize } f(x) \\ &\text{subject to } g_i(x) \in -\Lambda_i, \quad i = 1, \cdots, m \\ &\qquad x \in \Omega. \end{aligned} \tag{1.3}$$

The set Θ is used to assign an order to the vector space Y, say, if $\Theta = C$ is a pointed convex cone (see Definition 2.10), then the relation $\leq_C \triangleq \leq_\Theta$ defined in (1.2) is a reflexive, antisymmetric and transitive relation. For example, if we let $Y = \mathbb{R}^n$ and $C = \mathbb{R}^n_+ \triangleq \{(y_1, \cdots, y_n) \in \mathbb{R}^n : y_i \geq 0 \text{ for all } i\}$ (where "$\geq$" is the usual "greater than or equal to" in real numbers), then $(y_1, \cdots, y_n) \leq_{\mathbb{R}^n_+} (y'_1, \cdots, y'_n)$ if and only if $y_i \leq y'_i$ for all i. Some examples of ordering cones in infinite dimensional spaces are also of our interest, say $l^p_+ \triangleq \{\{x_k\} \in l^p : x_k \geq 0 \text{ for all } k\}$ in l^p $(1 \leq p \leq \infty)$ and $L^p_+(\mathbb{R}^n) \triangleq \{[f] \in L^p(\mathbb{R}^n) : \exists \widetilde{f} \in [f] \text{ such that } \widetilde{f} \geq 0 \text{ a.e.}\}$ in $L^p(\mathbb{R}^n)$ $(1 \leq p \leq \infty)$ (where $[f]$ denotes the equivalence class under the equivalence relation on L^p identifying two functions which are equal almost everywhere).

Therefore if we consider $f : X \to \mathbb{R}$ and $g_i : X \to \mathbb{R}$ $(i = 1, \cdots, m)$, $\Theta = \mathbb{R}_+$, $\Lambda_i = \mathbb{R}_+ \triangleq \mathbb{R}^1_+$ for $i = 1, \cdots, p$ and $\Lambda_i = \{0\}$ for $i = p+1, \cdots, m$, then the problem (1.3) becomes:

$$\begin{aligned} &\text{minimize } f(x) \\ &\text{subject to } g_i(x) \leq 0, \quad i = 1, \cdots, p \\ &\qquad g_i(x) = 0, \quad i = p+1, \cdots, m \\ &\qquad x \in \Omega. \end{aligned} \tag{1.4}$$

If X is the Euclidean space $\mathbb{R}^n$ and $\Omega = X$, then the above problem is just the constrained optimization problem under equality and inequality constraints appeared in many standard undergraduate texts on nonlinear programming. How to tell the meaning of a local minimizer for this situation? We consider the feasible set $S \triangleq \{x \in \mathbb{R}^n : g_i(x) \leq 0, \text{ for all } i = 1, \cdots, n \text{ and } x \in \Omega\}$, and say $\bar{x} \in S$ is a local minimizer of the above problem if there exists a neighborhood U of $\bar{x}$ such that $f(\bar{x}) \leq f(x)$ for any $x \in S \cap U$. In other words, if we consider the image set $f(S \cap U) \subset \mathbb{R}$, then $f(\bar{x})$ is the smallest number in the set $f(S \cap U)$. We can do similarly when we study the problem (1.3) (indeed, (1.1) too) , that is to say, first we let $S \triangleq \{x \in X : g_i(x) \in -\Lambda_i \text{ for all } i \text{ and } x \in \Omega\}$ be the feasible set, next we call $\bar{x} \in S$ a local minimizer of the problem (1.3) if there exists a neighborhood U of $\bar{x}$ such that "the element $f(\bar{x})$ is the smallest element in the image set $f(S \cap U) \subset Y$". Nevertheless, what is the meaning of the quoted part in the last sentence? We now raise the first question.

(1) What is meant by the "smallest element" of a nonempty set A in Y?

To define the "smallest element" we require that the space Y has an order (cf. the set Θ given before). We remark that finding out such element is the core problem in vector optimization.

Next we consider the underlying spaces. If all spaces are assumed to be the very general topological vector spaces, then it is difficult to do hard analysis or to make estimates as we have to work without norms or seminorms. On the other hand, as we want to generalize the previous results to a large class of underlying spaces, we usually work on Banach spaces. This is not a must and we will discuss case by case.

Then we look at the set Θ used for ordering elements in Y. The previous examples are all convex cones (see Definition 2.12). Nonetheless the first cone $\mathbb{R}^n_+$ has nonempty interior while the remaining two cones l^p_+ and L^p_+ $(1 \leq p < \infty)$ can be checked to have empty interior. Commonly we can obtain better results if the cone is assumed to have nonempty interior.

The mappings given are also very important. We want to understand what is essential about the set-valued mappings in order to reach the desired results. On the other hand, to formulate the necessary conditions, one way we expect is to introduce derivative-like objects. To conclude, we may ask:

(2) What restrictions should be given to the underlying spaces in order to work out what we want to?

(3) What restrictions should be given to the "ordering set" Θ in order to obtain results?

(4) How to introduce derivative notions to a general function which can be single-valued or set-valued?

The answers of the following questions constitute the core of the thesis.

(5) What are the methods used to solve the problem?

(6) How do our main results improve the previously known results?

We attempt to answer the above questions in the subsequent discussion.

The outline of the thesis is as follows.

In Chapter 2, we give an overview of the basic definitions and results needed for later discussion. They are related to functional analysis, convex analysis, and variational analysis introduced by Mordukhovich and others [25].

In Chapter 3, we introduce an optimality notion describing the minimal points of a set unifying the well-known notions in previous literature (e.g. Pareto minimality and weak minimality). We also describe what is meant by a local minimizer of the problem (1.1) in this chapter where there are no operator constraints (i.e., the only constraint is $x \in \Omega$). These notions were introduced by Bao and Mordukhovich [3].

In Chapter 4, we study the most important theorems that contribute to formulating the first order necessary conditions for the optimality notion given in Chapter 3. Roughly speaking, we call all of them separation theorems. We shall give the proof of a recent separation theorem by Zheng and Ng [36] and see how it improves other previous separation theorems.

In Chapter 5, we first focus on a new property about a set, which is proposed by Bao and Mordukhovich [3], called the local asymptotic closedness in generalizing several well-known notions. Second we will prove the main theorems concerning the necessary conditions for a local minimizer of the problem (1.1). They include the fuzzy, Fritz John-like, Lagrange-like conditions and a condition related to the inverse of the objective set-valued function. Finally we compare the main results to some preceding and related results so as to study the improvements. The results in this chapter are mainly modified from [3]. For one of the main results in [3], we suggest an improved version yet the original method given in [3] does not work for the improved version (see Theorem 5.21).

In Chapter 6, we turn to study an optimality notion which is weaker than Pareto minimality called approximate minimality. We go into the construction of a notion of approximate minimality which is more general than the standard one by Kutateladze [22]. The formulation comes from Gutiérrez, Jiménez and Novo [16]. Then we will find out the necessary conditions for approximate minimality by two approaches. The first way is the scalarization approach by Gutiérrez, Jiménez and Novo. We shall illustrate their results in [16]. They did everything related to single-valued mappings but we extend some of their results to the setting of set-valued mappings. The second way is using variational analysis. Zheng and Ng [36] formulated a fuzzy version of necessary condition for approximate minimality by using their separation theorem (given in Chapter 4), and we will demonstrate their proof.

Chapter 2

Preliminaries

Throughout the whole thesis, we consider vector spaces of which the underlying field must be real numbers, and we denote by $\mathbb{N}$ the set $\{1, 2, 3, \cdots\}$ of natural numbers.

2.1 Functional analysis

Definition 2.1. *Let X be a vector space.*

(1) The space X is called a topological vector space (TVS) if X is a vector space with a Hausdorff topology such that both the addition and scalar multiplication are continuous. Denote by X^* the continuous dual which is the collection of all continuous linear functionals on X. We usually use $\langle x^*, x \rangle$ to denote $x^*(x)$ whenever $x^* \in X^*$ and $x \in X$.

(2) The weak topology on X^* is the smallest topology on X^* such that all evaluation maps $X^* \to \mathbb{R} : x^* \mapsto \langle x^*, x \rangle$ (for all $x \in X$) are continuous.*

(Hausdorff) locally convex spaces (abbreviation: LCS) and normed (vector) spaces are important examples of TVS.

Without otherwise specified, we let X be a TVS throughout this section.

Given any subset A of X, denote by $\text{int}(A)$, $\text{cl}(A)$, $\text{bd}(A)$ the interior, closure, boundary of A respectively. For a sequence $\{x_k\}$ in X, the expression $x_k \xrightarrow{A} \bar{x}$ means that $x_k \to \bar{x}$ and $x_k \in A$. For a sequence $\{t_n\}$ of real numbers, $t_n \downarrow 0$ means $t_n \to 0$ as $n \to \infty$ and $t_n \geq 0$ for all n.

We use the notation w^* to indicate the concepts related to the weak* topology, for example, weak* closure of $A \subset X$ (w^*-cl(A)), weak* convergence ($\xrightarrow{w^*}$) etc.

Assume X is a normed space. The closed ball in X centered at x with radius r is denoted by $B(x,r)$ while the corresponding open ball is denoted by $D(x,r)$. We use $\mathbb{B}$ (resp. $\mathbb{B}^*$) to denote the closed unit ball in X (resp. in X^*) while $\mathbb{D}$ (resp. $\mathbb{D}^*$) denotes the open unit ball in X (resp. in X^*). The distance function from x to A is denoted by $d(x,A)$ or $d_A(x) \triangleq \inf\{\|x-y\| : y \in A\}$, which is Lipschitz in x of modulus 1. The distance between two subsets A_1 and A_2 of X is $d(A_1,A_2) \triangleq \inf\{\|x_1 - x_2\| : x_1 \in A_1, x_2 \in A_2\}$. .

If $X_1, \cdots, X_m$ are normed spaces, then so does the Cartesian product $\prod_{i=1}^{m} X_i$ under the norm $\|(x_1,\cdots,x_m)\| = \|x_1\| + \cdots + \|x_m\|$. Its continuous dual is isomorphic to $\prod_{i=1}^{m} X_i^*$ under the norm $\|(x_1^*,\cdots,x_m^*)\| = \max\{\|x_1^*\|,\cdots,\|x_m^*\|\}$.

Theorem 2.2 (Banach-Alaoglu). *[24, P. 229, Theorem 2.6.18] Let X be a normed space. Then the closed unit ball $\mathbb{B}^*$ in X^* is weak* compact.*

Definition 2.3. *Let X and Y be two normed spaces and S be a nonempty open subset of X. Let $f : S \to Y$ be a mapping and $\bar{x} \in S$ be given. We say f is Fréchet differentiable at $\bar{x}$ if there exists a continuous linear mapping $\nabla f(\bar{x}) : X \to Y$, called the Fréchet derivative of f at $\bar{x}$, with the property*

$$\lim_{\|h\|\to 0} \frac{\|f(\bar{x}+h) - f(x) - \nabla f(\bar{x})h\|}{\|h\|} = 0.$$

Definition 2.4. *A Fréchet differentiable mapping $f : X \to Y$ is strictly differentiable at $\bar{x} \in X$ if*

$$\lim_{\substack{x\to\bar{x}\\ u\to\bar{x}}} \frac{f(x) - f(u) - \nabla f(\bar{x})(x-u)}{\|x-u\|} = 0,$$

where $\nabla f(\bar{x})$ denotes the Fréchet derivative of f.

Notice that any mapping f continuously Fréchet differentiable around $\bar{x}$ is strictly differentiable at $\bar{x}$, but not vice versa. About this point readers can refer to [25, P. 19].

Definition 2.5. *Let X be a vector space, Y be a TVS, S be a nonempty open subset of X, $f : S \to Y$ be a given map and $\bar{x}$ be an element in S. If for all $h \in X$ the limit*

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

exists and if $f'(\bar{x})$ is a continuous linear map from X into Y, then $f'(\bar{x})$ is called the Gâteaux derivative of f at $\bar{x}$ and f is called Gâteaux differentiable at $\bar{x}$.

We mention two important classical separation theorems. Their proofs can be found in [21, P. 74, Theorem 3.16; P. 76, Theorem 3.20].

Theorem 2.6 (Eidelhelt's separation theorem). *Let X be a topological vector space, A and B be two convex subsets of X. Assume that A has nonempty interior. Then $\operatorname{int}(A) \cap B = \emptyset$ if and only if there exist $x^* \in X^* \setminus \{0\}$ and $\alpha \in \mathbb{R}$ such that*

$$\langle x^*, a\rangle \le \alpha \le \langle x^*, b\rangle \text{ for all } a \in A,\ b \in B, \text{ and}$$
$$\langle x^*, a\rangle < \alpha \text{ for all } a \in \operatorname{int}(A).$$

Theorem 2.7 (Strict separation theorem). *Let A and B be nonempty convex subsets of a locally convex space X where A is compact and B is closed. Then $A \cap B = \emptyset$ if and only if there is $x^* \in X^* \setminus \{0\}$ such that*

$$\sup_{a\in A}\langle x^*, a\rangle < \inf_{b\in B}\langle x^*, b\rangle.$$

A special class of Banach spaces called Asplund spaces (introduced by Asplund [1]) enjoys many nice properties that we are interested in.

Definition 2.8. *[29, P. 14, Definitions 1.22] A Banach space X is called an Asplund space provided that every continuous convex function defined on a nonempty open convex subset D of X is Fréchet differentiable at each point of some dense G_δ subset of D.*

For example, any reflexive space is Asplund (see [31, P. 88, Proposition 4.7.14, 4.7.15]). Also, a Banach space is an Asplund space if its dual space is separable (see [29, P. 23, Theorem 2.12]). Therefore c_0, which is non-reflexive, is an Asplund space. From [29, P. 26, Theorem 2.19], a separable Banach space X is an Asplund space if and only if its dual X^* is separable. Hence l^1 is not Asplund. Moreover, the following result is well-known and useful (cf. Theorem 2.2).

Theorem 2.9. *[25, P. 196, second last paragraph] If X is an Asplund space, then the closed unit ball $\mathbb{B}^*$ in X^* is weak* sequentially compact.*

2.2 Convex analysis

Without otherwise specified we let X be a vector space.

Definition 2.10. *A nonempty set $C \subset X$ is called a cone if $\alpha c \in C$ for all $\alpha \geq 0$ and $c \in C$ (Note that a cone must contain the origin). It is pointed if $C \cap (-C) = \{0\}$. It is solid if X is a TVS and $\text{int}(C)$ is nonempty. We call $\{0\}$ the trivial cone. A cone $C \subset Y$ is proper if $\{0\} \neq C \neq Y$.*

The next proposition follows immediately from definitions:

Proposition 2.11. *If $C \subset X$ is a cone, then C is convex if and only if $C+C = C$. Also, a convex cone $C \subset X$ is a vector subspace if and only if $C \setminus (-C) = \emptyset$.*

Definition 2.12.

(1) Given a nonempty set C in X. We define a relation $\leq_C$ by: $x_1 \leq_C x_2$ if and only if $x_2 - x_1 \in C$. We call C a relation set.

(2) We call X an ordered vector space if there exists a convex cone $C \subset X$ inducing a relation "$\leq_C$" on X. We also say X is a vector space ordered by C. The set C is called an ordering cone.

(3) A binary relation "$\sim$" on X is said to be linear if it satisfies

(a) $x + z \sim y + z$ *whenever* $x, y, z \in X$ *and* $x \sim y$.

(b) $\alpha x \sim \alpha y$ *whenever* $x, y \in X$, $\alpha \geq 0$ *and* $x \sim y$.

(4) A relation "$\sim$" on X is called a pre-order if it is reflexive, transitive and linear. If in addition it is antisymmetric, then it is called a partial order.

Next, we suggest a relation between the order and the relation set in X. The following result can be proved easily.

Proposition 2.13.

(1) If $C \subset Y$ is a convex cone (resp. convex pointed cone), then the relation induced from C is a pre-order (resp. partial order) on Y.

(2) Given a pre-order (resp. partial order) "$\leq$" on X. Then the relation set induced from $\leq$, $C \triangleq \{x \in X : 0 \leq x\}$ is a convex cone (resp. convex pointed cone) in Y.

Definition 2.14. *Let $C \subset X$ be a cone. The dual cone of C is $C^* \triangleq \{x^* \in X^* : \langle x^*, x\rangle \geq 0$ for all $x \in C\}$.*

Definition 2.15. *Given a subset A of X. The core (or algebraic interior) of A, written by $\operatorname{cor}(A)$, is defined by*

$$\operatorname{cor}(A) \triangleq \{a \in A : \text{ for all } x \in X, \text{ there exists } \delta > 0 \\ \text{such that } a + \delta' x \in A \text{ for all } 0 \leq \delta' < \delta\}.$$

Proposition 2.16. *[32, P. 4, Theorem 1.1.2] Let X be a TVS and $A \subset X$ be a nonempty convex set. Then we have:*

(1) $\operatorname{int}(A) \subset \operatorname{cor}(A)$.

(2) $\operatorname{cl}(A)$ *is convex.*

(3) If $x \in \operatorname{int}(A)$ and $y \in \operatorname{cl}(A)$, then $\{\alpha x + (1-\alpha)y : \alpha \in (0,1]\} \subset \operatorname{int}(A)$.

(4) If A is solid then

(a) $\text{int}(A) = \text{cor}(A)$ *and* $\text{int}(A)$ *is convex.*

(b) $\text{cl}(A) = \text{cl}(\text{int}(A))$ *and* $\text{int}(A) = \text{int}(\text{cl}(A))$.

(5) If A is a cone, then $\text{int}(A)$ *is also a cone and* $\text{cl}(A) + \text{int}(A) = \text{int}(A)$.

Example 2.17. *The standard ordering cone in $\mathbb{R}$ is $\mathbb{R}_+ \triangleq \{x \in \mathbb{R} : x \geq 0\}$. For $\mathbb{R}^n$, it is $\mathbb{R}^n_+ \triangleq \{(x_1, \cdots, x_n) \in \mathbb{R}^n : x_i \geq 0$ for any $i\}$.*

Definition 2.18. *The affine hull and conic hull of a set $A \subset X$ are*

$$\text{aff}(A) \triangleq \bigcap\{V \subset X : A \subset V, V \textit{ is affine}\}$$

and

$$\text{cone}(A) \triangleq \bigcap\{V \subset X : A \subset C, C \textit{ is a cone}\} = \mathbb{R}_+ \cdot A$$

respectively. The closed affine hull and the closed conic hull of a set $A \subset X$ is $\overline{\text{aff}}(A) = \text{cl}(\text{aff}(A))$ *and* $\overline{\text{cone}}(A) = \text{cl}(\text{cone}(A))$ *respectively.*

Given a function $f : S \to \overline{\mathbb{R}}$ where S is a nonempty set and $\overline{\mathbb{R}}$ indicates the set of extended real numbers $\mathbb{R} \cup \{\pm\infty\}$. The domain of f is $\text{dom}(f) \triangleq \{x \in S : f(x) \in \mathbb{R}\}$. The epigraph of f is $\text{epi}(f) \triangleq \{(x, t) \in S \times \mathbb{R} : f(x) \leq t\}$. We say f is proper when $\text{dom}(f) \neq \emptyset$ and $f(x) > -\infty$ for all $x \in S$.

Definition 2.19. *Let X be a vector space and S be a nonempty convex subset of X. A function $f : S \to \overline{\mathbb{R}}$ is said to be convex if*

$$f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha) f(y)$$

for any $x, y \in S$, $\alpha \in [0, 1]$ with the convention $(+\infty) - (-\infty) = +\infty$, $0 \cdot (+\infty) = +\infty$, $0 \cdot (-\infty) = 0$.

Proposition 2.20. *[32, P. 40, Theorem 2.1.1] Let $f : S \subset X \to \overline{\mathbb{R}}$. The following statements are equivalent:*

(1) f *is convex.*

(2) $\mathrm{dom}(f)$ *is a convex set and* $f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y)$ *for any* $x, y \in \mathrm{dom}(f)$, $\alpha \in [0,1]$.

(3) $\mathrm{epi}(f)$ *is a convex subset of* $S \times \mathbb{R}$.

Definition 2.21. *Let A be a nonempty subset of a vector space X. We denote by $\delta_A : X \to \mathbb{R} \cup \{+\infty\}$ the indicator function of A defined by $\delta_A(x) = 0$ if $x \in A$ and $\delta_A(x) = +\infty$ otherwise.*

Note $\mathrm{dom}(\delta_A) = A$ and $\mathrm{epi}(\delta_A) = A \times \mathbb{R}_+$. By Proposition 2.20, δ_A is a convex function on X if A is convex.

Definition 2.22. *Let X be a vector space and Y be a vector space ordered by a convex cone C. Also let S be a nonempty subset of X. The mapping $f : S \to Y$ is said to be convex (or C-convex) if for any $x, y \in S$, $\alpha \in [0,1]$,*

$$f(\alpha x + (1-\alpha)y) \leq_C \alpha f(x) + (1-\alpha)f(y).$$

For the above f, we denote the epigraph of f by $\mathrm{epi}(f) \triangleq \{(x,y) \in S \times Y : f(x) \leq_C y\}$

Proposition 2.23. *[21, P. 41, Theorem 2.6] Let X be a vector space and Y be a vector space ordered by a convex cone C. Also let S be a nonempty subset of X and $f : S \to Y$ be a map. Then f is convex if and only if $\mathrm{epi}(f)$ is a convex set in $S \times Y$.*

The following variational principle introduced by Ekeland in [9, 10] is renowned.

Theorem 2.24 (Ekeland's variational principle). *[32, P. 30, Corollary 1.4.2] Let (X,d) be a complete metric space and $f : X \to \overline{\mathbb{R}}$ be a bounded below lower semicontinuous proper function. Let also $\varepsilon > 0$ and $x_0 \in \mathrm{dom}(f)$ be such that $f(x_0) \leq \inf_X f + \varepsilon$. Then for every $\lambda > 0$ there exists $x_\lambda \in X$ such that*

$$f(x_\lambda) \leq f(x_0), \quad d(x_\lambda, x_0) \leq \lambda$$

and

$$f(x_\lambda) < f(x) + \varepsilon\lambda^{-1}d(x, x_\lambda) \text{ for all } x \in X \setminus \{x_\lambda\}$$

(Therefore x_λ is a unique global minimizer of the function $f(\cdot) + \varepsilon\lambda^{-1}d(\cdot, x_\lambda)$ on X).

Definition 2.25 (Fenchel subdifferential and normal cone).

(1) *Let $f : X \to \overline{\mathbb{R}}$ be a convex function and $\bar{x} \in \operatorname{dom}(f)$. The (Fenchel) subdifferential of f at $\bar{x}$ is $\partial f(\bar{x}) \triangleq \{x^* \in X^* : \langle x^*, x - \bar{x}\rangle \leq f(x) - f(\bar{x})$ for all $x \in X\}$. Every element of $\partial f(\bar{x})$ is called a subgradient of f at $\bar{x}$.*

(2) *Let A be a nonempty convex subset of X and $\bar{x} \in A$. The normal cone to A at $\bar{x}$ is*

$$N(\bar{x}; A) \triangleq \partial\delta_A(\bar{x}) = \{x^* \in X^* : \langle x^*, x - \bar{x}\rangle \leq 0 \text{ for all } x \in A\}$$

which is a weak closed convex cone.*

Remark 2.26. *If $\bar{x} \in \operatorname{int}(A)$, one can use the fact that $\operatorname{int}(A) = \operatorname{cor}(A)$ to show $N(\bar{x}; A) = \{0\}$, which is the trivial cone, so only the normal cone to A at a boundary point in A can be nontrivial.*

On the other hand, by the definitions we have:

Proposition 2.27. *If C is a convex cone, then $-N(c; C) \subset C^*$ for any $c \in C$. Furthermore $-N(0; C) = C^*$.*

The following estimate of the subdifferential of the norm will be useful:

Proposition 2.28. *[31, P. 79, Proposition 4.6.2] Let X be a normed space and $\bar{x} \in X$. Define $f : X \to \mathbb{R}$ by $f(x) = \|x - \bar{x}\|$ for all $x \in X$. Then for any $x \in X$, $\partial f(x) = \{x^* \in X^* : \|x^*\| = 1$ and $\langle x^*, x - \bar{x}\rangle = \|x - \bar{x}\|\}$ if $x \neq \bar{x}$, and $\partial f(\bar{x}) = \mathbb{B}^*$.*

Proposition 2.29. *[32, P. 85, Theorem 2.4.4(i)] If $f : X \to \mathbb{R}$ is convex and Gâteaux differentiable at $\bar{x} \in X$, then $\partial f(\bar{x}) = \{f'(\bar{x})\}$.*

If the two given functions are convex, then under a mild assumption of continuity, we have the subdifferential sum rule [31, P. 76, Proposition 4.5.1] described as follows.

Theorem 2.30 (Subdifferential sum rule). *Let $\phi_1, \phi_2 : X \to \overline{\mathbb{R}}$ be two proper convex functions. Suppose there exists $\bar{x} \in \mathrm{dom}(\phi_1) \cap \mathrm{int}(\mathrm{dom}(\phi_2))$ such that ϕ_i is continuous at $\bar{x}$ for $i = 1, 2$. Then for each $x \in \mathrm{dom}(\phi_1) \cap \mathrm{dom}(\phi_2)$, one has $\partial(\phi_1 + \phi_2)(x) = \partial(\phi_1)(x) + \partial(\phi_2)(x)$.*

2.3 Relative interiors

Without otherwise specified we assume X to be a TVS. The following definitions of relative interiors can be found in [32, P. 3, line 3; P. 14, last line; P. 15, 12th last line].

Definition 2.31. *Let A be a nonempty subset of X.*

(1) The <u>relative interior</u> of A, denoted by $\mathrm{ri}(A)$, is the interior of A regarded as a subset of $\overline{\mathrm{aff}}(A)$ with the relative topology. In other words, $a \in \mathrm{ri}(A)$ if and only if there exists a neighborhood of a, V in X such that $V \cap \overline{\mathrm{aff}}(A) \subset A$.

(2) The <u>intrinsic relative interior</u> of A is defined by

$$\mathrm{iri}(A) \triangleq \{a \in A : \mathrm{cone}(A - a) \text{ is a subspace of } X\}.$$

(3) The <u>quasi relative interior</u> of A is defined by

$$\mathrm{qri}(A) \triangleq \{a \in A : \overline{\mathrm{cone}}(A - a) \text{ is a subspace of } X\}.$$

Proposition 2.32. *Let A be a nonempty convex subset of X. Then the following statements hold:*

(1) $\mathrm{int}(A) \subset \mathrm{ri}(A)$. *Equality holds if* $\mathrm{int}(A) \neq \emptyset$.

(2) $\mathrm{ri}(A) \subset \mathrm{iri}(A) \subset \mathrm{qri}(A)$. *Equality holds if X is finite dimensional or* $\mathrm{int}(A) \neq \emptyset$ *or* $\mathrm{ri}(A) \neq \emptyset$.

Proof. Clearly $\mathrm{int}(A) \subset \mathrm{ri}(A)$. If $\mathrm{int}(A) \neq \emptyset$ then by Proposition 2.16 (4a) $X = \mathrm{aff}(A)$. Thus $\mathrm{int}(A) = \mathrm{ri}(A)$. (1) holds. For (2), it follows directly that $\mathrm{ri}(A) \subset \mathrm{iri}(A) \subset \mathrm{qri}(A)$. The remaining result is not needed in our subsequent discussion, so we shall not present the proof but refer the reader to [5, Theorem 2.12]. □

Proposition 2.33. *Let A be a nonempty convex subset of X.*

(1) If $X = \mathbb{R}^n$ then $\mathrm{ri}(A) \neq \emptyset$.

(2) If X is a separable Banach space and A is closed, then $\mathrm{qri}(A) \neq \emptyset$.

Proof. (1) follows from [19, P. 103, Theorem 2.1.3] or [20, P. 34, Theorem 2.1.3] (Note: the set $\mathrm{aff}(A) \subset \mathbb{R}^n$ must be closed). (2) is a special case of [32, P. 18, Proposition 1.2.9]. □

Proposition 2.34. *Let A be a nonempty convex subset of X.*

(1) If $x \in \mathrm{ri}(A)$ and $y \in \mathrm{cl}(A)$, then $\{\alpha x + (1-\alpha)y : \alpha \in (0,1]\} \subset \mathrm{ri}(A)$.

(2) The set $\mathrm{ri}(A)$ *is convex and, when nonempty, is dense in* $\mathrm{cl}(A)$.

(3) If A is a cone, then $\mathrm{ri}(A)$ *is also a cone and* $\mathrm{cl}(A) + \mathrm{ri}(A) = \mathrm{ri}(A)$.

Proof. For (1), translating the set A such that it contains the origin and then applying Proposition 2.16 (3) with $X = \overline{\mathrm{aff}}(A)$, we obtain the result. (2) and (3) are consequences of (1). □

Proposition 2.35. *Let A be a nonempty convex subset of X.*

(1) If $x \in \mathrm{iri}(A)$ (resp. $\mathrm{qri}(A)$) and $y \in A$, then $\{\alpha x + (1-\alpha)y : \alpha \in (0,1]\} \subset \mathrm{iri}(A)$ *(resp.* $\mathrm{qri}(A)$*).*

(2) *The sets* $\mathrm{iri}(A)$ *and* $\mathrm{qri}(A)$ *are convex and, when nonempty, are dense in* $\mathrm{cl}(A)$.

(3) *If* A *is a cone, then* $\mathrm{iri}(A)$ *and* $\mathrm{qri}(A)$ *are also cones and* $A + \mathrm{iri}(A) = \mathrm{iri}(A)$ *and* $A + \mathrm{qri}(A) = \mathrm{qri}(A)$.

Proof. (1) follows from [32, P. 3, 12th last line (viii); P. 15, Proposition 1.2.7]. (2) and (3) are consequences of (1). □

2.4 Multifunctions

Let X and Y be nonempty sets. A mapping $F : X \to 2^Y$ is called a multifunction. Usually we denote it by $F : X \rightrightarrows Y$. The domain of F is $\mathrm{Dom}(F) \triangleq \{x \in X : F(x) \neq \emptyset\}$. The range of F is $F(X) \triangleq \bigcup_{x \in X} F(x)$ which is a subset of Y. The (direct) image of $A \subset X$ under F is $F(A) \triangleq \bigcup_{x \in A} F(x)$. The inverse image of $B \subset Y$ under F is $F^{-1}(B) = \{x \in X : F(x) \cap B \neq \emptyset\}$. The graph of F is $\mathrm{gph}(F) \triangleq \{(x, y) \in X \times Y : y \in F(x)\}$. The inverse of the multifunction F is a multifunction $F^{-1} : Y \rightrightarrows X$ given by $F^{-1}(y) = \{x \in X : y \in F(x)\}$ for all $y \in Y$. The closure of the multifunction F is a multifunction $\mathrm{cl}(F) : X \rightrightarrows Y$ given by $\mathrm{cl}(F)(x) = \{y \in Y : (x, y) \in \mathrm{cl}(\mathrm{gph}(F))\}$. Therefore we have $\mathrm{gph}(\mathrm{cl}(F)) = \mathrm{cl}(\mathrm{gph}(F))$. If Y is a vector space ordered by a convex cone C, we define the epigraph of F by $\mathrm{epi}(F) \triangleq \{(x, y) \in X \times Y : y \in F(x) + C\}$. Given two multifunctions $F : X \rightrightarrows Y$ and $G : Y \rightrightarrows Z$, the composition of G and F is the multifunction $G \circ F : X \rightrightarrows Z$ given by $(G \circ F)(x) \triangleq \bigcup_{y \in F(x)} G(y)$.

Remark 2.36. *Any single-valued function* $f : X \to Y$ *can be viewed as a multifunction* $\widetilde{f} : X \rightrightarrows Y$ *by* $\widetilde{f}(x) = \{f(x)\}$ *for all* $x \in X$. *We shall use back the notation* f *to denote* $\widetilde{f}$. *Under this identification,* $\mathrm{Dom}(f) = X$, $\mathrm{Im}(f)$ *is the image (or range) of* f, $\mathrm{gph}(f)$ *is the graph of* f *and* $f^{-1}(y)$ *is the inverse image of* $\{y\}$ *under* f. *We should keep in mind that many properties on multifunctions make sense on single-valued functions under this identification.*

2.5 Variational analysis

In this section, we assume all spaces concerned are Banach spaces. Most of the definitions and basic results come from [25, Chapter 1-3].

Given a multifunction $F : X \rightrightarrows X^*$, the sequential Painleve-Kuratowski upper/outer limit with respect to the norm topology of X and the weak* topology of X^* is

$$\operatorname*{Limsup}_{x\to\bar{x}} F(x) \triangleq \{x^* \in X^* : \text{there exist sequences } x_k \to \bar{x} \text{ and } x_k^* \xrightarrow{w^*} x^* \text{ with } x_k^* \in F(x_k) \text{ for all } k \in \mathbb{N}\}.$$

Definition 2.37. *Let Ω be a subset of X and $\bar{x} \in \Omega$. Define*

(1) the ε-normal ($\varepsilon \geq 0$) to Ω at $\bar{x}$ by

$$\hat{N}_\varepsilon(\bar{x};\Omega) \triangleq \left\{ x^* \in X^* : \limsup_{x \xrightarrow{\Omega} \bar{x}} \frac{\langle x^*, x-\bar{x}\rangle}{\|x-\bar{x}\|} \leq \varepsilon \right\},$$

(2) the Fréchet normal cone to Ω at $\bar{x}$ by $\widehat{N}(\bar{x};\Omega) \triangleq \hat{N}_0(\bar{x};\Omega)$, and

(3) the Mordukhovich normal cone to Ω at $\bar{x}$ by

$$N(\bar{x};\Omega) \triangleq \operatorname*{Limsup}_{\substack{x \xrightarrow{\Omega} \bar{x} \\ \varepsilon > 0}} \hat{N}_\varepsilon(\bar{x};\Omega).$$

Definition 2.38. *A set $\Omega \subset X$ is said to be locally closed at $\bar{x} \in \Omega$ if there exists a neighborhood U of $\bar{x}$ such that $\Omega \cap U$ is closed.*

Remark 2.39. *If $\Omega \subset X$ is locally closed at $\bar{x}$, then there exists an open neighborhood of $\bar{x}$ such that* $\mathrm{cl}(\Omega) \cap U \subset \Omega$.

Proposition 2.40. *Let $\Omega \subset X$ and $\bar{x} \in \Omega$. The following statements hold:*

(1) Both $\hat{N}(\bar{x};\Omega)$ and $N(\bar{x};\Omega)$ are cones.

(2) $\hat{N}(\bar{x};\Omega) \subset N(\bar{x};\Omega)$.

(3) For any $c \in X$*, we have:* $\hat{N}(\bar{x};\Omega) = \hat{N}(\bar{x}+c;\Omega+c)$ *and* $N(\bar{x};\Omega) = N(\bar{x}+c;\Omega+c)$*;* $\hat{N}(-\bar{x};-\Omega) = -\hat{N}(\bar{x};\Omega)$ *and* $N(-\bar{x};-\Omega) = -N(\bar{x};\Omega)$.

(4) For any $\varepsilon \geq 0$, $\hat{N}_\varepsilon(\bar{x};\mathrm{cl}(\Omega)) = \hat{N}_\varepsilon(\bar{x};\Omega)$.

(5) $N(\bar{x};\Omega) \subset N(\bar{x};\mathrm{cl}(\Omega))$ *and equality holds when* Ω *is locally closed at* $\bar{x}$.

(6) If X *is an Asplund space, then* $N(\bar{x};\Omega) = \operatorname*{Limsup}_{x\to\bar{x}} \widehat{N}(x;\Omega)$.

Proof. The checking of (1) to (5) is routine. For (6), the inclusion "$\supset$" is clear. For the inclusion "$\subset$", by [25, P. 221, Theorem 2.35] we know $N(\bar{x};\mathrm{cl}(\Omega)) = \operatorname*{Limsup}_{x\to\bar{x}} \widehat{N}(x;\mathrm{cl}(\Omega))$. The desired conclusion follows by using (4) and (5). □

Proposition 2.41. *[25, P.6, Proposition 1.2] Given a point* $\bar{x} = (\bar{x}_1,\bar{x}_2) \in \Omega_1 \times \Omega_2 \subset X_1 \times X_2$*. Then* $\hat{N}(\bar{x};\Omega_1\times\Omega_2) = \hat{N}(\bar{x}_1;\Omega_1)\times\hat{N}(\bar{x}_2;\Omega_2)$ *and* $N(\bar{x};\Omega_1\times\Omega_2) = N(\bar{x}_1;\Omega_1)\times N(\bar{x}_2;\Omega_2)$.

The following result [25, P. 7, Proposition 1.5] shows that the normal cones $N(\bar{x};\Omega)$ and $\hat{N}(\bar{x};\Omega)$ coincide with the normal cone defined in Definition 2.25 (2) when Ω is convex.

Proposition 2.42. *Let* U *be a neighborhood of* $\bar{x} \in \Omega \subset X$ *such that* $\Omega \cap U$ *is convex. Then* $N(\bar{x};\Omega) = \hat{N}(\bar{x};\Omega) = \{x^* \in X^* : \langle x^*, x-\bar{x}\rangle \leq 0$ *for all* $x \in \Omega\cap U\}$.

Next we introduce two criteria concerning when strong convergence is implied from weak* convergence.

Definition 2.43. *(SNC and PSNC property)*

(1) A set $\Omega \subset X$ *is said to be* *sequentially normally compact* *(SNC) at* $\bar{x} \in \Omega$ *if for any sequence* $\{(\varepsilon_k, x_k, x_k^*)\} \subset [0,\infty)\times X\times X^*$ *such that* $\varepsilon_k \downarrow 0$, $x_k \xrightarrow{\Omega} \bar{x}$, $x_k^* \in \hat{N}_{\varepsilon_k}(x_k;\Omega)$ *and* $x_k^* \xrightarrow{w^*} 0$ *one has* $\|x_k^*\| \to 0$.

(2) Given $\Omega \subset X \triangleq \prod_{j=1}^{m} X_j$, $\bar{x} \in \Omega$ *and* $J \subset \{1,\cdots,m\}$*. Then* Ω *is said to be* *partially sequentially normally compact* *(PSNC) at* $\bar{x}$ *with respect to* $\{X_j :$

$j \in J\}$ (or $\prod_{j\in J} X_j$ or just J) if for any sequence $\{(\varepsilon_k, x_k, x_{1k}^*, \cdots, x_{mk}^*)\} \subset [0,\infty) \times X \times \prod_{j=1}^{m} X_j^*$ such that $\varepsilon_k \downarrow 0$, $x_k \xrightarrow{\Omega} \bar{x}$, $(x_{1k}^*, \cdots, x_{mk}^*) \in \hat{N}_{\varepsilon_k}(x_k;\Omega)$, $x_{jk}^* \xrightarrow{w^*} 0$ for $j \in J$ and $\|x_{jk}^*\| \to 0$ for $j \in \{1,\cdots,m\} \setminus J$ one has $\|x_{jk}^*\| \to 0$ for $j \in J$.

(3) *A multifunction $F : X \rightrightarrows Y$ is said to be SNC at $(\bar{x},\bar{y}) \in \mathrm{gph}(F)$ if $\mathrm{gph}(F)$ is SNC at $(\bar{x},\bar{y})$.*

(4) *A multifunction $F : X \rightrightarrows Y$ is said to be PSNC at $(\bar{x},\bar{y}) \in \mathrm{gph}(F)$ if $\mathrm{gph}(F)$ is PSNC at $(\bar{x},\bar{y})$ with respect to X.*

Remark 2.44. *If X is an Asplund space, then by Proposition 2.40 (6), Definition 2.43 (1) and (2) can be modified as follows: keeping the same statements except removing $\varepsilon_k \downarrow 0$ and replacing $\hat{N}_{\varepsilon_k}(x_k;\Omega)$ by $\hat{N}(x_k;\Omega)$.*

Proposition 2.45. *Suppose $\Omega \subset X$ is locally closed at $\bar{x} \in \Omega$. We have: if Ω is SNC at $\bar{x}$, then $\mathrm{cl}(\Omega)$ is SNC at $\bar{x}$.*

Proof. Remark 2.39 tells if $x_k \xrightarrow{\mathrm{cl}(\Omega)} \bar{x}$, then $x_k \in \Omega$ for large k. Also invoking Proposition 2.40 (4) with $\varepsilon \geq 0$, we get the conclusion. □

The next theorem [25, P. 27, Theorem 1.21; P. 31, Proposition 1.25, Theorem 1.26] gives some necessary and sufficient conditions for the SNC of a set.

Theorem 2.46. *Given a convex set $\Omega \subset X$. We have:*

(1) *If $\mathrm{ri}(\Omega) \neq \emptyset$, then Ω is SNC at every $\bar{x} \in \Omega$ if and only if $\mathrm{codim}(\overline{\mathrm{aff}}(\Omega)) < \infty$.*

(2) *If $\mathrm{int}(\Omega) \neq \emptyset$, then Ω is SNC at any $\bar{x} \in \Omega$.*

We now introduce a nice type of mappings which are widely studied in literature called Lipschitz-like mappings.

Definition 2.47. *Let $F : X \rightrightarrows Y$ and $(\bar{x}, \bar{y}) \in \operatorname{gph}(F)$. We say that F is Lipschitz-like (or pseudo-Lipschitzian) at $(\bar{x}, \bar{y})$ with modulus $l \geq 0$ if there are neighborhoods U of $\bar{x}$ and V of $\bar{y}$ such that*

$$F(x) \cap V \subset F(u) + l\|x - u\|\mathbb{B} \text{ for all } x, u \in U.$$

Lipschitz-like mappings enjoy the PSNC property. We describe this, together with other situations, precisely as follows.

Theorem 2.48. *Let $F : X \rightrightarrows Y$ be a multifunction and $(\bar{x}, \bar{y}) \in \operatorname{gph}(F)$. Then the following statements hold:*

(1) If F is SNC at $(\bar{x}, \bar{y})$, then it is PSNC at this point.

(2) F is PSNC at $(\bar{x}, \bar{y})$ if X is finite-dimensional.

(3) If Y is finite-dimensional, then F is SNC at $(\bar{x}, \bar{y})$ if and only if it is PSNC at this point.

(4) If F is Lipschitz-like at $(\bar{x}, \bar{y})$ then it is PSNC at this point.

(5) If $F = f$ is single valued and locally Lipschitz at $\bar{x}$, then f is PSNC at $(\bar{x}, f(\bar{x}))$. Moreover it is SNC at this point if Y is finite-dimensional. If in addition f is strictly differentiable at $\bar{x}$ with the surjective derivative $\nabla f(\bar{x})$, then $f^{-1} = F^{-1} : Y \rightrightarrows X$ is PSNC at $(f(\bar{x}), \bar{x})$.

Proof. (1) to (3) follow directly from definitions and (4) and (5) follow from [25, P. 76, Proposition 1.68, Corollary 1.69]. □

We employ coderivatives instead of the more natural graphical derivatives (see [30, P. 324, Definition 8.33] or [21]) on the construction of derivative-like objects for multifunctions.

Definition 2.49 (coderivatives). *Let $F : X \rightrightarrows Y$ and $(\bar{x}, \bar{y}) \in \operatorname{gph}(F)$.*

(1) The <u>Fréchet coderivative</u> of F at $(\bar{x},\bar{y})$ is a multifunction $\hat{D}^*F(\bar{x},\bar{y}) : Y^* \rightrightarrows X^*$ defined by: for any $y^* \in Y^*$,

$$\hat{D}^*F(\bar{x},\bar{y})(y^*) \triangleq \{x^* \in X^* : (x^*, -y^*) \in \hat{N}((\bar{x},\bar{y}); \mathrm{gph}(F))\}.$$

(2) The <u>Mordukhovich coderivative</u> of F at $(\bar{x},\bar{y})$ is a multifunction $D^*F(\bar{x},\bar{y}) : Y^* \rightrightarrows X^*$ defined by: for any $y^* \in Y^*$,

$$\begin{aligned} D^*F(\bar{x},\bar{y})(y^*) &\triangleq \{x^* \in X^* : (x^*, -y^*) \in N((\bar{x},\bar{y}); \mathrm{gph}(F))\} \\ &= \{x^* \in X^* : \text{ there exist sequences } \varepsilon_k \downarrow 0, (x_k, y_k) \to (\bar{x},\bar{y}) \\ &\quad \text{and } (x_k^*, y_k^*) \xrightarrow{w^*} (x^*, y^*) \text{ with } (x_k, y_k) \in \mathrm{gph}(F) \text{ and} \\ &\quad (x_k^*, -y_k^*) \in \hat{N}_{\varepsilon_k}((x_k, y_k); \mathrm{gph}(F)) \text{ for all } k\}. \end{aligned}$$

(3) The <u>mixed coderivative</u> of F at $(\bar{x},\bar{y})$ is a multifunction $D_M^*F(\bar{x},\bar{y}) : Y^* \rightrightarrows X^*$ defined by: for any $y^* \in Y^*$,

$$\begin{aligned} D_M^*F(\bar{x},\bar{y})(y^*) \triangleq \{x^* \in X^* : &\text{ there exist sequences } \varepsilon_k \downarrow 0, (x_k, y_k) \to (\bar{x},\bar{y}), \\ &x_k^* \xrightarrow{w^*} x^* \text{ and } y_k^* \to y^* \text{ with } (x_k, y_k) \in \mathrm{gph}(F) \text{ and} \\ &(x_k^*, -y_k^*) \in \hat{N}_{\varepsilon_k}((x_k, y_k); \mathrm{gph}(F)) \text{ for all } k\}. \end{aligned}$$

Remark 2.50. *Clearly $\hat{D}^*F(\bar{x},\bar{y})(y^*) \subset D_M^*F(\bar{x},\bar{y})(y^*) \subset D^*F(\bar{x},\bar{y})(y^*)$ for any $y^* \in Y^*$. If X and Y are Asplund spaces, then by Proposition 2.40 (6), the definitions of Mordukhovich coderivative and mixed coderivative can be simplified as: keeping the same statements but removing $\varepsilon_k \downarrow 0$ and replacing $\hat{N}_{\varepsilon_k}((x_k, y_k); \mathrm{gph}(F))$ by $\hat{N}((x_k, y_k); \mathrm{gph}(F))$. Moreover, if $f : X \to Y$ is single-valued, then we write $\hat{D}^*f(\bar{x})$ (resp. $D^*f(\bar{x})$, $D_M^*f(\bar{x})$) instead of $\hat{D}^*f(\bar{x}, f(\bar{x}))$ (resp. $D^*f(\bar{x}, f(\bar{x}))$, $D_M^*f(\bar{x}, f(\bar{x}))$).*

Next we show some special situations which the coderivatives can be easily computed. All types of coderivatives coincide when the graph of the multifunction is convex. This result just follows from Proposition 2.42 and the first sentence of Remark 2.50.

Proposition 2.51. *[25, P. 45, Proposition 1.37] Let $F : X \rightrightarrows Y$ be a multifunction with convex graph. Then for any $(\bar{x}, \bar{y}) \in \mathrm{gph}(F)$ and $y^* \in Y^*$, we have*

$$\begin{aligned}\hat{D}^*F(\bar{x},\bar{y})(y^*) &= D^*_M F(\bar{x},\bar{y})(y^*) = D^*F(\bar{x},\bar{y})(y^*)\\ &= \left\{x^* \in X^* : \langle x^*, \bar{x}\rangle - \langle y^*, \bar{y}\rangle \le \max_{(x,y)\in \mathrm{gph}(F)} [\langle x^*, x\rangle - \langle y^*, y\rangle]\right\}.\end{aligned}$$

Theorem 2.52 (coderivatives of differentiable mappings). *[25, P. 45, Theorem 1.38]*

(1) Let $f : X \to Y$ be Fréchet differentiable at $\bar{x}$, then $\hat{D}^*f(\bar{x})(y^*) = \{\nabla f(\bar{x})^* y^*\}$ for all $y^* \in Y^*$.

(2) If f is strictly differentiable at $\bar{x} \in X$, then $D^*f(\bar{x})(y^*) = D^*_M f(\bar{x})(y^*) = \{\nabla f(\bar{x})^* y^*\}$ for all $y^* \in Y^*$.

Theorem 2.53. *[25, P. 53, Theorem 1.44] If $F : X \rightrightarrows Y$ is Lipschitz-like at $(\bar{x}, \bar{y}) \in \mathrm{gph}(F)$, then $D^*_M F(\bar{x}, \bar{y})(0) = \{0\}$.*

Definition 2.54. *For a function $f : X \to \overline{\mathbb{R}}$ and $\bar{x} \in \mathrm{dom}(f)$, we define*

$$\hat{\partial} f(\bar{x}) \triangleq \{x^* \in X^* : (x^*, -1) \in \hat{N}((\bar{x}, \bar{y}); \mathrm{epi}(f))\} \text{ and}$$
$$\partial f(\bar{x}) \triangleq \{x^* \in X^* : (x^*, -1) \in N((\bar{x}, \bar{y}); \mathrm{epi}(f))\}.$$

Remark 2.55. *If $f : X \to \overline{\mathbb{R}}$ is convex and $\bar{x} \in \mathrm{dom}(f)$, then owing to the convexity of the set $\mathrm{epi}(f)$ (see Proposition 2.20), it follows from Proposition 2.42 that two types of subdifferential of f at $\bar{x}$ defined in Definition 2.54 coincide with the Fenchel subdifferential given in Definition 2.25 (1).*

The mixed/Mordukhovich coderivative and Mordukhovich subdifferential have the following good relation [25, P. 93, Theorem 1.90; P. 291, Theorem 3.28]:

Theorem 2.56.

(1) Let $f : X \to Y$ be continuous at $\bar{x}$. Then $\partial(y^* \circ f)(\bar{x}) \subset D^*_M f(\bar{x})(y^*)$ for all $y^* \in Y^*$. Moreover, the equality holds for all $y^* \in Y^*$ if in addition f is Lipschitz at $\bar{x}$.

(2) Suppose the mapping $f : X \to Y$ between an Asplund space X and a Banach space Y is strictly differentiable at $\bar{x}$. Then $\partial(y^* \circ f)(\bar{x}) = D^* f(\bar{x})(y^*) = D^*_M f(\bar{x})(y^*) = \{\nabla f(\bar{x})^* y^*\}$ for all $y^* \in Y^*$.

The following fuzzy subdifferential sum rule can be found in [6, Theorem 6.1.11]. Originally it was proved by Fabian (see [11, Theorem 2]) by using a separable reduction argument.

Theorem 2.57. *Let X be an Asplund space, $\phi_1 : X \to \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous function and $\phi_2 : X \to \mathbb{R}$ be a Lipschitz function. Suppose x is a minimizer of $\phi_1 + \phi_2$. Then for any $\varepsilon > 0$, there exist $x_1, x_2 \in D(x, \varepsilon)$ such that $|\phi_i(x_i) - \phi_i(x)| < \varepsilon$ $(i = 1, 2)$ and $0 \in \hat{\partial}\phi_1(x_1) + \hat{\partial}\phi_2(x_2) + \varepsilon\mathbb{D}^*$.*

Chapter 3

A unified notion of optimality

In this chapter we study the following constrained vector optimization problem:

$$\begin{aligned} &\text{minimize } F(x) \\ &\text{subject to } x \in \Omega. \end{aligned} \tag{3.1}$$

where $F : X \rightrightarrows Y$ is a multifunction between two TVSs and $\Omega \subset X$ is a nonempty constraint set. The set Ω can be a geometric set or described by operators. What we do first is to introduce what is meant by "minimize". We are going to study an optimality notion introduced in [3] which unifies several well-known notions.

First we need to recall some basic optimality notions.

3.1 Basic notions of minimality

In this section, we assume Y is a TVS ordered by a convex cone $C \subset Y$. Also, A is always a nonempty subset of Y. The most basic notion first introduced by the economists Edgeworth and Pareto is defined as follows:

Definition 3.1. *Let $\bar{y} \in A$. Then $\bar{y}$ is a Pareto minimal point of A with respect* *to C if $A \cap (\bar{y} - C) = \{\bar{y}\}$. Note that*

$$\begin{aligned} & A \cap (\bar{y} - C) = \{\bar{y}\} \\ \Leftrightarrow \quad & (\bar{y} - A) \cap (C \setminus \{0\}) = \emptyset \\ \Leftrightarrow \quad & y = \bar{y} \text{ whenever } y \in A \text{ and } y \leq_C \bar{y} \end{aligned}$$

Denote by $\mathrm{Min}(A, C)$ *the collection of all Pareto minimal points of A with respect to C.*

The words "with respect to C" in the above definition may be omitted in later discussion if it is understood (similar for other optimality notions).

Definition 3.2. *Let $\bar{y} \in A$. Then $\bar{y}$ is an ideal minimal point of A if $A \subset \bar{y} + C$. Note that*

$$\begin{aligned} & A \subset \bar{y} + C \\ \Leftrightarrow \quad & A \cap (\bar{y} - (Y \setminus (-C))) = \emptyset \\ \Leftrightarrow \quad & A \cap (\bar{y} - (Y \setminus (-C)) \cup \{0\}) = \{\bar{y}\} \\ \Leftrightarrow \quad & \bar{y} \leq_C y \text{ for all } y \in A \end{aligned}$$

Denote by $\mathrm{IMin}(A, C)$ *the collection of all ideal minimal points of A.*

It is clear that $\mathrm{IMin}(A, C) \subset \mathrm{Min}(A, C)$ if in addition C is pointed.

Consider a function $f : X \to Y$ where X is a TVS and Y is as before. Also let $S \subset X$ be a nonempty set.

Definition 3.3. *An element $\bar{x} \in S$ is called a (global) Pareto minimizer of f on S with respect to C if $f(\bar{x}) \in \mathrm{Min}(f(S), C)$. Denote by $\mathrm{Min}(f, S, C)$ the collection of these minimizers. We call $\bar{x} \in S$ to be a local Pareto minimizer of f with respect to C if $\bar{x} \in \mathrm{Min}(f, S \cap U, C)$ for some neighborhood U of $\bar{x}$. The notions of (local and global) ideal minimizers of f are similarly defined, and we omit them here.*

Remark 3.4.

(1) *We have illustrated that, once we defined the meaning of "Pareto minimal point" of a set as in Definition 3.1, we can define what is meant by "local Pareto minimizers" of a function* $f : X \to Y$ *(Definition 3.3). In fact Definition 3.1 can be viewed as a special case of Definition 3.3 by letting* $X = Y$ *and* f *be the identity map.*

(2) *Given a multifunction* $F : X \rightrightarrows Y$ *and a set* $S \subset X$*, we can also define the meaning of "*$(\bar{x}, \bar{y})$ *is a (global) Pareto minimizer for a multifunction* F*" by replacing the reference set* $f(S) \subset Y$ *in Definition 3.3 by* $F(S) \subset Y$*. The collection of such minimizers is denoted by* $\mathrm{Min}(F, S, C)$ *(a subset of* $X \times Y$*). It is similar for the case of local Pareto minimizers. Using the identification described in Remark 2.36, we see that Definition 3.3 is a special case of this. Usually this relation and the relation explained in (1) still apply to other optimality notions.*

Definition 3.5. *Assume that* $\mathrm{int}(C) \neq \emptyset$ *(resp.* $\mathrm{ri}(C) \neq \emptyset$, $\mathrm{iri}(C) \neq \emptyset$, $\mathrm{qri}(C) \neq \emptyset$*). We say that* $\bar{y} \in A$ *is a* <u>*weak*</u> *(resp.* <u>*relative*</u>*,* <u>*intrinsic*</u>*,* <u>*quasi*</u>*)* <u>*minimal point*</u> *if* $A \cap (\bar{y} - \mathrm{int}(C)) = \emptyset$ *(resp.* $A \cap (\bar{y} - \mathrm{ri}(C)) = \emptyset$, $A \cap (\bar{y} - \mathrm{iri}(C)) = \emptyset$, $A \cap (\bar{y} - \mathrm{qri}(C)) = \emptyset$ *). Denote the collection of such minimal points by* $\mathrm{WMin}(A, C)$ *(resp.* $\mathrm{RMin}(A, C)$, $\mathrm{IRMin}(A, C)$, $\mathrm{QRMin}(A, C)$*).*

If $\bar{y}$ is a weak minimal point of A, then by the definition we see that $0 \notin \mathrm{int}(C)$. In the next section, sometimes we may add the restriction $0 \notin \mathrm{int}(C)$ on the convex cone C. This restriction is redundant when there exists one weak minimal point. Readers should keep this in mind, and this also works for other notions defined in Definition 3.5.

As before, we can define (local and global) weak minimizers, relative minimizers etc. of a function $f : X \to Y$ and a multifunction $F : X \rightrightarrows Y$. Noting that $\mathrm{int}(C) \subset \mathrm{ri}(C) \subset \mathrm{iri}(C) \subset \mathrm{qri}(C)$ by Proposition 2.32, we can easily check the following:

Proposition 3.6. *Let* C *be a convex cone and* A *be a nonempty subset of* Y.

(1) $\mathrm{QRMin}(A,C) \subset \mathrm{IRMin}(A,C)$ *(resp.* $\mathrm{IRMin}(A,C) \subset \mathrm{RMin}(A,C)$, $\mathrm{RMin}(A,C) \subset \mathrm{WMin}(A,C)$*) provided that* $\mathrm{iri}(C) \neq \emptyset$ *(resp.* $\mathrm{ri}(C) \neq \emptyset$, $\mathrm{int}(C) \neq \emptyset$*)*

(2) If in addition C is a proper solid cone, then $\mathrm{Min}(A,C) \subset \mathrm{WMin}(A,C)$.

(3) If in addition C satisfies $\mathrm{ri}(C) \neq \emptyset$ *and* $0 \notin \mathrm{ri}(C)$ *(resp.* $\mathrm{iri}(C) \neq \emptyset$ *and* $0 \notin \mathrm{iri}(C)$, $\mathrm{qri}(C) \neq \emptyset$ *and* $0 \notin \mathrm{qri}(C)$*), then* $\mathrm{Min}(A,C) \subset \mathrm{RMin}(A,C)$ *(resp.* $\mathrm{Min}(A,C) \subset \mathrm{IRMin}(A,C)$, $\mathrm{Min}(A,C) \subset \mathrm{QRMin}(A,C)$*).*

3.2 A unified notion

In this section, we let Y be a TVS and A be a nonempty subset of Y.

Definition 3.7. *Let $\Theta \subset Y$ be an relation set containing the origin. Given $\bar{y} \in A$. We call the point $\bar{y}$ a local minimal point of A with respect to Θ (or local Θ-minimal point of A) if there exists a neighborhood V of $\bar{y}$ such that*

$$A \cap (\bar{y} - \Theta) \cap V = \{\bar{y}\}.$$

The point $\bar{y}$ is called a (global) minimal point of A with respect to Θ if we can choose the neighborhood V in (3.7) to be Y. In other words,

$$A \cap (\bar{y} - \Theta) = \{\bar{y}\}.$$

The notion unifies the following previously defined notions as follows.

Proposition 3.8. *Let $C \subset Y$ be convex cone. For each of the cases (3) to (6) below, we assume the corresponding interior is nonempty and does not contain the origin. Then $\bar{y} \in A$ is a minimal point of A with respect to*

(1) $\Theta_m = C$ *if and only if* $\bar{y} \in \mathrm{Min}(A,C)$.

(2) $\Theta_i = (Y \setminus (-C)) \cup \{0\}$ *if and only if* $\bar{y} \in \mathrm{IMin}(A,C)$.

(3) $\Theta_w = \mathrm{int}(C) \cup \{0\}$ *if and only if* $\bar{y} \in \mathrm{WMin}(A,C)$.

(4) $\Theta_r = \text{ri}(C) \cup \{0\}$ *if and only if* $\bar{y} \in \text{RMin}(A, C)$.

(5) $\Theta_{ir} = \text{iri}(C) \cup \{0\}$ *if and only if* $\bar{y} \in \text{IRMin}(A, C)$.

(6) $\Theta_q = \text{qri}(C) \cup \{0\}$ *if and only if* $\bar{y} \in \text{QRMin}(A, C)$.

Proof.

(1) This follows directly from Definition 3.1 and Definition 3.7.

(2) This follows directly from Definition 3.2 and Definition 3.7.

(3)-(6) For (3), since it is assumed $0 \notin \text{int}(C)$ so $(\text{int}(C) \cup \{0\}) \setminus \{0\} = \text{int}(C) \setminus \{0\} = \text{int}(C)$. This is true also for other interiors. Bearing this in mind, the results follow from Definition 3.5 and Definition 3.7. □

Besides the known orders given in the above proposition, in our discussion we hope to involve other kinds of orders. We illustrate two examples which are related to economics and are special cases of the unified notion.

The first example is the lexicographical order. Intuitively, we consider two goods A and B. A is considered to be much more important than B by Mr. Lee. Then when comparing whether 3 quantities of A and 2 quantities of B is better than 2 quantities of A and 3 quantities of B, he will compare the quantities of A first to conclude that the former combination is better. In other cases if the quantities of A in two choices are the same, then Mr. Lee will compare the quantities of B. This situation is described mathematically as follows.

Example 3.9 (Lexicographical order in $\mathbb{R}^n$). *Let* $x = (x_1, \cdots, x_n)$, $y = (y_1, \cdots, y_n) \in \mathbb{R}^n$. *Define the* <u>*lexicographical order*</u> $\prec_l$ *by:* $x \prec_l y$ *if and only if* $x_1 < y_1$ *or* $x_1 = y_1, \cdots, x_k = y_k, x_{k+1} < y_{k+1}$ *for some* $1 \leq k \leq n-1$. *Define* $\Theta_l \triangleq \{y \in \mathbb{R}^n : 0 \prec_l y\} \cup \{0\}$, *which is a convex solid cone containing the origin and not being a subspace. Then we can discuss the* Θ_l*-minimal points.*

Example 3.10 (Preference relation). *Let Y be a set and $Q \subset Y \times Y$ (Cartesian product). Let R be a binary relation on Y given by xRy if and only if $(x, y) \in Q$. Given $x, y \in Y$.*

(1) A strict preference on Y, denoted by $\prec$, is defined by: $x \prec y$ if and only if xRy and not yRx.

(2) An indifference on Y, denoted by $\sim$, is defined by: $x \sim y$ if and only if xRy and yRx. In particular when R is reflexive and antisymmetric, then $x \sim y$ if and only if $x = y$.

(3) A preference on Y, denoted by $\preceq$, is the disjoint union $R = \prec \cup \sim$.

To describe the set of points being better than (or "better than or equal to") a given point, we let $P(y) \triangleq \{x \in Y : x \preceq y\}$ and $P_{\prec}(y) \triangleq \{x \in Y : x \prec y\}$. Given $A \subset Y$. We say that $\bar{y} \in A$ is a preference point with respect to preference $\preceq$ if

$$A \cap P(\bar{y}) = \{\bar{y}\}.$$

Also, we say that $\bar{y} \in A$ is a weak preference point with respect to strict preference $\prec$ if

$$A \cap P_{\prec}(\bar{y}) = \emptyset.$$

In practice economists assume Y is the set of alternatives, also the preference relation on Y is transitive and any two elements in Y can be compared.

There is a correspondence [3, Remark 3.2(e)] between Θ-minimality and preference point. The proof is routine.

Proposition 3.11. *Every Θ-minimal point $\bar{y}$ of $A \subset Y$ is a preference point of A with respect to the preference defined by $P(\bar{y}) \triangleq \bar{y} - \Theta$. Every preference point $\bar{y}$ of A with respect to some preference $\preceq$ is a Θ-minimal point of A, where*

$$\Theta \triangleq \bar{y} - P(\bar{y}).$$

The following part discusses the most important notion in this chapter. The setting is as follows.

Given two real TVSs X and Y, $F : X \rightrightarrows Y$, $\emptyset \neq \Omega \subset X$ and a relation set Θ in Y containing the origin. We want to solve the problem

$$\begin{aligned} &\text{minimize } F(x) \\ &\text{subject to } x \in \Omega \end{aligned} \tag{3.2}$$

where the meaning of "minimize" is explained as follows.

Definition 3.12. *Let* $(\bar{x}, \bar{y}) \in \operatorname{gph}(F)$ *and* $\bar{x} \in \Omega$.

(1) The point $(\bar{x}, \bar{y})$ *is a* <u>*fully local minimizer*</u> *(or fully local* Θ*-minimizer if emphasizing the relation set* Θ*. This works also for the notions in (2) and (3)) of the problem (3.2) if there exist neighborhoods* U *of* $\bar{x}$ *and* V *of* $\bar{y}$ *such that*

$$F(\Omega \cap U) \cap (\bar{y} - \Theta) \cap V = \{\bar{y}\}. \tag{3.3}$$

(2) The point $(\bar{x}, \bar{y})$ *is a* <u>*partially local minimizer*</u> *of the problem (3.2) if it is a fully local minimizer which the neighborhood* V *in (3.3) is* Y.

(3) The point $(\bar{x}, \bar{y})$ *is a* <u>*global minimizer*</u> *of the problem (3.2) if it is a fully local minimizer which the neighborhoods* U *and* V *in (3.3) are* X *and* Y *respectively.*

The notion of partially local minimizer indeed imitates the conventional notions discussed in Section 3.1. For example, if the relation set $\Theta = C$ is a convex cone, then $(\bar{x}, \bar{y})$ is a partial local minimizer of the problem (3.2) if and only if it is a local Pareto minimizer of the same problem.

Sometimes we do not need to distinguish between fully and partially local minimizers.

Proposition 3.13. *[3, Proposition 4.2] Given* $F = f : X \to Y$ *which is continuous at a given* $\bar{x} \in X$*. Then* $\bar{x}$ *is a fully local minimizer of the problem (3.2) if and only if it is a partially local minimizer.*

Proof. The "if" part is clear. For the converse, we assume that $\bar{x}$ is a fully local minimizer of the problem (3.2). Then there exist a neighborhood U of $\bar{x}$ and an open neighborhood V of $\bar{y} = f(\bar{x})$ such that

$$f(\Omega \cap U) \cap (\bar{y} - \Theta) \cap V = \{\bar{y}\}. \tag{3.4}$$

Let $\tilde{U} \triangleq U \cap f^{-1}(V)$ which is a neighborhood of $\bar{x}$. It remains to show

$$f(\Omega \cap \tilde{U}) \cap (\bar{y} - \Theta) = \{\bar{y}\}. \tag{3.5}$$

Suppose there exists $y \neq \bar{y}$ such that $y = f(x) \in V$ for some $x \in \Omega \cap U$ and $y \in \bar{y} - \Theta$. This implies $y \in f(\Omega \cap U) \cap (\bar{y} - \Theta) \cap V$ which contradicts (3.4). Thus (3.5) holds. □

Example 3.14. *Proposition 3.13 may not hold for discontinuous functions. Consider $\varphi : \mathbb{R} \to \mathbb{R}$ given by*

$$\varphi(x) = \begin{cases} \log(|x|), & \text{if } x \neq 0 \\ 0, & \text{if } x = 0. \end{cases}$$

Then $(0,0)$ is a fully local minimizer for φ, but is not a partially local minimizer.

The following result [3, Proposition 4.4] converts the notion just discussed to standard one in Definition 3.7. Then the aforementioned problem (3.2) can be reduced to the problem of finding a minimal point of an a priori given set $\mathrm{gph}(F)$. Indeed, this proposition shows the first step of proving one of the main results: Theorem 5.12 in Chapter 5.

Proposition 3.15. *The point $(\bar{x}, \bar{y})$ is a fully local minimizer of the problem (3.2) if and only if it is a local Θ_{fm}-minimal point of $\mathrm{gph}(F)$ where $\Theta_{fm} \subset X \times Y$ is the product of $\bar{x} - \Omega$ and $\Theta \setminus \{0\}$ together with the point $(0,0)$. In other words,*

$$\Theta_{fm} \triangleq ((\bar{x} - \Omega) \times (\Theta \setminus \{0\})) \cup \{(0,0)\}. \tag{3.6}$$

Proof. We have:

$$
\begin{aligned}
& (\bar{x}, \bar{y}) \text{ is a fully local minimizer for the problem (3.2)} \\
\text{if and only if } & F(\Omega \cap U) \cap (\bar{y} - \Theta) \cap V = \{\bar{y}\} \\
& \text{for some neighborhoods } U \text{ of } \bar{x} \text{ and } V \text{ of } \bar{y} \\
\text{if and only if } & \operatorname{gph}(F) \cap [(\Omega \times (\bar{y} - (\Theta \setminus \{0\}))) \cup \{(\bar{x}, \bar{y})\}] \cap (U \times V) = \{(\bar{x}, \bar{y})\} \\
& \text{for some neighborhoods } U \text{ of } \bar{x} \text{ and } V \text{ of } \bar{y} \\
\text{if and only if } & \operatorname{gph}(F) \cap ((\bar{x}, \bar{y}) - \Theta_{fm}) \cap (U \times V) = \{(\bar{x}, \bar{y})\} \\
& \text{for some neighborhood } U \times V \text{ of } (\bar{x}, \bar{y}) \\
\text{if and only if } & (\bar{x}, \bar{y}) \text{ is a local } \Theta_{fm}\text{-minimal point of } \operatorname{gph}(F).
\end{aligned}
$$

□

Chapter 4

Separation theorems

4.1 Zheng and Ng fuzzy separation theorem

We are going to prove a fuzzy separation theorem [36, Theorem 3.4] which generalizes the extremal principles and the strict separation theorem.

We employ the following notation to reflect how far a given finite collection of sets is separated. Given finitely many closed subsets $A_1, \cdots, A_n$ of a Banach space X. Let $1 \le p \le +\infty$. We denote by $\gamma_p(A_1, \cdots, A_n)$ the non-intersecting index defined by

$$\gamma_p(A_1, \cdots, A_n) \triangleq \inf \left\{ \left(\sum_{i=1}^{n-1} \|x_i - x_n\|^p \right)^{\frac{1}{p}} : x_i \in A_i \text{ for all } i = 1, \cdots, n \right\} \quad (4.1)$$

where $\left(\sum_{i=1}^{n-1} \|x_i - x_n\|^p \right)^{\frac{1}{p}}$ means $\max_{1 \le i \le n-1} \|x_i - x_n\|$ when $p = +\infty$. A special case is that $\gamma_1(A_1, A_2) = d(A_1, A_2)$ (Note $d(A_1, A_2)$ is defined in P. 12). Note that $\gamma(A_1, \cdots, A_n) = 0$ if $\bigcap_{i=1}^{n} A_i \ne \emptyset$ and that for any $\varepsilon > 0$ there exist $a_i \in A_i$ $(i = 1, \cdots, n)$ such that $\left(\sum_{i=1}^{n-1} \|a_i - a_n\|^p \right)^{\frac{1}{p}} < \gamma(A_1, \cdots, A_n) + \varepsilon$.

Theorem 4.1. *Let $A_1, \cdots, A_n$ be nonempty closed subsets of an Asplund space X such that $\bigcap_{i=1}^{n} A_i = \emptyset$. Let $1 \le p, q \le +\infty$ with $\frac{1}{p} + \frac{1}{q} = 1$, $\varepsilon > 0$ and $a_i \in A_i$*

(for all $1 \le i \le n$) be such that

$$\left(\sum_{i=1}^{n-1} \|a_i - a_n\|^p\right)^{\frac{1}{p}} < \gamma_p(A_1, \cdots, A_n) + \varepsilon. \tag{4.2}$$

Then for any $\lambda > 0$ and $\rho \in (0,1)$ there exist $\widetilde{a}_i \in A_i$ and $a_i^* \in X^*$ with the following properties:

(1) $\left(\sum_{i=1}^{n} \|\widetilde{a}_i - a_i\|^p\right)^{\frac{1}{p}} < \lambda$;

(2) $\left(\sum_{i=1}^{n-1} \|a_i^*\|^q\right)^{\frac{1}{q}} = 1$, $\sum_{i=1}^{n} a_i^* = 0$, $a_i^* \in \hat{N}(\widetilde{a}_i; A_i) + \frac{\varepsilon}{\lambda}\mathbb{D}^*$ *and* $\left(\sum_{i=1}^{n} d(a_i^*, \hat{N}(\widetilde{a}_i; A_i))^q\right)^{\frac{1}{q}} < \frac{\varepsilon}{\lambda}$;

(3) $\rho\left(\sum_{i=1}^{n-1} \|\widetilde{a}_i - \widetilde{a}_n\|^p\right)^{\frac{1}{p}} \le \sum_{i=1}^{n-1}\langle a_i^*, \widetilde{a}_n - \widetilde{a}_i\rangle$.

Proof. Define a function $\phi : X^n \to \mathbb{R} \cup \{+\infty\}$ by

$$\phi(x_1, \cdots, x_n) \triangleq \left(\sum_{i=1}^{n-1} \|x_i - x_n\|^p\right)^{\frac{1}{p}} + \delta_{A_1 \times \cdots \times A_n}(x_1, \cdots, x_n) \tag{4.3}$$

for all $(x_1, \cdots, x_n) \in X^n$, where X^n is equipped with the p-norm

$$\|(x_1, \cdots, x_n)\| \triangleq \left(\sum_{i=1}^{n} \|x_i\|^p\right)^{\frac{1}{p}} \quad \text{for any } (x_1, \cdots, x_n) \in X^n.$$

Then in view of $\inf\limits_{(x_1,\cdots,x_n)\in X} \phi(x_1, \cdots, x_n) = \inf\limits_{(x_1,\cdots,x_n)\in A_1\times\cdots\times A_n} \phi(x_1, \cdots, x_n)$ and the assumption we have $\phi(a_1, \cdots, a_n) < \inf\{\phi(x_1, \cdots, x_n) : (x_1, \cdots, x_n) \in X^n\} + \varepsilon$. Choose $\varepsilon' \in (0, \varepsilon)$ such that $\phi(a_1, \cdots, a_n) < \inf\{\phi(x_1, \cdots, x_n) : (x_1, \cdots, x_n) \in X^n\} + \varepsilon'$. Then take $\lambda' \in (0, \lambda)$ such that $\frac{\varepsilon'}{\lambda'} < \frac{\varepsilon}{\lambda}$. It follows from the Ekeland's variational principle (Theorem 2.24) that there exists $(\bar{a}_1, \cdots, \bar{a}_n) \in X^n$ such that

$$\left(\sum_{i=1}^{n} \|\bar{a}_i - a_i\|^p\right)^{\frac{1}{p}} < \lambda' \quad \text{and} \tag{4.4}$$

$$\phi(\bar{a}_1, \cdots, \bar{a}_n) \le \phi(x_1, \cdots, x_n) + \frac{\varepsilon'}{\lambda'}\left(\sum_{i=1}^{n} \|x_i - \bar{a}_i\|^p\right)^{\frac{1}{p}} \text{ for all } (x_1, \cdots, x_n) \in X^n. \tag{4.5}$$

From (4.3) we know $\phi(x_1, \cdots, x_n)$ is finite if and only if $(x_1, \cdots, x_n) \in A_1 \times \cdots \times A_n$. Thus putting any point $(x_1, \cdots, x_n) \in A_1 \times \cdots \times A_n$ on the right hand side of (4.5) we see that $\phi(\bar{a}_1, \cdots, \bar{a}_n) < +\infty$ or $\bar{a}_i \in A_i$ for all i. As $\bigcap_{i=1}^{n} A_i = \emptyset$, we have

$$\left(\sum_{i=1}^{n-1} \|\bar{a}_i - \bar{a}_n\|^p\right)^{\frac{1}{p}} > 0 \tag{4.6}$$

This implies that

$$(\bar{a}_1, \cdots, \bar{a}_n) \neq (\bar{a}_n, \cdots, \bar{a}_n).$$

Then there exists $\mu > 0$ such that

$$((\bar{a}_1, \cdots, \bar{a}_n) + \mu\mathbb{D}) \cap ((\bar{a}_n, \cdots, \bar{a}_n) + \mu\mathbb{D}) = \emptyset. \tag{4.7}$$

For each $(x_1, \cdots, x_n) \in X^n$, we let

$$f(x_1, \cdots, x_n) \triangleq \left(\sum_{i=1}^{n-1} \|x_i - x_n\|^p\right)^{\frac{1}{p}} + \frac{\varepsilon'}{\lambda'}\left(\sum_{i=1}^{n} \|x_i - \bar{a}_i\|^p\right)^{\frac{1}{p}}.$$

Then f is a continuous convex real-valued function and (4.5) implies f attains a minimum on $A_1 \times \cdots \times A_n$ at $(\bar{a}_1, \cdots, \bar{a}_n)$. Hence from the definition of Fréchet subdifferential we have $0 \in \hat{\partial}(f + \delta_{A_1\times\cdots\times A_n})(\bar{a}_1, \cdots, \bar{a}_n)$. Let $\beta \in \left(0, \min\left\{\frac{\varepsilon}{\lambda} - \frac{\varepsilon'}{\lambda'}, \lambda - \lambda', \frac{\mu}{n}\right\}\right)$ be given. By applying Theorem 2.57, there exist $\bar{x}_i \in X$ and $\widetilde{a}_i \in A_i$ (for all i) such that

$$\left(\sum_{i=1}^{n} \|\bar{x}_i - \bar{a}_i\|^p\right)^{\frac{1}{p}} < \beta, \left(\sum_{i=1}^{n} \|\widetilde{a}_i - \bar{a}_i\|^p\right)^{\frac{1}{p}} < \beta, \tag{4.8}$$

and

$$0 \in \hat{\partial} f(\bar{x}_1, \cdots, \bar{x}_n) + \hat{N}((\widetilde{a}_1, \cdots, \widetilde{a}_n); A_1 \times \cdots \times A_n) + \beta\mathbb{D}^*. \tag{4.9}$$

From (4.4), the second inequality of (4.8), Minkowski inequality and the estimate $\beta < \lambda - \lambda'$ we obtain (1). By the first inequality of (4.8) and the estimate $\beta < \frac{\mu}{n}$, $(\bar{x}_1, \cdots, \bar{x}_n) \in (\bar{a}_1, \cdots, \bar{a}_n) + \mu\mathbb{D}$ and $(\bar{x}_n, \cdots, \bar{x}_n) \in (\bar{a}_n, \cdots, \bar{a}_n) + \mu\mathbb{D}$. By (4.7),

$$(\bar{x}_1, \cdots, \bar{x}_n) \neq (\bar{x}_n, \cdots, \bar{x}_n) \tag{4.10}$$

Define $g(x_1, \cdots, x_n) \triangleq \left(\sum_{i=1}^{n-1} \|x_i - x_n\|^p\right)^{\frac{1}{p}}$ for all $(x_1, \cdots, x_n) \in X^n$. Then we have $f = g + \frac{\varepsilon'}{\lambda'}\|\cdot\|_{X^n}$. Notice that g and $\|\cdot\|_{X^n}$ are convex and continuous. It follows from (4.9), Theorem 2.30, Proposition 2.28 and Proposition 2.41 that

$$0 \in \partial g(\bar{x}_1, \cdots, \bar{x}_n) + \hat{N}(\widetilde{a}_1; A_1) \times \cdots \times \hat{N}(\widetilde{a}_n; A_n) + \left(\beta + \frac{\varepsilon'}{\lambda'}\right)\mathbb{D}^*.$$

Hence there exists

$$-(a_1^*, \cdots, a_n^*) \in \partial g(\bar{x}_1, \cdots, \bar{x}_n) \tag{4.11}$$

such that

$$\left(\sum_{i=1}^{n} d(a_i^*, \hat{N}(\widetilde{a}_i; A_i))^q\right)^{\frac{1}{q}} \leq \beta + \frac{\varepsilon'}{\lambda'} < \frac{\varepsilon}{\lambda}.$$

By (4.11) and the definition of the (Fenchel) subdifferential of g, we have

$$\sum_{i=1}^{n}\langle -a_i^*, x_i - \bar{x}_i\rangle \leq \left(\sum_{i=1}^{n-1} \|x_i - x_n\|^p\right)^{\frac{1}{p}} - \left(\sum_{i=1}^{n-1} \|\bar{x}_i - \bar{x}_n\|^p\right)^{\frac{1}{p}} \tag{4.12}$$

for all $(x_1, \cdots, x_n) \in X^n$. Setting $x_1 = \cdots = x_n = x$, one has

$$\sum_{i=1}^{n}\langle -a_i^*, x - \bar{x}_i\rangle \leq -\left(\sum_{i=1}^{n-1} \|\bar{x}_i - \bar{x}_n\|^p\right)^{\frac{1}{p}}$$

for all $x \in X$. As the right hand side of the above inequality is a fixed number, we must have $\sum_{i=1}^{n} a_i^* = 0$. This and (4.12) imply that

$$\sum_{i=1}^{n-1}\langle -a_i^*, x_i - x_n - (\bar{x}_i - \bar{x}_n)\rangle \leq \left(\sum_{i=1}^{n-1} \|x_i - x_n\|^p\right)^{\frac{1}{p}} - \left(\sum_{i=1}^{n-1} \|\bar{x}_i - \bar{x}_n\|^p\right)^{\frac{1}{p}}$$

for any $(x_1, \cdots, x_n) \in X^n$. By taking an arbitrary element $(u_1, \cdots, u_{n-1}) \in X^{n-1}$ and letting $x_i \triangleq u_i + x_n$ $(1 \leq i \leq n-1)$, it follows that

$$\sum_{i=1}^{n-1}\langle -a_i^*, u_i - (\bar{x}_i - \bar{x}_n)\rangle \leq \left(\sum_{i=1}^{n-1} \|u_i\|^p\right)^{\frac{1}{p}} - \left(\sum_{i=1}^{n-1} \|\bar{x}_i - \bar{x}_n\|^p\right)^{\frac{1}{p}}.$$

Consequently

$$(-a_1^*, \cdots, -a_{n-1}^*) \in \partial\| \cdot \|_{X^{n-1}}(\bar{x}_1 - \bar{x}_n, \cdots, \bar{x}_{n-1} - \bar{x}_n). \tag{4.13}$$

By (4.10) we know $(\bar{x}_1 - \bar{x}_n, \cdots, \bar{x}_{n-1} - \bar{x}_n) \neq (0, \cdots, 0)$. Thus we get from (4.13) and Proposition 2.28

$$\left(\sum_{i=1}^{n-1} \|a_i^*\|^q\right)^{\frac{1}{q}} = 1 \text{ and } \sum_{i=1}^{n-1} \langle a_i^*, \bar{x}_n - \bar{x}_i \rangle = \left(\sum_{i=1}^{n-1} \|\bar{x}_i - \bar{x}_n\|^p\right)^{\frac{1}{p}}. \tag{4.14}$$

We now arrive at (2). By the second equality of (4.14), the implication $\|a_i^*\| \leq 1$ $(1 \leq i \leq n-1)$ from the first equality of (4.14), Minkowski inequality and (4.8),

$$\begin{aligned}
\sum_{i=1}^{n-1} \langle a_i^*, \widetilde{a}_n - \widetilde{a}_i \rangle &= \sum_{i=1}^{n-1} \langle a_i^*, \bar{x}_n - \bar{x}_i \rangle + \sum_{i=1}^{n-1} \langle a_i^*, \widetilde{a}_n - \bar{x}_n - (\widetilde{a}_i - \bar{x}_i) \rangle \\
&\geq \left(\sum_{i=1}^{n-1} \|\bar{x}_i - \bar{x}_n\|^p\right)^{\frac{1}{p}} - \left(\sum_{i=1}^{n-1} \|\widetilde{a}_n - \bar{x}_n - (\widetilde{a}_i - \bar{x}_i)\|^p\right)^{\frac{1}{p}} \\
&\geq \left(\sum_{i=1}^{n-1} \|\widetilde{a}_i - \widetilde{a}_n\|^p\right)^{\frac{1}{p}} - 2\left(\sum_{i=1}^{n-1} \|\widetilde{a}_n - \bar{x}_n - (\widetilde{a}_i - \bar{x}_i)\|^p\right)^{\frac{1}{p}} \\
&\geq \left(\sum_{i=1}^{n-1} \|\widetilde{a}_i - \widetilde{a}_n\|^p\right)^{\frac{1}{p}} - 2\left(\sum_{i=1}^{n-1} (2\beta)^p\right)^{\frac{1}{p}} \\
&= \left(\sum_{i=1}^{n-1} \|\widetilde{a}_i - \widetilde{a}_n\|^p\right)^{\frac{1}{p}} - 4(n-1)^{\frac{1}{p}}\beta.
\end{aligned} \tag{4.15}$$

Since β is arbitrary in $\left(0, \min\left\{\frac{\varepsilon}{\lambda} - \frac{\varepsilon'}{\lambda'}, \lambda - \lambda', \frac{\mu}{n}\right\}\right)$, using the second inequality of (4.8), and (4.6) we see that

$$\lim_{\beta \to 0^+} \left(\left(\sum_{i=1}^{n-1} \|\widetilde{a}_i - \widetilde{a}_n\|^p\right) - 4(n-1)^{\frac{1}{p}}\beta\right) = \left(\sum_{i=1}^{n} \|\bar{a}_i - \bar{a}_n\|^p\right)^{\frac{1}{p}} > 0.$$

It follows that for a given $\rho \in (0,1)$, one has

$$\begin{aligned}
\lim_{\beta \to 0^+} \rho \left(\sum_{i=1}^{n-1} \|\widetilde{a}_i - \widetilde{a}_n\|\right)^{\frac{1}{p}} &= \rho \left(\sum_{i=1}^{n-1} \|\bar{a}_i - \bar{a}_n\|\right)^{\frac{1}{p}} \\
&< \lim_{\beta \to 0^+} \left(\left(\sum_{i=1}^{n-1} \|\bar{a}_i - \bar{a}_n\|^p\right) - 4(n-1)^{\frac{1}{p}}\beta\right).
\end{aligned}$$

This implies that there exists a sufficiently small $\beta > 0$ such that

$$\rho\left(\sum_{i=1}^{n-1}\|\widetilde{a}_i - \widetilde{a}_n\|\right)^{\frac{1}{p}} < \left(\sum_{i=1}^{n-1}\|\bar{a}_i - \bar{a}_n\|^p\right) - 4(n-1)^{\frac{1}{p}}\beta \leq \sum_{i=1}^{n-1}\langle a_i^*, \widetilde{a}_n - \widetilde{a}_i\rangle,$$

where the last inequality follows from (4.15). We now have (3) and the whole proof. □

4.2 Extremal principles and other consequences

Extremal principles can be viewed as some sort of nonconvex separation theorems. They are discussed in the Chapter 2 of [25]. The reference point we study is called the local extremal point which is defined as follows:

Definition 4.2. *Let $\Omega_1, \cdots, \Omega_n$ ($n \geq 2$) be nonempty subsets of a Banach space X. We say that $x \in \bigcap_{i=1}^{n} \Omega_i$ is a local extremal point of the set system $\{\Omega_1, \cdots, \Omega_n\}$ if there are sequences $\{a_{ik}\} \subset X$ ($i = 1, \cdots, n$), and a neighborhood U of $\bar{x}$ such that $a_{ik} \to 0$ as $k \to \infty$ for all i and $\bigcap_{i=1}^{n}(\Omega_i + a_{ik}) \cap U = \emptyset$ for all k. In this case $\{\Omega_1, \cdots, \Omega_n; \bar{x}\}$ is said to be an extremal system in X.*

The approximate extremal principle can be derived from the preceding theorem. The proof can be found in [35, P. 1161, Remark], however we will do it here for completeness. Readers can find this extremal principle in [25, P. 199, Theorem 2.20], but its origin is [27].

Theorem 4.3 (Approximate extremal principle). *Let X be an Asplund space and $\{\Omega_1, \cdots, \Omega_n; \bar{x}\}$ be an extremal system in X. Suppose all Ω_i is locally closed at $\bar{x}$. Then for every $\varepsilon > 0$ there are $x_i \in \Omega_i \cap (\bar{x} + \varepsilon\mathbb{B})$ and $x_i^* \in \hat{N}(x_i; \Omega_i) + \varepsilon\mathbb{B}^*$ for $i = 1, \cdots, n$ such that $\sum_{i=1}^{n} x_i^* = 0$ and $\sum_{i=1}^{n}\|x_i^*\| = 1$.*

Proof. By the local closedness assumption, choose $r_0 > 0$ such that for all $i = 1, \cdots, n$, $\Omega_i \cap B(\bar{x}, r_0)$ is closed. As $\bar{x}$ is a local extremal point of the system $\{\Omega_1, \cdots, \Omega_n\}$, there exists $r \in (0, r_0)$ satisfying: for all $\sigma > 0$, there exist

$c_1, \cdots, c_n \in X$ such that $\|c_i\| < \dfrac{\sigma^2}{4n}$ for all $i = 1, \cdots, n$ and $\bigcap_{i=1}^{n}(\Omega_i + c_i) \cap B(\bar{x}, r) = \emptyset$. This implies that

$$\sum_{i=1}^{n-1} \|(x + c_i) - (x + c_n)\| \le \sum_{i=1}^{n-1}(\|c_i\| + \|c_n\|) < \frac{\sigma^2}{2} \le \gamma_1(A_1, \cdots, A_n) + \frac{\sigma^2}{2}.$$

Define for all i, $A_i \triangleq (\Omega_i + c_i) \cap B(\bar{x}, r)$. Then for all i, since $\Omega_i \cap B(\bar{x}, r_0)$ is closed, so does A_i. By applying Theorem 4.1 with the data $p = 1$, $A_i = (\Omega_i + c_i) \cap B(\bar{x}, r)$, $a_i = \bar{x} + c_i$ $(i = 1, \cdots, n)$, $\varepsilon = \dfrac{\sigma^2}{2}$ and $\lambda = \sigma$, there exist $x_i \in \Omega_i$ and $a_i^* \in X^*$ such that $\|x_i - \bar{x}\| < \sigma$, $a_i^* \in \hat{N}(x_i; \Omega_i) + \dfrac{\sigma}{2}\mathbb{B}^*$ $(i = 1, \cdots, n)$, $\max_{1 \le i \le n} \|a_i^*\| = 1$ and $\sum_{i=1}^{n} a_i^* = 0$. Note $\sum_{i=1}^{n} \|a_i^*\| \ge \max_{1 \le i \le n} \|a_i^*\| = 1$ and so let $\widetilde{a}_i^* \triangleq \dfrac{a_i^*}{\sum_{i=1}^{n} \|a_i^*\|}$ for all i. We obtain $\widetilde{a}_i^* \in \hat{N}(x_i; \Omega_i) + \dfrac{\sigma}{2\sum_{i=1}^{n} \|a_i^*\|}\mathbb{B}^* \subset \hat{N}(x_i; \Omega_i) + \dfrac{\sigma}{2}\mathbb{B}^*$, $\sum_{i=1}^{n} \widetilde{a}_i^* = 0$ and $\sum_{i=1}^{n} \|\widetilde{a}_i^*\| = 1$. The proof is completed. □

Corollary 4.4. *Let X be an Asplund space and $\{\Omega_1, \cdots, \Omega_n; \bar{x}\}$ be an extremal system in X. Then for any $\varepsilon > 0$ there are $x_i \in \Omega \cap (\bar{x} + \varepsilon\mathbb{B})$ and $y_i^* \in \hat{N}(x_i; \Omega_i)$ for $i = 1, \cdots, n$ such that $\left\| \sum_{i=1}^{n} y_i^* \right\| \le \varepsilon$ and $1 - \varepsilon \le \sum_{i=1}^{n} \|y_i^*\| \le 1 + \varepsilon$.*

Proof. For a given $\varepsilon > 0$, by Theorem 4.3 there are $x_i \in \Omega \cap \left(\bar{x} + \dfrac{\varepsilon}{n}\mathbb{B}\right) \subset \Omega \cap (\bar{x} + \varepsilon\mathbb{B})$ and $y_i^* \in \hat{N}(x_i; \Omega_i)$ and $z_i^* \in \mathbb{B}^*$ for $i = 1, \cdots, n$ such that $\sum_{i=1}^{n} y_i^* = -\dfrac{\varepsilon}{n}\left(\sum_{i=1}^{n} z_i^*\right)$ and $\sum_{i=1}^{n} \left\| y_i^* + \dfrac{\varepsilon}{n} z_i^* \right\| = 1$. Hence by the triangle inequality,

$$\left\| \sum_{i=1}^{n} y_i^* \right\| = \left\| -\frac{\varepsilon}{n}\left(\sum_{i=1}^{n} z_i^*\right) \right\| \le \frac{\varepsilon}{n}\left(\sum_{i=1}^{n} \|z_i^*\|\right) \le \varepsilon,$$
$$1 \le \sum_{i=1}^{n} \left(\|y_i^*\| + \frac{\varepsilon}{n}\|z_i^*\|\right) \le \left(\sum_{i=1}^{n} \|y_i^*\|\right) + \varepsilon \text{ and}$$
$$\sum_{i=1}^{n} \|y_1^*\| \le 1 + \frac{\varepsilon}{n} \sum_{i=1}^{n} \|z_i^*\| \le 1 + \varepsilon.$$

We then reach the desired conclusion. □

If we further add SNC assumptions on the above collection $\{\Omega_1, \cdots, \Omega_n\}$, then we can derive the so called exact extremal principle as follows:

Theorem 4.5 (Exact extremal principle). *[25, P. 201, Theorem 2.22(i)] Let X be an Asplund space and $\{\Omega_1, \cdots, \Omega_n; \bar{x}\}$ be an extremal system in X. Assume both Ω_i are locally closed at $\bar{x}$ for all i and at least any $n-1$ sets from the collection $\{\Omega_i : 1 \leq i \leq n\}$ are SNC at $\bar{x}$. Then there are $x_i^* \in N(\bar{x}; \Omega_i)$ for $i = 1, \cdots, n$ such that $\sum_{i=1}^{n} x_i^* = 0$ and $\sum_{i=1}^{n} \|x_i^*\| = 1$.*

Proof. By Theorem 4.3 and the fact that $\hat{N}(a; A) + \varepsilon\mathbb{B}^* \subset \hat{N}_\varepsilon(a; A)$ for any $A \subset X$, $a \in A$ and $\varepsilon \geq 0$, we have for any $k \in \mathbb{N}$, there exist sequences $\{x_{ik}\} \subset \Omega_i \cap \left(\bar{x} + \frac{1}{k}\mathbb{B}\right)$, $\{x_{ik}^*\} \subset \hat{N}_{\frac{1}{k}}(x_{ik}; \Omega_i)$ (for all i) such that for all i,

$$\sum_{i=1}^{n} x_{ik}^* = 0 \tag{4.16}$$

and

$$\sum_{i=1}^{n} \|x_{ik}^*\| = 1. \tag{4.17}$$

Then for all i, $x_{ik} \to \bar{x}$ as $k \to \infty$ and $\{x_{ik}^*\}_{k\in\mathbb{N}}$ is a bounded sequence in X^*. Since X is Asplund, applying Theorem 2.9 we have that for all i, $x_{ik}^* \xrightarrow{w^*} x_i^*$ as $k \to \infty$ (by passing to a subsequence when necessary). By passing to limit as $k \to \infty$, we obtain

$$x_i^* \in N(x_i; \Omega_i) \text{ for all } i. \tag{4.18}$$

Also by (4.16),

$$\sum_{i=1}^{n} x_i^* = 0. \tag{4.19}$$

Now we claim that $(x_1^*, \cdots, x_n^*) \neq (0, \cdots, 0)$. Indeed if $x_i^* = 0$ for all i, then by the SNC assumptions we have for all $1 \leq i \leq n-1$, $\|x_{ik}^*\| \to 0$ as $k \to \infty$. Then $\|x_{nk}^*\| = \|\sum_{i=1}^{n-1} x_{ik}^*\| \leq \sum_{i=1}^{n-1} \|x_{ik}^*\| \to 0$ as $k \to \infty$, contradicting (4.17). The claim is justified. Hence we let $\widetilde{x}_i^* \triangleq \dfrac{x_i^*}{\sum_{i=1}^{n} \|x_i^*\|}$ for $i = 1, \cdots, n$. Then $\sum_{i=1}^{n} \|\widetilde{x}_i^*\| = 1$, and by (4.18) and (4.19) we have $\widetilde{x}_i^* \in N(x_i; \Omega_i)$ for all i and $\sum_{i=1}^{n} \widetilde{x}_i^* = 0$. □

If we consider the above exact extremal principle for two convex sets Ω_1 and Ω_2 in X and $\bar{x} \in \Omega_1 \cap \Omega_2$, then the conclusion means that there exists $0 \neq x^* \in \hat{N}(\bar{x};\Omega_1) \cap (-\hat{N}(\bar{x};\Omega_2))$ which means $\langle x^*, x_1\rangle \leq \langle x^*, \bar{x}\rangle \leq \langle x^*, x_2\rangle$ for all $x_i \in \Omega_i$ $(i = 1, 2)$. This is the classical separation property of convex sets.

Not only the extremal principles, the strict separation theorem (Theorem 2.7) can also be derived from Theorem 4.1 where the underlying space is Asplund. First we prove the following corollary [36, Corollary 3.3]:

Corollary 4.6. *Let A_1 be a nonempty closed subset of an Asplund space X. And let A_2 be a nonempty, closed, bounded and convex subset of X. Suppose that $A_1 \cap A_2 = \emptyset$. Then for any $\varepsilon > 0$ there exist $a_1 \in A_1$ and $a^* \in \hat{N}(a_1; A_1)$ with $\|a^*\| = 1$ such that*

$$d(A_1, A_2) - \varepsilon < \inf_{x\in A_2} \langle a^*, x\rangle - \langle a^*, a_1\rangle.$$

Consequently, if in addition A_1 is convex then

$$d(A_1, A_2) - \varepsilon < \inf_{x\in A_2} \langle a^*, x\rangle - \max_{x\in A_1}\langle a^*, x\rangle.$$

Proof. Let k be an arbitrary natural number and take $a_i(k) \in A_i$ such that

$$d(A_1, A_2) \leq \|a_1(k) - a_2(k)\| < d(A_1, A_2) + \frac{1}{k^2} = \gamma_1(A_1, A_2) + \frac{1}{k^2} \tag{4.20}$$

(cf. the remark after (4.1)). By Theorem 4.1 (take $p = 1$), there exist $\widetilde{a}_i(k) \in A_i$ and $a_i^*(k) \in X^*$ $(i = 1, 2)$ such that

$$\|\widetilde{a}_1(k) - a_1(k)\| + \|\widetilde{a}_2(k) - a_2(k)\| < \frac{1}{k}, \tag{4.21}$$

$$\max_{i=1,2} \|a_i^*(k)\| \leq 1, a_1^*(k) + a_2^*(k) = 0, a_i^*(k) \in \hat{N}(\widetilde{a}_i(k); A_i) + \frac{1}{k}\mathbb{D}^*,\ i = 1, 2, \text{ and} \tag{4.22}$$

$$\left(1 - \frac{1}{k}\right) \|\widetilde{a}_1(k) - \widetilde{a}_2(k)\| \leq \langle a_1^*(k), \widetilde{a}_2(k) - \widetilde{a}_1(k)\rangle. \tag{4.23}$$

By (4.22) we take $\widetilde{a}_i^*(k) \in \hat{N}(\widetilde{a}_i(k); A_i)$ such that $\|\widetilde{a}_i^*(k) - a_i^*(k)\| < \frac{1}{k}$ $(i = 1, 2)$. Then by (4.22) and triangle inequality,

$$1 - \frac{1}{k} < \|\widetilde{a}_i^*(k)\| < 1 + \frac{1}{k},\ \|\widetilde{a}_1^*(k) + \widetilde{a}_2^*(k)\| < \frac{2}{k}. \tag{4.24}$$

Notice that (4.23) implies

$$\langle a_1^*(k), \widetilde{a}_2(k) - \widetilde{a}_1(k)\rangle \geq 0. \tag{4.25}$$

As $\widetilde{a}_2^*(k) \in \hat{N}(\widetilde{a}_2(k); A_2)$ and A_2 is convex, we have

$$\langle \widetilde{a}_2^*(k), \widetilde{a}_2(k)\rangle = \max_{x\in A_2}\langle \widetilde{a}_2^*(k), x\rangle. \tag{4.26}$$

Let $L \triangleq \sup_{x\in A_2} \|x\| < \infty$ (note A_2 is bounded). Then by the second equality of (4.22) we have

$$-\max_{x\in A_2}\langle \widetilde{a}_2^*(k), x\rangle \leq \inf_{x\in A_2}\langle \widetilde{a}_1^*(k), x\rangle + \frac{L}{k} + \frac{L}{k} = \inf_{x\in A_2}\langle \widetilde{a}_1^*(k), x\rangle + \frac{2L}{k}. \tag{4.27}$$

Using (4.23), (4.25), (4.26) and (4.27) we derive that

$$\begin{aligned}
&\left(1-\frac{1}{k}\right)^2 \|\widetilde{a}_1(k) - \widetilde{a}_2(k)\| \\
&\leq \left(1-\frac{1}{k}\right)\langle a_1^*(k), \widetilde{a}_2(k) - \widetilde{a}_1(k)\rangle \\
&\leq \langle a_1^*(k), \widetilde{a}_2(k) - \widetilde{a}_1(k)\rangle \\
&= \langle -\widetilde{a}_2^*(k), \widetilde{a}_2(k)\rangle - \langle \widetilde{a}_1^*(k), \widetilde{a}_1(k)\rangle + \langle \widetilde{a}_1^*(k) + \widetilde{a}_2^*(k), \widetilde{a}_2(k)\rangle \\
&\leq \langle -\widetilde{a}_2^*(k), \widetilde{a}_2(k)\rangle - \langle \widetilde{a}_1^*(k), \widetilde{a}_1(k)\rangle + \|\widetilde{a}_1^*(k) + \widetilde{a}_2^*(k)\|\|\widetilde{a}_2(k)\| \\
&\leq -\max_{x\in A_2}\langle \widetilde{a}_1^*(k), x\rangle - \langle \widetilde{a}_1^*(k), \widetilde{a}_1(k)\rangle + \frac{2L}{k} \\
&\leq \inf_{x\in A_2}\langle \widetilde{a}_1^*(k), x\rangle - \langle \widetilde{a}_1^*(k), \widetilde{a}_1(k)\rangle + \frac{4L}{k}.
\end{aligned} \tag{4.28}$$

By the first two equalities of (4.22) we see that $\widetilde{a}_1^*(k) \neq 0$ and so we let $\widetilde{a}^*(k) \triangleq \frac{\widetilde{a}_1^*(k)}{\|\widetilde{a}_1^*(k)\|}$. Then $\widetilde{a}^*(k) \in \hat{N}(\widetilde{a}_1(k); A_1)$, $\|\widetilde{a}^*(k)\| = 1$ and

$$\frac{\left(1-\frac{1}{k}\right)^2 \|\widetilde{a}_1(k) - \widetilde{a}_2(k)\| - \frac{4L}{k}}{\|\widetilde{a}_1^*(k)\|} \leq \inf_{x\in A_2}\langle \widetilde{a}^*(k), x\rangle - \langle \widetilde{a}^*(k), \widetilde{a}_1(k)\rangle. \tag{4.29}$$

By (4.20), (4.21) and the first inequality of (4.24), one has

$$\frac{\left(1-\frac{1}{k}\right)^2 \|\widetilde{a}_1(k) - \widetilde{a}_2(k)\| - \frac{4L}{k}}{\|\widetilde{a}_1^*(k)\|} \to \frac{1\cdot d(A_1, A_2) - 0}{1} = d(A_1, A_2)$$

as $k \to \infty$. Now let $\varepsilon > 0$ be given. Then for all k sufficiently large, we have

$$d(A_1, A_2) - \varepsilon < \frac{\left(1 - \frac{1}{k}\right)^2 \|\widetilde{a}_1(k) - \widetilde{a}_2(k)\| - \frac{4L}{k}}{\|\widetilde{a}_1^*(k)\|} \leq \inf_{x \in A_2} \langle \widetilde{a}^*(k), x\rangle - \langle \widetilde{a}^*(k), \widetilde{a}_1(k)\rangle$$

where the last inequality follows from (4.29). We have proved the first assertion. If in addition A_1 is convex, then since $\widetilde{a}_1^*(k) \in \hat{N}(\widetilde{a}_1(k); A_1)$, it follows that $\langle \widetilde{a}_1^*(k), \widetilde{a}_1(k)\rangle = \max_{x \in A_1} \langle \widetilde{a}_1^*(k), x\rangle$. Therefore, from (4.28) we get

$$\begin{aligned}\left(1 - \frac{1}{k}\right)^2 \|\widetilde{a}_1(k) - \widetilde{a}_2(k)\| &\leq \inf_{x \in A_2} \langle \widetilde{a}_1^*(k), x\rangle - \langle \widetilde{a}_1^*(k), \widetilde{a}_1(k)\rangle + \frac{4L}{k} \\ &= \inf_{x \in A_2} \langle \widetilde{a}_1^*(k), x\rangle - \max_{x \in A_1} \langle \widetilde{a}_1^*(k), x\rangle + \frac{4L}{k}.\end{aligned}$$

Carrying on the argument as above, we obtain the second assertion. □

Now we recover the strict separation theorem in the Asplund space setting using the above corollary (this is done in [36, Remark after the proof of Corollary 3.3]).

Theorem 4.7. *Let A and B be nonempty convex subsets of an Asplund space X where A is compact and B is closed. Then $A \cap B = \emptyset$ if and only if there is $x^* \in X^* \setminus \{0\}$ such that*

$$\sup_{a \in A} \langle x^*, a\rangle < \inf_{b \in B} \langle x^*, b\rangle.$$

Proof. The "if" part is clear. For the "only if" part, we know that $d(A, B) > 0$ since $A \cap B = \emptyset$ and A is compact. By taking $A_1 = B$, $A_2 = A$ and $\varepsilon = \dfrac{d(A_1, A_2)}{2} > 0$ in Corollary 4.6 we obtain the desired conclusion. □

A separation theorem similar to Theorem 4.1 except that the Asplund space setting is replaced by the Banach space setting and the Fréchet normal cone is replaced by the Clarke normal cone, is shown in Zheng and Ng's recent paper [36, Theorem 3.1]. By using that result the strict separation theorem under the Banach space setting can be derived. Since this similar result is not used in the later discussion, we omit the details here.

Chapter 5

Necessary conditions for the unified notion of optimality

We now formulate the necessary conditions for the unified notion defined in Chapter 3. Note that in this chapter, the underlying spaces are always Banach spaces.

5.1 Local asymptotic closedness

In this section we let Y be a Banach space. Bao and Mordukhovich [3, Definition 3.3] introduced the following LAC property which is an essential assumption to the following results.

Definition 5.1. *Let $A \subset Y$ and $\bar{y} \in \text{cl}(A)$. The set A is said to have the local asymptotic closedness (LAC) property at $\bar{y}$ (or A is local asymptotic closed (LAC) at $\bar{y}$) if there exist a neighborhood V of $\bar{y}$ and a sequence $\{c_k\} \subset Y$ with $\|c_k\| \to 0$ as $k \to \infty$ satisfying*

$$(\text{cl}(A) + c_k) \cap V \subset A \setminus \{\bar{y}\} \text{ for all } k \in \mathbb{N}. \tag{5.1}$$

Remark 5.2. *If A is LAC at $\bar{y}$, then* $\text{cl}(A)$ *is also LAC at $\bar{y}$ (because* $\text{cl}(\text{cl}(A)) = \text{cl}(A)$*). However the converse may not be true. For example, let $A \triangleq \{(y_1, y_2) \in \mathbb{Q}^2 : y_1 \geq 0\}$ and $\bar{y} = (0,0) \in \text{cl}(A) = \mathbb{R}_+ \times \mathbb{R}$. Then consider $c_k \triangleq \left(\frac{1}{k}, 0\right)$ so*

$c_k \to 0$ and $\mathrm{cl}(A) + c_k = \left(\mathbb{R}_+ + \frac{1}{k}\right) \times \mathbb{R} \subset \mathrm{cl}(A) \setminus \{0\}$ for all k. This means $\mathrm{cl}(A)$ is LAC at $\bar{y}$. However for any $c \in \mathbb{R}^2$, $\mathrm{cl}(A) + c$ contains some point not in A. Therefore A is not LAC at $\bar{y}$.

Mathematically speaking, the importance of LAC assumption is that it guarantees the local extremality of the Θ-minimal points introduced in Chapter 3 (see Definition 3.7).

Theorem 5.3. *[3, Theorem 3.4] Let Θ be a relation set in Y and $\bar{y} \in A$ be a local Θ-minimal point of A. We consider the following statements:*

(A) The point $\bar{y}$ is a local extremal point of the set system $\{A, \bar{y} - \mathrm{cl}(\Theta)\}$.

(B) The point $\bar{y}$ is a local extremal point of the set system $\{\mathrm{cl}(A), \bar{y} - \Theta\}$.

(C) The point $\bar{y}$ is a local extremal point of the set system $\{\mathrm{cl}(A), \bar{y} - \mathrm{cl}(\Theta)\}$.

Then the following assertions hold:

(1) If Θ is LAC at 0, then (A) is true. If in addition A is locally closed at $\bar{y}$, then (C) is true.

(2) If A is LAC at $\bar{y}$, then (B) is true. If in addition Θ is locally closed at 0, then (C) is true.

(3) If both A and Θ are LAC at $\bar{y}$ and 0 respectively, then (C) is true.

Proof. Since $\bar{y} \in A$ be a local Θ-minimal point of A, then there exists a neighborhood V of $\bar{y}$ such that $A \cap (\bar{y} - \Theta) \cap V = \{\bar{y}\}$ implying

$$A \cap (\bar{y} - \Theta \setminus \{0\}) \cap V = \emptyset, \quad \text{and} \tag{5.2}$$

$$A \setminus \{\bar{y}\} \cap (\bar{y} - \Theta) \cap V = \emptyset. \tag{5.3}$$

(1) As Θ is LAC at 0, $\bar{y} - \Theta$ is LAC at $\bar{y}$. Hence there exist a sequence $\{c_k\}$ in Y converging to 0 and a neighborhood V_1 of $\bar{y}$ such that

$$(\bar{y} - \mathrm{cl}(\Theta) + c_k) \cap V_1 \subset (\bar{y} - \Theta) \setminus \{\bar{y}\} = \bar{y} - \Theta \setminus \{0\} \text{ for all } k \in \mathbb{N}.$$

Hence, using (5.2), we see that for each k,

$$A \cap (\bar{y} - \text{cl}(\Theta) + c_k) \cap (V \cap V_1) \subset A \cap (\bar{y} - \Theta \setminus \{0\}) \cap V = \emptyset.$$

Thus (A) is true. Suppose further A is locally closed at $\bar{y}$. Then there exists a neigborhood V_2 of $\bar{y}$ such that $\text{cl}(A) \cap V_2 \subset A$. As a result, by (5.2), for each k,

$$\begin{aligned} &\text{cl}(A) \cap (\bar{y} - \text{cl}(\Theta) + c_k) \cap (V \cap V_1 \cap V_2) \\ &= (\text{cl}(A) \cap V_2) \cap ((\bar{y} - \text{cl}(\Theta) + c_k) \cap V_1) \cap V \\ &\subset A \cap (\bar{y} - \Theta \setminus \{0\}) \cap V = \emptyset. \end{aligned}$$

This implies (C).

(2) The desired result follows by writing a similar proof as (1) and using (5.3) instead of (5.2).

(3) As $\text{cl}(A)$ is LAC at $\bar{y}$ and $\bar{y} - \Theta$ is LAC at $\bar{y}$, there exist two sequences $\{a_k\}$ and $\{c_k\}$ converging to 0 and a neighborhood V of $\bar{y}$ such that for all k,

$$(\text{cl}(A) + a_k) \cap V \subset A \setminus \{\bar{y}\}$$

and

$$(\bar{y} - \text{cl}(\Theta) + c_k) \cap V \subset \bar{y} - \Theta \setminus \{0\}.$$

Consequently, using (5.2), for each k,

$$\begin{aligned} (\text{cl}(A) + a_k) \cap (\text{cl}(\bar{y} - \Theta) + c_k) \cap V &= (\text{cl}(A) + a_k) \cap (\bar{y} - \text{cl}(\Theta) + c_k) \cap V \\ &\subset (A \setminus \{\bar{y}\}) \cap (\bar{y} - \Theta \setminus \{0\}) \cap V = \emptyset. \end{aligned}$$

Hence we get the desired conclusion.

□

The next result [3, Proposition 3.8, Corollary 3.9] illustrates the LAC property of the ordering cones discussed in Section 3.2.

Proposition 5.4. *Let $C \subset Y$ be a convex cone. Then the following statements hold:*

(1) If $\operatorname{int}(C) \neq \emptyset$ and $0 \notin \operatorname{int}(C)$, then both $\operatorname{int}(C)$ and $\Theta_w \triangleq \operatorname{int}(C) \cup \{0\}$ are LAC at 0.

(2) If $\operatorname{ri}(C) \neq \emptyset$ and $0 \notin \operatorname{ri}(C)$, then both $\operatorname{ri}(C)$ and $\Theta_r \triangleq \operatorname{ri}(C) \cup \{0\}$ are LAC at 0.

(3) If $C \setminus (-C) \neq \emptyset$ and C is locally closed at 0, then $\Theta_m \triangleq C$ is LAC at 0.

(4) If $\operatorname{int}(C) \neq \emptyset$ then $\Theta_i \triangleq (Y \setminus (-C)) \cup \{0\}$ is LAC at 0.

(5) The lexicographical ordering cone (cf. Example 3.9) Θ_l is LAC at 0.

Proof.

(1) Take an arbitrary $c \in \operatorname{int}(C) \setminus \{0\}$. Define for all k, $c_k \triangleq k^{-1}c$. Then as $\operatorname{int}(C)$ is a cone we have $c_k \in \operatorname{int}(C)$ for all k. Also $c_k \to 0$ as $k \neq \infty$. Invoking Proposition 2.16 (4b) and (5), and noting $0 \notin \operatorname{int}(C)$, we have for all k,

$$\operatorname{cl}(\operatorname{int}(C)) + c_k = \operatorname{cl}(C) + c_k \subset \operatorname{cl}(C) + \operatorname{int}(C) = \operatorname{int}(C) = \operatorname{int}(C) \setminus \{0\}. \quad (5.4)$$

Thus $\operatorname{int}(C)$ is LAC at 0. With the observations that $\operatorname{cl}(\operatorname{int}(C) \cup \{0\}) = \operatorname{cl}(\operatorname{int}(C))$ and $(\operatorname{int}(C) \cup \{0\}) \setminus \{0\} = \operatorname{int}(C) \setminus \{0\}$, it follows from (5.4) that $\operatorname{int}(C) \cup \{0\}$ is also LAC at 0.

(2) The conclusion is obtained by writing a similar proof as that of (1), with applying Proposition 2.34 (2) and (3).

(3) Since C is locally closed at 0, then there exists $\varepsilon > 0$ such that

$$\operatorname{cl}(C) \cap \varepsilon\mathbb{B} \subset C. \quad (5.5)$$

Then pick $c \in C \setminus (-C)$ (then $c \neq 0$) and consider $c_k \triangleq k^{-1}c$ for all k. We have $c_k \in C \setminus (-C)$ and $c_k \to 0$ as $k \to \infty$. For each k, $c_k \notin -C$ so $0 \notin C + c_k$.

As C is a convex cone, by using Proposition 2.11,

$$C + c_k \subset C \setminus \{0\} \text{ for all } k. \tag{5.6}$$

Choose $K \in \mathbb{N}$ such that for all $k \geq K$, $\|c_k\| \leq \frac{1}{2}\varepsilon$. Consider the sequence $\{d_n\} \triangleq \{c_k\}_{k \geq K}$ converging to 0, and let $V \triangleq \frac{1}{2}\varepsilon\mathbb{B}$. Then for all n,

$$(\mathrm{cl}(C) + d_n) \cap V \subset \mathrm{cl}(C) \cap \varepsilon\mathbb{B} + d_n. \tag{5.7}$$

Indeed, if $y \in \mathrm{cl}(C)$ and $\|y + d_n\| \leq \frac{1}{2}\varepsilon$, then $\|y\| \leq \|y + d_n\| + \|d_n\| \leq \varepsilon$. Hence $y \in \varepsilon\mathbb{B}$ and (5.7) is proved. Therefore by (5.7), (5.5) and (5.6),

$$(\mathrm{cl}(C) + d_n) \cap V \subset \mathrm{cl}(C) \cap \varepsilon\mathbb{B} + d_n \subset C + d_n \subset C \setminus \{0\} \text{ for all } n$$

which implies that C is LAC at 0.

(4) Pick $c \in \mathrm{int}(C)$ and consider $c_k \triangleq k^{-1}c$ for all k. Notice

$$\mathrm{cl}(\Theta_i) = \mathrm{cl}((Y \setminus (-C)) \cup \{0\}) = \mathrm{cl}(Y \setminus (-C)) = Y \setminus (-\mathrm{int}(C)) \tag{5.8}$$

and as $0 \notin Y \setminus (-C)$,

$$\Theta_i \setminus \{0\} = ((Y \setminus (-C)) \cup \{0\}) \setminus \{0\} = (Y \setminus (-C)) \setminus \{0\} = Y \setminus (-C). \tag{5.9}$$

We claim that

$$Y \setminus (-\mathrm{int}(C)) + c_k \subset Y \setminus (-C) \text{ for all } k. \tag{5.10}$$

Indeed, if there exist $k \in \mathbb{N}$ and $y \in Y$ such that $y - c_k \in Y \setminus -\mathrm{int}(C)$ but $y \in -C$, then since $c_k \in \mathrm{int}(C)$, $y - c_k \in -C - c_k \subset -\mathrm{int}(C)$ which is a contradiction. (5.10) is seen to be true. Finally, using (5.8), (5.10) and (5.9), we have for all k,

$$\mathrm{cl}(\Theta_i) + c_k = Y \setminus (-\mathrm{int}(C)) + c_k \subset Y \setminus (-C) = \Theta_i \setminus \{0\}.$$

This means Θ_i is LAC at 0.

(5) Note that Θ_l has nonempty interior and $0 \notin \mathrm{int}(\Theta_l)$. By applying (1) the set $\mathrm{int}(\Theta_l)$ is LAC at 0. It follows from the fact $\mathrm{cl}(\Theta_l) = \mathrm{cl}(\mathrm{int}(\Theta_l))$ (see Proposition 2.16 (4b)) that Θ_l is also LAC at 0.

□

Remark 5.5. *If in addition the above convex cone C is locally closed at 0, then we can imitate the proofs of (1) and (3) with using Proposition 2.35 (2) and (3) to show that $\mathrm{iri}(C)$ (resp. $\mathrm{qri}(C)$) is LAC at 0 provided that $\mathrm{iri}(C) \neq \emptyset$ and $0 \notin \mathrm{iri}(C)$ (resp. $\mathrm{qri}(C) \neq \emptyset$ and $0 \notin \mathrm{qri}(C)$). Here we omit the details.*

Before we prove the next proposition, let us recall the following evident facts.

Lemma 5.6. *Let $X, X_1, \cdots, X_r$ be topological spaces. Then:*

(1) If $S_1, \cdots, S_q$ are all subsets of X, then $\mathrm{cl}(S_1 \cup \cdots \cup S_p) = \mathrm{cl}(S_1) \cup \cdots \cup \mathrm{cl}(S_q)$.

(2) If $T_1 \subset X_1, \cdots, T_r \subset X_r$, then $\mathrm{cl}(T_1 \times \cdots \times T_r) = \mathrm{cl}(T_1) \times \cdots \times \mathrm{cl}(T_r)$.

The following result is very useful in the subsequent sections.

Proposition 5.7. *Let A_p be a subset of the real Banach space Y_p ($p = 0, \cdots, n$). Suppose $0 \in A_p$ for all p, and let*

$$A \triangleq \left(A_0 \setminus \{0\} \times \prod_{p=1}^{n} A_p \right) \cup \{(0, \cdots, 0)\} \subset \prod_{p=0}^{n} Y_p \triangleq Y. \tag{5.11}$$

Then the following statements hold:

(1) $\mathrm{cl}(A) = \left(\mathrm{cl}(A_0 \setminus \{0\}) \times \prod_{p=1}^{m} \mathrm{cl}(A_p) \right) \cup \{(0, \cdots, 0)\}.$

(2) For any $(c_0, \cdots, c_n) \in \prod_{p=0}^{n} Y_p$,

$$\mathrm{cl}(A) + (c_0, \cdots, c_n) = \left((\mathrm{cl}(A_0 \setminus \{0\}) + c_0) \times \prod_{p=1}^{m} (\mathrm{cl}(A_p) + c_p) \right) \cup \{(c_0, \cdots, c_n)\}.$$

(3) Suppose $I \subset \{0, \cdots, n\}$ *is a nonempty index set containing 0, and* $J = \{0, \cdots, n\} \setminus I$. *Assume that the sets* A_i *are LAC at 0 for* $i \in I$ *while other sets* A_j *are locally closed at 0 for* $j \in J$. *Then the set* A *is LAC at* $(0, \cdots, 0)$.

(4) If $0 \in \mathrm{cl}(A_0 \setminus \{0\})$, *then* $\mathrm{cl}(A) = \prod_{p=0}^{n} \mathrm{cl}(A_p)$.

Proof.

(1) It can be derived using Lemma 5.6.

(2) Given $(c_0, \cdots, c_n) \in \prod_{p=0}^{n} Y_p$, the conclusion follows by the translating the equality in (1) by the element $(c_0, \cdots, c_n)$.

(3) We may assume that $I = \{0, \cdots, m\}$ for some $0 \leq m \leq n$. Since for each $i \in I$, A_i is LAC at 0, there exist a neighborhood V_i of 0 and a sequence $\{c_i^k\} \subset Y_i$ converging to 0 as $k \to \infty$ such that for all $k \in \mathbb{N}$,

$$(\mathrm{cl}(A_0) + c_0^k) \cap V_0 \subset A_0 \setminus \{0\} \tag{5.12}$$

$$(\mathrm{cl}(A_i) + c_i^k) \cap V_i \subset A_i \setminus \{0\} \ (1 \leq i \leq m). \tag{5.13}$$

Hence the sequence $\{c^k\} \subset Y$ given by $c^k = (c_0^k, c_1^k, \cdots, c_m^k, 0, \cdots, 0)$ converges to 0 as $k \to \infty$. On the other hand as A_j is locally closed at 0 for each $j \in J$, there exists a neighborhood V_j of 0 such that

$$\mathrm{cl}(A_j) \cap V_j \subset A_j. \tag{5.14}$$

Thus $V \triangleq \prod_{i=0}^{m} V_i \times \prod_{j=m+1}^{n} V_j$ is a neighborhood of $(0, \cdots, 0)$. As a result,

$$\begin{aligned}
&(\mathrm{cl}(A) + c^k) \cap V \\
&= \left[(\mathrm{cl}(A_0 \setminus \{0\}) + c_0^k) \cap V_0) \times \left(\prod_{i=1}^{m} (\mathrm{cl}(A_i) + c_i^k) \cap V_i \right) \times \left(\prod_{j=m+1}^{n} \mathrm{cl}(A_j) \cap V_j \right) \right] \\
&\quad \cup (\{c_0^k\} \cap V_0) \times \left(\prod_{i=1}^{m} \{c_i^k\} \cap V_i \right) \times \left(\prod_{j=m+1}^{n} V_j \right) \\
&\subset (A_0 \setminus \{0\}) \times \left(\prod_{i=1}^{m} (A_i \setminus \{0\}) \right) \times \left(\prod_{j=m+1}^{n} A_j \right) \\
&\subset A \setminus \{0\},
\end{aligned} \tag{5.15}$$

where the first equality follows by (2) and the first inclusion follows from (5.12), (5.13), (5.14) and the assumptions $0 \in A_p$ for all p. Therefore A is LAC at $(0, \cdots, 0)$.

(4) First observe that by Lemma 5.6 (1) and the assumption $0 \in \mathrm{cl}(A_0 \setminus \{0\})$

$$\mathrm{cl}(A_0) = \mathrm{cl}(A_0 \setminus \{0\} \cup \{0\}) = \mathrm{cl}(A_0 \setminus \{0\}) \cup \{0\} = \mathrm{cl}(A_0 \setminus \{0\}).$$

Then, by (1) we obtain $\mathrm{cl}(A) = \left(\prod_{p=0}^{n} \mathrm{cl}(A_p) \right) \cup \{(0, \cdots, 0)\}$. The desired result follows as $0 \in A_p$ for all p.

□

5.2 First order necessary conditions

5.2.1 Introductory remark

We are going to consider the main results giving the first order necessary conditions for a pair of two problems, one without operator constraints and one with operator constraints, namely,

(i)

$$\begin{aligned} &\text{minimize } F(x) \\ &\text{subject to } x \in \Omega, \end{aligned} \tag{5.16}$$

where X and Y are Banach spaces, $F : X \rightrightarrows Y$ is a multifunction, Ω is a nonempty subset of X and Θ is a relation set in Y containing the origin; and

(ii)

$$\begin{aligned} &\text{minimize } F(x) \\ &\text{subject to } G_i(x) \cap (-\Lambda_i) \neq \emptyset, \quad i = 1, \cdots, m \\ &\qquad x \in \Omega, \end{aligned} \tag{5.17}$$

where $X, Y, Y_1, \cdots, Y_m$ are real Banach spaces, $F : X \rightrightarrows Y$, $G_i : X \rightrightarrows Y_i$ for $i = 1, \cdots, m$ are multifunctions, $\emptyset \neq \Omega \subset X$, $\emptyset \neq \Lambda_i \subset Y_i$ for all $i = 1, \cdots, m$, and Θ is a relation set in Y containing the origin. We let $G_0 \triangleq F$ and $Y_0 \triangleq Y$.

In this section and the next section, the language of variational analysis will be heavily used. Readers may refer to Section 2.5 for the definitions of various derivatives, normal cones and related properties.

Respectively for each of above two problems, we introduce the following definitions for a feasible pair $(\bar{x}, \bar{y}) \in \text{gph}(F)$:

Definition 5.8. *The point $(\bar{x}, \bar{y})$ is*

(1) a fuzzy Fritz John point with respect to (5.16) if it satisfies: for every $\varepsilon > 0$, there exist $(x_1, y_1) \in \text{gph}(\text{cl}(F)) \cap ((\bar{x}, \bar{y}) + \varepsilon\mathbb{B})$, $x_2 \in \text{cl}(\Omega) \cap (\bar{x} + \varepsilon\mathbb{B})$, $y_2 \in \text{cl}(\Theta) \cap \varepsilon\mathbb{B}$ and $(x^*, y^*) \in X^* \times Y^*$ with $\|(x^*, y^*)\| = 1$ such that

$$\begin{aligned} (x^*, -y^*) &\in \hat{N}((x_1, y_1); \text{gph}(\text{cl}(F))) + \varepsilon\mathbb{B}^*, \\ -x^* &\in \hat{N}(x_2; \text{cl}(\Omega)) + \varepsilon\mathbb{B}^*, \\ y^* &\in -\hat{N}(y_2; \text{cl}(\Theta)) + \varepsilon\mathbb{B}^*, \text{ and} \\ 0 &\in \hat{D}^*\text{cl}(F)(x_1, y_1)(y^*) + \hat{N}(x_2; \text{cl}(\Omega)) + \varepsilon\mathbb{B}^*. \end{aligned}$$

(2) a Fritz John point with respect to (5.16) if it satisfies: there exists $(x^*, y^*) \in X^* \times Y^*$ *with* $\|(x^*, y^*)\| = 1$ *such that*

$$\begin{aligned} &x^* \in D^*\mathrm{cl}(F)(\bar{x}, \bar{y})(y^*), -x^* \in N(\bar{x}; \mathrm{cl}(\Omega)), y^* \in -N(0; \mathrm{cl}(\Theta)), \text{ and} \\ &0 \in D^*\mathrm{cl}(F)(\bar{x}, \bar{y})(y^*) + N(\bar{x}; \mathrm{cl}(\Omega)). \end{aligned} \tag{5.18}$$

(3) a Lagrange point with respect to (5.16) if it satisfies: there exists $y^* \in -N(0; \mathrm{cl}(\Theta))$ *with* $\|y^*\| = 1$ *such that*

$$0 \in D^*\mathrm{cl}(F)(\bar{x}, \bar{y})(y^*) + N(\bar{x}; \mathrm{cl}(\Omega)). \tag{5.19}$$

(4) an inverse point with respect to (5.16) if it satisfies: there exists $x^* \in N(\bar{x}; \mathrm{cl}(\Omega))$ *with* $\|x^*\| = 1$ *such that*

$$0 \in D^*\mathrm{cl}(F^{-1})(\bar{y}, \bar{x})(x^*) + N(0; -\mathrm{cl}(\Theta)). \tag{5.20}$$

(5) a fuzzy Fritz John point with respect to (5.17) if it satisfies: for any $\varepsilon > 0$, *there exist* $(x_0, y_0) \in \mathrm{gph}(F) \cap ((\bar{x}, \bar{y}) + \varepsilon\mathbb{B})$, $(x_i, y_i) \in \mathrm{gph}(G_i) \cap ((\bar{x}, \bar{y}_i) + \varepsilon\mathbb{B})$ $(i = 1, \cdots, m)$, $\widetilde{x} \in \Omega \cap (\bar{x} + \varepsilon\mathbb{B})$, $\widetilde{y}_0 \in \Theta \cap (\varepsilon\mathbb{B})$, $\widetilde{y}_i \in (-\Lambda_i) \cap (\bar{y}_i + \varepsilon\mathbb{B})$ $(1 \le i \le m)$ *and* $(x_0^*, \cdots, x_m^*, y_0^*, \cdots, y_m^*) \in (X^*)^{m+1} \times (Y^*)^{m+1}$ *of unit norm such that* $(x_0^*, -y_0^*) \in \hat{N}((x_0, y_0); \mathrm{gph}(F)) + \varepsilon\mathbb{B}^*$, $(x_i^*, -y_i^*) \in \hat{N}((x_i, y_i); \mathrm{gph}(G_i)) + \varepsilon\mathbb{B}^*$ $(1 \le i \le m)$, $y_0^* \in -\hat{N}(\widetilde{y}_0; \Theta) + \varepsilon\mathbb{B}^*$, $y_i^* \in -\hat{N}(-\widetilde{y}_i; \Lambda_i) + \varepsilon\mathbb{B}^*$ $(1 \le i \le m)$, $-\sum_{i=0}^{m} x_i^* \in \hat{N}(\widetilde{x}; \Omega) + \varepsilon\mathbb{B}^*$, *and*

$$0 \in \hat{D}^*F(x_0, y_0)(y_0^*) + \sum_{i=1}^{m} \hat{D}^*G_i(x_i, y_i)(y_i^*) + \hat{N}(\widetilde{x}; \Omega) + \varepsilon\mathbb{B}^*.$$

(6) a Fritz John point with respect to (5.17) if it satisfies: there exists

$$(x_0^*, \cdots, x_m^*, y_0^*, \cdots, y_m^*) \in (X^*)^{m+1} \times \prod_{i=0}^{m} Y_i^*$$

with

$$\|(x_0^*, \cdots, x_m^*, y_0^*, \cdots, y_m^*)\| = 1$$

such that

$$x_0^* \in D^*F(\bar{x},\bar{y})(y_0^*), x_i^* \in D^*G_i(\bar{x},\bar{y}_i)(y_i^*) \ (i=1,\cdots,m), y_0^* \in -N(0;\Theta),$$
$$y_i^* \in -N(-\bar{y}_i;\Lambda_i) \ (i=1,\cdots,m), -\sum_{i=0}^{m} x_i^* \in N(\bar{x};\Omega) \text{ and}$$

$$0 \in D^*F(\bar{x},\bar{y})(y_0^*) + \sum_{i=1}^{m} D^*G_i(\bar{x},\bar{y}_i)(y_i^*) + N(\bar{x};\Omega). \qquad (5.21)$$

(7) *a Lagrange point with respect to (5.17) if it satisfies: there exist $y_0^* \in -N(0;\Theta)$ and $y_i^* \in -N(-\bar{y}_i;\Lambda_i)$ $(i=1,\cdots,m)$ with $\|(y_0^*,\cdots,y_m^*)\| = 1$ such that*

$$0 \in D^*F(\bar{x},\bar{y})(y_0^*) + \sum_{i=1}^{m} D^*G_i(\bar{x},\bar{y}_i)(y_i^*) + N(\bar{x};\Omega). \qquad (5.22)$$

We leave the interpretation of a Fritz John point in (6) of the above definition to the Subsection 5.3.1 by considering a special case of which the functions are defined on $\mathbb{R}^n$ and smooth. Other definitions can be interpreted similarly.

5.2.2 Without operator constraints

In this subsection we prove a main theorem of this chapter, which is a recent result by Bao and Mordukhovich [3, Theorem 4.6] in necessary conditions for a fully local minimizer of the problem (5.16) (cf. Definition 3.12 (1)).

The following two lemmas are consequences of Proposition 5.7 (3) and (4).

Lemma 5.9. *Let Θ and Ω be the sets given in problem (5.16). Given $\bar{x} \in \Omega$. Suppose Θ is LAC at 0. If either Ω is locally closed at $\bar{x}$ or Ω is LAC at $\bar{x}$, then the set (cf. (3.6))*

$$\Theta_{fm} \triangleq ((\bar{x}-\Omega)\times(\Theta\setminus\{0\}))\cup\{(0,0)\}$$

is LAC at $(0,0)$.

Lemma 5.10. *If $0 \in \mathrm{cl}(\Theta\setminus\{0\})$, then $\mathrm{cl}(\Theta_{fm}) = (\bar{x}-\mathrm{cl}(\Omega))\times\mathrm{cl}(\Theta)$, where Θ_{fm} is the set given in Lemma 5.9.*

Remark 5.11. *In Lemma 5.9, the conclusion may not hold if the assumption "Θ is LAC at 0" is dropped. We let $\Omega \triangleq \mathbb{R}$, $\bar{x} \triangleq 0$ and $\Theta \triangleq \{0\} \cup [1, +\infty) \subset \mathbb{R}$. Note Θ is not LAC at 0. Then the above set $\Theta_{fm} = (\mathbb{R} \times [1, +\infty)) \cup \{(0, 0)\}$ which is not LAC at (0,0). In the proof of [3, Theorem 4.6], Bao and Mordukhovich asserted that if Ω is LAC at $\bar{x}$ and Θ is locally closed at 0, then the above set Θ_{fm} is LAC at $(0, 0)$. The example just given indeed refutes their assertion. On the other hand, they also claimed that the conclusion of Lemma 5.10 must hold without mentioning the assumption $0 \in \mathrm{cl}(\Theta \setminus \{0\})$. We see that our example is a counter example to this claim.*

Theorem 5.12. *Let $(\bar{x}, \bar{y})$ be a fully local minimizer to the problem (5.16), where X and Y are Asplund spaces, and the relation set Θ satisfies $0 \in \mathrm{cl}(\Theta \setminus \{0\})$. Assume that either one of the following conditions (1)-(4) is satisfied:*

(1) $\mathrm{gph}(F)$ and Ω are locally closed at $(\bar{x}, \bar{y})$ and $\bar{x}$ respectively. Moreover Θ is LAC at 0.

(2) $\mathrm{gph}(F)$ is locally closed at $(\bar{x}, \bar{y})$, Ω is LAC at $\bar{x}$ and Θ is LAC at 0.

(3) $\mathrm{gph}(F)$ is LAC at $(\bar{x}, \bar{y})$, Ω is locally closed at $\bar{x}$ and Θ is LAC at 0.

(4) $\mathrm{gph}(F)$, Ω and Θ are LAC at $(\bar{x}, \bar{y})$, $\bar{x}$ and 0 respectively.

Then the following versions of necessary conditions for the point $(\bar{x}, \bar{y})$ hold:

A. FUZZY VERSION

The point $(\bar{x}, \bar{y})$ is a fuzzy Fritz John point.

B. FRITZ JOHN VERSION

Assume further that either one of the following assumptions is fulfilled:

(I) Both $\mathrm{cl}(\Omega)$ and $\mathrm{cl}(\Theta)$ are SNC at $\bar{x}$ and 0 respectively.

(II) $\mathrm{cl}(\Theta)$ is SNC at 0 amd $\mathrm{cl}(F)$ is PSNC at $(\bar{x}, \bar{y})$.

(III) $\mathrm{cl}(\Omega)$ is SNC at $\bar{x}$ and $\mathrm{cl}(F^{-1})$ is PSNC at $(\bar{y}, \bar{x})$.

(IV) $\mathrm{cl}(F)$ *is SNC at* $(\bar{x}, \bar{y})$.

Then $(\bar{x}, \bar{y})$ *is a Fritz John point.*

C. LAGRANGE VERSION

Assume either (I) or (II) holds. Suppose also that the mixed qualification condition

$$D_M^*\mathrm{cl}(F)(\bar{x}, \bar{y})(0) \cap (-N(\bar{x}; \mathrm{cl}(\Omega))) = \{0\} \tag{5.23}$$

is satisfied. Then $(\bar{x}, \bar{y})$ *is a Lagrange point.*

D. INVERSE VERSION

Assume either (I) or (III) holds. Suppose also that the inverse qualification condition

$$D_M^*\mathrm{cl}(F^{-1})(\bar{y}, \bar{x})(0) \cap N(0; \mathrm{cl}(\Theta)) = \{0\} \tag{5.24}$$

is satisfied. Then $(\bar{x}, \bar{y})$ *is an inverse point.*

Proof. We let

$$\Theta_{fm} \triangleq ((\bar{x} - \Omega) \times (\Theta \setminus \{0\})) \cup \{(0,0)\}.$$

By Lemma 5.10 under the assumption $0 \in \mathrm{cl}(\Theta \setminus \{0\})$, we get

$$(\bar{x}, \bar{y}) - \mathrm{cl}(\Theta_{fm}) = \mathrm{cl}(\Omega) \times (\bar{y} - \mathrm{cl}(\Theta)). \tag{5.25}$$

As $(\bar{x}, \bar{y})$ is a fully local minimizer for problem (5.16), by Proposition 3.15 it is a local Θ_{fm}-minimal point for $\mathrm{gph}(F)$. If either (1) or (2) holds, then $\mathrm{gph}(F)$ is locally closed at $(\bar{x}, \bar{y})$ and Θ_{fm} is LAC at $(0,0)$ by Lemma 5.9. If either (3) or (4) holds, then by Lemma 5.9 again, $\mathrm{gph}(F)$ and Θ_{fm} are LAC at $(\bar{x}, \bar{y})$ and $(0,0)$ respectively. It follows that any one of cases (1) to (4) implies one of the assumptions in Theorem 5.3 (1) to (3) (putting $A = \mathrm{gph}(F)$ and $\Theta = \Theta_{fm}$ in Theorem 5.3). Hence, by applying Theorem 5.3 and taking account of (5.25), $(\bar{x}, \bar{y})$ is a local extremal point of the set system $\{\mathrm{cl}(\mathrm{gph}(F)), \mathrm{cl}(\Omega) \times (\bar{y} - \mathrm{cl}(\Theta))\}$. By the approximate extremal principle (Theorem 4.3), for all $\varepsilon > 0$, there exist $(x_1, y_1) \in \mathrm{gph}(\mathrm{cl}(F)) \cap ((\bar{x}, \bar{y}) + \varepsilon\mathbb{B})$, $(x_2, \bar{y} - y_2) \in (\mathrm{cl}(\Omega) \times (\bar{y} - \mathrm{cl}(\Theta))) \times ((\bar{x}, \bar{y}) + \varepsilon\mathbb{B})$

and $(x^*, y^*) \in X^* \times Y^*$ with $\|(x^*, y^*)\| = 1$ such that

$$\begin{aligned}(x^*, -y^*) &\in \hat{N}((x_1, y_1); \text{gph}(\text{cl}(F))) + \frac{\varepsilon}{2}\mathbb{B}^*,\\ -x^* &\in \hat{N}(x_2; \text{cl}(\Omega)), \text{ and}\\ y^* &\in \hat{N}(\bar{y} - y_2; \bar{y} - \text{cl}(\Theta)) + \frac{\varepsilon}{2}\mathbb{B}^* = -\hat{N}(y_2; \text{cl}(\Theta)) + \frac{\varepsilon}{2}\mathbb{B}^*.\end{aligned}$$

We arrive at the fuzzy version.

The next step is to prove the Fritz John version. Applying Corollary 4.4 for the above settings, we have for all $k \in \mathbb{N}$, there exist sequences $\{(x_{1k}, y_{1k}, x_{2k}, y_{2k})\} \subset X \times Y \times X \times Y$ and $\{(x^*_{1k}, y^*_{1k}, x^*_{2k}, y^*_{2k})\} \subset X^* \times Y^* \times X^* \times Y^*$ where for all k

$$\begin{aligned}&(x_{1k}, y_{1k}) \in \text{gph}(\text{cl}(F)), x_{2k} \in \text{cl}(\Omega), y_{2k} \in \bar{y} - \text{cl}(\Theta),\\ &(x^*_{1k}, y^*_{1k}) \in \hat{N}((x_{1k}, y_{1k}); \text{gph}(\text{cl}(F))),\\ &x^*_{2k} \in \hat{N}(x_{2k}; \text{cl}(\Omega)) \text{ and } y^*_{2k} \in \hat{N}(y_{2k}; \bar{y} - \text{cl}(\Theta))\end{aligned} \tag{5.26}$$

such that for all k,

$$\begin{aligned}&\|(x_{1k}, y_{1k}) - (\bar{x}, \bar{y}))\| \le k^{-1}, \|x_{2k} - \bar{x}\| \le k^{-1}, \|y_{2k} - \bar{y}\| \le k^{-1}\\ &\|(x^*_{1k}, y^*_{1k}) + (x^*_{2k}, y^*_{2k})\| \le k^{-1} \text{ and}\\ &1 - k^{-1} \le \|(x^*_{1k}, y^*_{1k})\| + \|(x^*_{2k}, y^*_{2k})\| \le 1 + k^{-1}.\end{aligned} \tag{5.27}$$

By the last line of (5.27) we know the sequences $\{(x^*_{1k}, y^*_{2k})\}$ and $\{(x^*_{2k}, y^*_{2k})\}$ are bounded. As X and Y are Asplund spaces, any bounded set in X^* or Y^* is weak* sequentially compact (see Theorem 2.9). Therefore we may assume there exist $x^*_1, x^*_2 \in X^*$ and $y^*_1, y^*_2 \in Y^*$ such that

$$(x^*_{1k}, y^*_{1k}, x^*_{2k}, y^*_{2k}) \xrightarrow{w^*} (x^*_1, x^*_2, y^*_1, y^*_2) \text{ as } k \to \infty. \tag{5.28}$$

The second line of (5.27) implies for all k, $\|x^*_{1k} + x^*_{2k}\| \le k^{-1}$ and $\|y^*_{1k} + y^*_{2k}\| \le k^{-1}$. Passing to limit we get $x^*_1 + x^*_2 = 0$ and $y^*_1 + y^*_2 = 0$. Let $x^* = x^*_1 = -x^*_2$ and $y^* = -y^*_1 = y^*_2$. Using (5.26) and the first line of (5.27) and passing to limit, we arrive at

$$\begin{aligned}&(x^*, -y^*) \in N((\bar{x}, \bar{y}); \text{gph}(\text{cl}(F))) \Leftrightarrow x^* \in D^*\text{cl}(F)(\bar{x}, \bar{y})(y^*),\\ &-x^* \in N(\bar{x}; \text{cl}(\Omega)) \text{ and}\\ &y^* \in N(\bar{y}, \bar{y} - \text{cl}(\Theta)) = N(0; -\text{cl}(\Theta)) = -N(0; \text{cl}(\Theta)).\end{aligned} \tag{5.29}$$

This proves (5.18). It remains to show $\|(x^*, y^*)\| = 1$. As the Mordukhovich normal cone is a cone, if we can show $(x^*, y^*) \neq 0$, then after scaling (x^*, y^*) has unit norm and (5.29) still holds. Then we finish the proof of the Fritz John version. To verify $(x^*, y^*) \neq 0$, we assume that $(x^*, y^*) = 0$. Then $x_1^* = x_2^* = 0$ and $y_1^* = y_2^* = 0$. If (I) holds then $\|x_{2k}^*\| \to 0$ and $\|y_{2k}^*\| \to 0$ as $k \to \infty$. If (II) holds then $\|y_{2k}^*\| \to 0$ as $k \to \infty$ by the SNC assumption and $\|x_{1k}^*\| \to 0$ as $k \to \infty$ by the PSNC assumption. If (III) holds then $\|x_{2k}^*\| \to 0$ as $k \to \infty$ by the SNC assumption and $\|y_{1k}^*\| \to 0$ as $k \to \infty$ by the PSNC assumption. If (IV) holds then $\|(x_{1k}^*, y_{1k}^*)\| \to 0$ as $k \to \infty$ by the SNC assumption. Therefore, also considering the second line of (5.27), if either (I) or (II) or (III) or (IV) holds we obtain $\|(x_{1k}^*, y_{1k}^*, x_{2k}^*, y_{2k}^*)\| \to 0$ as $k \to \infty$ contradicting the third line of (5.27). The proof of the Fritz John version is completed.

Now we prove the Lagrange version. We use the same notations in the proof of Fritz John version. Then by (5.18), $0 \in D^*\mathrm{cl}(F)(\bar{x}, \bar{y})(y^*) + N(\bar{x}; \mathrm{cl}(\Omega))$. Thus (5.23) holds for $y^* \in -N(0; \mathrm{cl}(\Theta))$ (by third line of (5.29). To show $\|y^*\| = 1$, it suffices to show $y^* \neq 0$ (the reason is similar to the case in the proof of the Fritz John version). Suppose $y^* = 0$. If (I) or (II) hold, then $\mathrm{cl}(\Theta)$ is SNC at 0 and thus $\|y_{2k}^*\| \to 0$ as $k \to \infty$. By the second line of (5.27) $\|y_{1k}^*\| \to 0$ as $k \to \infty$. Consequently, from the second line of (5.26) and (5.28), $(x_{1k}^*, y_{1k}^*) \in \hat{N}((x_{1k}, y_{1k}); \mathrm{cl}(\mathrm{gph}(F)))$, $x_{1k}^* \xrightarrow{w^*} x^*$ and $\|y_{1k}^*\| \to 0$ as $k \to \infty$. This implies $x^* \in D_M^*\mathrm{cl}(F)(\bar{x}, \bar{y})(0)$. As $x^* \in -N(\bar{x}; \mathrm{cl}(\Omega))$ by the second line of (5.29), invoking the mixed qualification condition (5.23) we get

$$x^* \in D_M^*\mathrm{cl}(F)(\bar{x}, \bar{y})(0) \cap -N(\bar{x}; \mathrm{cl}(\Omega)) = \{0\}$$

implying $x^* = 0$. Thus $(x^*, y^*) = (0, 0)$, which contradicts the fact that $\|(x^*, y^*)\| = 1$ shown in the proof of the Fritz John version. This completes the proof of the Lagrange version.

For the inverse version, we first convert the problem such that the objective function becomes F^{-1} instead of F. As $(\bar{x}, \bar{y})$ is a local Θ_{fm}-minimal point of $\text{gph}(F)$, there exist neighborhoods U of $\bar{x}$ and V of $\bar{y}$ such that

$$\text{gph}(F) \cap [(\bar{x}, \bar{y}) - ((\bar{x} - \Omega) \times (\Theta \setminus \{0\})) \cup \{(0,0)\}] \cap (U \times V) = \{(\bar{x}, \bar{y})\}$$

implying

$$\text{gph}(F^{-1}) \cap [(\bar{y}, \bar{x}) - ((\Theta \setminus \{0\}) \times (\bar{x} - \Omega)) \cup \{(0,0)\}] \cap (V \times U) = \{(\bar{y}, \bar{x})\}.$$

Thus $(\bar{y}, \bar{x})$ is a local $\widetilde{\Theta}$-minimal point to the set $\text{gph}(F^{-1})$ where

$$\widetilde{\Theta} = ((\Theta \setminus \{0\}) \times (\bar{x} - \Omega)) \cup \{(0,0)\}.$$

Put another way, $(\bar{y}, \bar{x})$ is a fully local $(\bar{x} - \Omega)$-minimizer of the following "inverse" problem:

$$\begin{aligned} &\text{minimize } F^{-1}(y) \\ &\text{subject to } y \in \bar{y} - \Theta \end{aligned} \tag{5.30}$$

where $F^{-1} : Y \rightrightarrows X$ and the relation set on X is $\bar{x} - \Omega$. Now we consider the whole theorem for this new problem. We observe that if one of the assumptions (1) to (4) and one of (I) and (III) hold with respect to problem (5.16), then one of the assumptions (1) to (4) and one of (I) and (II) hold with respect to problem (5.30). However to apply the Lagrange version we need to check (5.23) is true with respect to problem (5.30). Notice $N(0; \text{cl}(\Theta)) = -N(0; -\text{cl}(\Theta)) = -N(\bar{y}, \bar{y} - \text{cl}(\Theta)) = -N(\bar{y}; \text{cl}(\bar{y} - \Theta))$. Therefore the left hand side of (5.23) for the problem (5.30) is

$$\begin{aligned} &D_M^*\text{cl}(F^{-1})(\bar{y}, \bar{x})(0) \cap -N(\bar{y}, \bar{y} - \text{cl}(\bar{y} - \text{cl}(\Theta))) \\ &= D_M^*\text{cl}(F^{-1})(\bar{y}, \bar{x})(0) \cap N(0; \text{cl}(\Theta)) = \{0\} \end{aligned}$$

by (5.24). Hence using the Lagrange version for the problem (5.30) we deduce that there exists

$$x^* \in -N(0; \text{cl}(\bar{x} - \Omega)) = -N(0; \bar{x} - \text{cl}(\Omega)) = N(\bar{x}; \text{cl}(\Omega))$$

of unit norm such that

$$\begin{aligned} 0 \in& D^*\mathrm{cl}(F^{-1})(\bar{y},\bar{x})(x^*) + N(\bar{y}, \mathrm{cl}(\bar{y}-\Theta)) \\ =& D^*\mathrm{cl}(F^{-1})(\bar{y},\bar{x})(x^*) + N(0; -\mathrm{cl}(\Theta)) \end{aligned}$$

which is precisely (5.20). We thus finish the proof of the inverse version and also the whole proof. □

If the constraint set $\Omega = X$, then applying the previous theorem we obtain the following generalized Fermat rule.

Corollary 5.13. *Consider the problem (5.16) where X and Y are Asplund spaces,* $\mathrm{gph}(F)$ *is locally closed at $(\bar{x},\bar{y})$, Θ are locally closed and LAC at 0, and $\Omega = X$. Given $(\bar{x},\bar{y}) \in$* $\mathrm{gph}(F)$. *Assume either Θ is SNC at 0 or F^{-1} is PSNC at $(\bar{y},\bar{x})$. Suppose $(\bar{x},\bar{y})$ is a fully local minimizer of the problem. Then*

$$0 \in D^*F(\bar{x},\bar{y})(y^*) \text{ for some } y^* \in -N(0;\Theta) \text{ with } \|y^*\| = 1.$$

Proof. By the given assumptions, the assumption (1) in Theorem 5.12 is satisfied. Clearly $\Omega = X$ is SNC at $\bar{x}$, so does its closure. Owing to the SNC property of Θ at 0, by Proposition 2.45, $\mathrm{cl}(\Theta)$ is also SNC at 0. This implies (I) in Theorem 5.12 holds. If F^{-1} is PSNC at $(\bar{y},\bar{x})$, then since $\mathrm{gph}(F^{-1})$ is locally closed at $(\bar{y},\bar{x})$, we have $\mathrm{cl}(F^{-1})$ is also PSNC at $(\bar{y},\bar{x})$. The condition (III) in Theorem 5.12 is satisfied. As a result, we can apply the Fritz John version of Theorem 5.12 that there exists $(x^*,y^*) \in X^* \times Y^*$ with $\|(x^*,y^*)\| = 1$ such that

$$\begin{aligned} x^* &\in D^*\mathrm{cl}(F)(\bar{x},\bar{y})(y^*) = D^*F(\bar{x},\bar{y})(y^*) \\ -x^* &\in N(\bar{x};\mathrm{cl}(\Omega)) = N(\bar{x};X) = \{0\} \text{ so } x^* = 0, \textit{and} \\ y^* &\in -N(0;\mathrm{cl}(\Theta)) = -N(0;\Theta) \end{aligned}$$

(cf. Proposition 2.40 (5)). The proof is thus completed. □

Indeed the well-known situation [25, Proposition 1.114] is a corollary [3, Remark after Corollary 4.9] of the previous result.

Corollary 5.14. *Given a lower semicontinuous function $f : X \to \overline{\mathbb{R}}$ where X is Asplund and let $\bar{x} \in \operatorname{dom}(f)$ be a minimizer of f. Then $0 \in \partial f(\bar{x})$.*

Proof. By applying Corollary 5.13 to the multifunction $F : X \rightrightarrows \mathbb{R}$ given by $F(x) = \{\alpha \in \mathbb{R} : f(x) \leq \alpha\}$ ($\operatorname{gph}(F) = \operatorname{epi}(f)$), we obtain $0 \in D^*F(\bar{x}, f(\bar{x}))(1) \triangleq \partial f(\bar{x})$. □

5.2.3 With operator constraints

In this subsection we discuss the vector optimization problem (5.17) consisting of both a geometric constraint and finitely many operator constraints. First we introduce what is meant by a fully local minimizer of the problem (5.17).

Definition 5.15. *The point $(\bar{x}, \bar{y}_0) \in \operatorname{gph}(F)$ is a fully local Θ-minimizer of the problem (5.17) if it is a fully local Θ-minimizer of problem (5.16) where the constraint set is*

$$\Omega \cap \bigcap_{i=1}^{m} G_i^{-1}(-\Lambda_i) = \{x \in X : x \in \Omega \text{ and } G_i(x) \cap (-\Lambda_i) \neq \emptyset \text{ for all } i = 1, \cdots, m\}.$$

To solve the problem (5.17) we want to transform it to a problem in the form of problem (5.16) of which the constraint set is kept to be the original constraint set Ω while F and Θ are replaced. If this can be done, the objective function and the relation set should be related to G_i's, Θ and Λ_i's. Then we can apply the previous results to this case. Indeed we have the following result.

Proposition 5.16. *If $(\bar{x}, \bar{y}_0)$ is a fully local Θ-minimizer of the problem (5.17) and let $\bar{y}_i \in G_i(\bar{x}) \cap (-\Lambda_i)$ $(i = 1, \cdots, m)$. Then $(\bar{x}, \bar{y}_0, \cdots, \bar{y}_m)$ is a fully local $\widetilde{\Theta}$-minimizer of the problem*

$$\begin{aligned} &\text{minimize } \widetilde{F}(x) \\ &\text{subject to } x \in \Omega, \end{aligned} \tag{5.31}$$

where $\widetilde{F} : X \rightrightarrows \prod_{i=0}^{m} Y_i$ and $\widetilde{\Theta} \subset \prod_{i=0}^{m} Y_i$ are respectively defined by

$$\widetilde{F}(x) \triangleq F(x) \times G_1(x) \times \cdots G_m(x) \text{ for all } x \in X \tag{5.32}$$

and

$$\widetilde{\Theta} \triangleq (\Theta \setminus \{0\} \times (\Lambda_1 + \bar{y}_1) \times \cdots \times (\Lambda_m + \bar{y}_m)) \cup \{(0, 0, \cdots, 0)\}. \tag{5.33}$$

Proof. As $(\bar{x}, \bar{y}_0)$ is a fully local Θ-minimizer of the problem (5.17), by definition there exist neighborhoods U of $\bar{x}$ and V of $\bar{y}_0$ such that

$$F(\widetilde{\Omega} \cap U) \cap (\bar{y}_0 - \Theta) \cap V = \{\bar{y}_0\} \tag{5.34}$$

where $\widetilde{\Omega} \triangleq \Omega \cap \bigcap_{i=1}^{m} G_i^{-1}(-\Lambda_i)$. We are going to show that

$$\widetilde{F}(\Omega \cap U) \cap ((\bar{y}_0, \cdots, \bar{y}_n) - \widetilde{\Theta}) \cap (V \times \prod_{i=1}^{m} Y_i) = \{(\bar{y}_0, \cdots, \bar{y}_m)\} \tag{5.35}$$

where $\widetilde{F}$ and $\widetilde{\Theta}$ are defined in the statement. Let $(y_0, \cdots, y_m)$ belongs to the left hand side of (5.35). To establish (5.35), we only need to show that

$$(y_0, \cdots, y_m) = (\bar{y}_0, \cdots, \bar{y}_m). \tag{5.36}$$

To do this, we note that $(y_0, \cdots, y_m) \in (\bar{y}_0, \cdots, \bar{y}_m) - \widetilde{\Theta}$,

$$y_0 \in V \tag{5.37}$$

and there exists $x \in \Omega \cap U$ such that

$$y_0 \in F(x), y_i \in G_i(x), \ i = 1, \cdots, m.$$

By (5.33) it follows that unless (5.36) holds,

$$\bar{y}_0 - y_0 \in \Theta \setminus \{0\} \tag{5.38}$$

and $\bar{y}_i - y_i \in \Lambda_i + \bar{y}_i$ $(1 \leq i \leq m)$ and so $x \in G_i^{-1}(-\Lambda_i)$ for each i. Consequently $x \in \widetilde{\Omega} \cap U$ and $y_0 \in F(\widetilde{\Omega} \cap U)$. By (5.37) and (5.38) it follows from (5.34) that $y_0 = \bar{y}_0$ contradicting $y_0 \neq \bar{y}_0$ (see (5.38)). This means that (5.36) must hold and the proof is completed. □

Remark 5.17. *Bao and Mordukhovich [3, Proposition 5.1(a)] provided the above result for the case* $m = 1$ *except the relation set* $\widetilde{\Theta}$ *is replaced by* $\Theta_0 \triangleq \Theta \times (\Lambda_1 + \bar{y}_1)$. *The following example shows that their result is not correct. We take* $X = Y = Y_0 = \Omega \triangleq \mathbb{R}$, $\bar{x} \triangleq 0$, $F \triangleq f$ *where* $f(x) = |x|$, $G_1 \triangleq g_1$ *where* $g_1(x) = x$, $\Theta = \mathbb{R}_+$ *and* $\Lambda_1 = \mathbb{R}$. *Then* 0 *is a fully local minimizer of the problem (5.17), but* $(0, 0, 0)$ *is not a* Θ_0*-minimal point of the problem (5.31).*

For the following three lemmas, the first one is a consequence of Proposition 5.7 (3), the second one is a a consequence of Proposition 5.7 (4), and the third one can be easily checked.

Lemma 5.18. *Under the same notations as Proposition 5.16, suppose that* Θ *and* Λ_i $(i = 1, \cdots, m)$ *are locally closed at 0 and* $-\bar{y}_i$ $(i = 1, \cdots, m)$ *respectively. If* Θ *is LAC at 0, then* $\widetilde{\Theta}$ *is LAC at 0.*

Lemma 5.19. *Under the same notations as Proposition 5.16, if* $0 \in \mathrm{cl}(\Theta \setminus \{0\})$, *then* $\mathrm{cl}(\widetilde{\Theta}) = \mathrm{cl}(\Theta) \times \prod_{i=1}^{m} (\mathrm{cl}(\Lambda_i) + \bar{y}_i)$.

Lemma 5.20. *Let* $S_1, \cdots, S_n$ *be subsets of a vector space* X. *Then for any* $x_0 \in X$, $\left(\bigcap_{i=1}^{n} S_i \right) + x_0 = \bigcap_{i=1}^{n} (S_i + x_0)$.

The following result modifies the one proved by Bao and Mordukhovich [3, Theorem 5.2] of which the proof is not the original one given there. Bao and Mordukhovich only considered one operator constraint and the necessary conditions are established through coderivative sum rule and chain rule (see Chapter 3 of [25]). The PSNC assumptions of the mappings and some constraint qualifications are crucial in order to apply those rules. As a result, they formulated only the Lagrange version (where $m = 1$) of the following theorem. Nevertheless, we consider m operator constraints, use the extremal principle where the set system concerned can have more than two sets, and provide the fuzzy, Fritz John, Lagrange version of necessary conditions like Theorem 5.12. Furthermore, our result is more useful to generalize the known results which will be discussed in the next section.

Theorem 5.21. *Let $(\bar{x}, \bar{y}_0)$ be a fully local Θ-minimizer of the problem (5.17), where X, Y and Y_i (for all i) are Asplund spaces, and the relation set Θ satisfies $0 \in \operatorname{cl}(\Theta \setminus \{0\})$. Also let $\bar{y}_i \in G_i(\bar{x}) \cap (-\Lambda_i)$ for each $i = 1, \cdots, m$. Suppose the following assumptions are true:*

(1) $\Omega, \Theta, \Lambda_i, \operatorname{gph}(F)$ and $\operatorname{gph}(G_i)$ (for all i) are locally closed at $\bar{x}$, 0, $-\bar{y}_i$, $(\bar{x}, \bar{y}_0)$ and $(\bar{x}, \bar{y}_i)$ (for all i) respectively.

(2) Θ is LAC at 0.

Then the following versions of necessary conditions for the point $(\bar{x}, \bar{y}_0, \cdots, \bar{y}_m)$ hold:

<u>A. Fuzzy version</u>

The point $(\bar{x}, \bar{y}_0)$ is a fuzzy Fritz John point.

<u>B. Fritz John version</u>

Assume further that either one of the following is satisfied:

(I) Θ is SNC at 0, Λ_i is SNC at $-\bar{y}_i$ for all i, and the mappings F and G_i (for all i) are PSNC at $(\bar{x}, \bar{y}_0)$ and $(\bar{x}, \bar{y}_i)$ (for all i) respectively.

(II) Ω is SNC at $\bar{x}$, Θ is SNC at 0, Λ_i is SNC at $-\bar{y}_i$ for all i, and there exist m mappings from the set $\{G_i : i = 0, \cdots, m\}$ so that they are PSNC at $(\bar{x}, \bar{y}_i)$ respectively.

Then $(\bar{x}, \bar{y}_0)$ is a Fritz John point.

<u>C. Lagrange version</u>

Assume either (I) or (II) holds. Suppose also that the following mixed qualification conditions are satisfied:

$$\left(D^*_M F(\bar{x}, \bar{y}_0)(0) + \sum_{i=1}^{m} D^*_M G_i(\bar{x}, \bar{y}_i)(0) \right) \cap (-N(\bar{x}; \Omega)) = \{0\}, \text{ and} \tag{5.39}$$

$$(D_M^* G_i(\bar{x},\bar{y}_i)(0)) \cap \left(\sum_{j=i+1}^{m} -D_M^* G_j(\bar{x},\bar{y}_j)(0) \right) = \{0\} \text{ for all } i = 0, \cdots, m-1. \tag{5.40}$$

Then $(\bar{x},\bar{y}_0)$ *is a Lagrange point.*

Proof. By Proposition 5.16, $(\bar{x},\bar{y}_0,\bar{y}_1,\cdots,\bar{y}_m)$ is a fully local $\widetilde{\Theta}$-minimizer of the problem (5.31) where $\widetilde{F}$ and $\widetilde{\Theta}$ are given in (5.32) and (5.33) respectively. From Proposition 3.15, Lemma 5.18 and Theorem 5.3 (1) we see that under the assumptions (1) and (2), $(\bar{x},\bar{y}_0,\bar{y}_1,\cdots,\bar{y}_m)$ is a local extremal point of the set system

$$\{\mathrm{cl}(\mathrm{gph}(\widetilde{F})), (\bar{x},\bar{y}_0,\bar{y}_1,\cdots,\bar{y}_m) - \mathrm{cl}(\Theta')\}$$

where $\Theta' \triangleq ((\bar{x}-\Omega)\times\widetilde{\Theta}\setminus\{0\})\cup\{(0,0)\}$. As $\mathrm{cl}(\Theta') = (\bar{x}-\mathrm{cl}(\Omega))\times\mathrm{cl}(\Theta)\times\prod_{i=1}^{m}(\mathrm{cl}(\Lambda_i)+\bar{y}_i)$ by Lemma 5.19, $(\bar{x},\bar{y}_0,\bar{y}_1,\cdots,\bar{y}_m)$ is a local extremal point of the set system $\{\mathrm{cl}(\mathrm{gph}(\widetilde{F})),\mathrm{cl}(B)\}$ where $B \triangleq \Omega\times(\bar{y}_0-\Theta)\times\prod_{i=1}^{m}(-\Lambda_i)$. Note that $\mathrm{gph}(\widetilde{F}) = \bigcap_{i=0}^{m} A_i$ where $A_i \triangleq \{(x_0,y_0,\cdots,y_m)\in X\times\prod_{j=0}^{m}Y_j : (x_0,y_i)\in \mathrm{cl}(\mathrm{gph}(G_i))\}$ for all i. Then by Lemma 5.20, we see that $(\bar{x},\bar{y}_0,\bar{y}_1,\cdots,\bar{y}_m)$ is a local extremal point of the set system $\{A_0,\cdots,A_m,\mathrm{cl}(B)\}$. For a given $\varepsilon>0$, by applying the approximate extremal principle (Theorem 4.3) with considering the assumption (1) and the fact that for all $i=0,\cdots,m$,

$$\begin{aligned}\hat{N}((\bar{x},\bar{y}_0,\cdots,\bar{y}_m);A_i) = \{&(x^*,y_0^*,\cdots,y_m^*)\in X^*\times\prod_{j=0}^{m}Y_i^* : \\ &(x^*,y_i^*)\in\hat{N}((\bar{x},\bar{y}_i);\mathrm{cl}(\mathrm{gph}(G_i))) \text{ and } y_j^*=0 \text{ for } j\neq i\},\end{aligned} \tag{5.41}$$

there are sequences

$$\begin{aligned}&\{(x_0,x_1,\cdots,x_m,\widetilde{x})\}\subset (X^*)^{m+2},\\ &\{(y_0,y_1,\cdots,y_m,\widetilde{y}_0,\widetilde{y}_1,\cdots,\widetilde{y}_m)\}\subset\left(\prod_{i=0}^{m}Y_i\right)^2,\\ &\{(x_0^*,x_1^*,\cdots,x_m^*,\widetilde{x}^*)\}\subset (X^*)^{m+2} \text{ and}\\ &\{(y_0^*,y_k^*,\cdots,y_k^*,\widetilde{y}_0^*,\widetilde{y}_1^*,\cdots,\widetilde{y}_m^*)\}\subset\left(\prod_{i=0}^{m}Y_i^*\right)^2\end{aligned}$$

such that

$$
\begin{aligned}
&(x_i, y_i) \in \mathrm{gph}(G_i) \cap ((\bar{x}, \bar{y}_i) + \varepsilon\mathbb{B}) \ (i = 0, \cdots, m), \\
&(\widetilde{x}_0, \bar{y}_0 - \widetilde{y}_0, \cdots, \widetilde{y}_m) \in B \cap ((\bar{x}, \bar{y}_0, \cdots, \bar{y}_m) + \varepsilon\mathbb{B}), \\
&(x_i^*, y_i^*) \in \hat{N}((x_i, y_i); \mathrm{gph}(G_i)) + \frac{\varepsilon}{8(m+1)(m+2)}\mathbb{B}^* \ (i = 0, \cdots, m) \text{ and} \\
&(\widetilde{x}^*, \widetilde{y}_0^*, \cdots, \widetilde{y}_m^*) \in \hat{N}((\widetilde{x}_0, \bar{y} - \widetilde{y}_0, \cdots, \widetilde{y}_m); B) + \frac{\varepsilon}{8(m+1)}\mathbb{B}^*
\end{aligned} \tag{5.42}
$$

with

$$
\begin{aligned}
&\left(\sum_{i=0}^{m} x_i^*\right) + \widetilde{x}^* = 0, y_i^* = -\widetilde{y}_i^* \text{ for all } i = 0, \cdots, m, \text{ and} \\
&\left(\sum_{i=0}^{m} \|(x_i^*, y_i^*)\|\right) + \|(\widetilde{x}^*, \widetilde{y}_0^*, \cdots, \widetilde{y}_m^*)\| = 1.
\end{aligned} \tag{5.43}
$$

Let $\nu \triangleq \|(x_0^*, \cdots, x_m^*, \widetilde{y}_0^*, \cdots, \widetilde{y}_m^*)\| = \max\{\|x_0^*\|, \cdots, \|x_m^*\|, \|\widetilde{y}_0^*\|, \cdots, \|\widetilde{y}_m^*\|\}$. From (5.43) and the triangle inequality, we have

$$
\begin{aligned}
1 &= \sum_{i=0}^{m} \max\{\|x_i^*\|, \|y_i^*\|\} + \max\{\|\widetilde{x}^*\|, \|\widetilde{y}_0^*\|, \cdots, \|\widetilde{y}_m^*\|\} \\
&\le \sum_{i=0}^{m} (\|x_i^*\| + \|y_i^*\|) + \|\widetilde{x}^*\| + \sum_{i=0}^{m} \|\widetilde{y}_i^*\| \\
&\le \sum_{i=0}^{m} (\|x_i^*\| + \|\widetilde{y}_i^*\|) + \sum_{i=0}^{m} \|x_i^*\| + \sum_{i=0}^{m} \|\widetilde{y}_i^*\| \\
&\le 4(m+1)\nu
\end{aligned}
$$

and so $\nu \neq 0$ and

$$
\nu \ge \frac{1}{4(m+1)}. \tag{5.44}
$$

We let $\mathcal{X}_i^* = \frac{x_i^*}{\nu}$ and $\mathcal{Y}_i^* = \frac{\widetilde{y}_i^*}{\nu}$ $(0 \le i \le m)$. It follows from (5.42), (5.43) and (5.44) that

$$
\begin{aligned}
&\|(\mathcal{X}_0^*, \mathcal{X}_1^*, \cdots, \mathcal{X}_m^*, \mathcal{Y}_0^*, \mathcal{Y}_1^*, \cdots, \mathcal{Y}_m^*)\| = 1, \\
&(\mathcal{X}_i^*, -\mathcal{Y}_i^*) \in \hat{N}((x_i, y_i); \mathrm{gph}(G_i)) + \frac{\varepsilon}{2(m+2)}\mathbb{B}^* \ (i = 0, \cdots, m) \\
&\left(\frac{\widetilde{x}^*}{\nu}, \mathcal{Y}_0^*, \cdots, \mathcal{Y}_m^*\right) \in \hat{N}((\widetilde{x}_0, \bar{y} - \widetilde{y}_0, \cdots, \widetilde{y}_m); B) + \frac{\varepsilon}{2(m+2)}\mathbb{B}^* \text{ and} \\
&\left(\sum_{i=0}^{m} \mathcal{X}_i^*\right) + \frac{\widetilde{x}^*}{\nu} = 0.
\end{aligned}
$$

As a result, $\mathcal{X}_i^* \in \hat{D}^*G(x_i, y_i)(\mathcal{Y}_i^* + \varepsilon\mathbb{B}^*) + \dfrac{\varepsilon}{2(m+2)}\mathbb{B}^*$, $-\sum_{i=0}^{m} \mathcal{X}_i^* \in \hat{N}(\widetilde{x}_0; \Omega) + \dfrac{\varepsilon}{2}\mathbb{B}^*$ $(0 \le i \le m)$, $\mathcal{Y}_0^* \in -\hat{N}(\widetilde{y}_0; \Theta) + \varepsilon\mathbb{B}^*$ and $\mathcal{Y}_i^* \in -\hat{N}(-\widetilde{y}_i; \Lambda_i) + \varepsilon\mathbb{B}^*$ $(1 \le i \le m)$. Therefore the fuzzy version is seen to be true.

By invoking Corollary 4.4 for the extremal system $\{A_0, A_1, \cdots, A_m, \mathrm{cl}(B); (\bar{x}, \bar{y}_0, \cdots, \bar{y}_m)\}$ with considering the local closedness assumption (1) and (5.41) for $i = 0, \cdots, m$, there exist sequences

$$\begin{aligned}
&\{(x_{0k}, x_{1k}, \cdots, x_{mk}, \widetilde{x}_k)\} \subset (X^*)^{m+2},\\
&\{(y_{0k}, y_{1k}, \cdots, y_{mk}, \widetilde{y}_{0k}, \widetilde{y}_{1k}, \cdots, \widetilde{y}_{mk})\} \subset \left(\prod_{i=0}^{m} Y_i\right)^2,\\
&\{(x_{0k}^*, x_{1k}^*, \cdots, x_{mk}^*, \widetilde{x}_k^*)\} \subset (X^*)^{m+2} \text{ and}\\
&\{(y_{0k}^*, y_{1k}^*, \cdots, y_{mk}^*, \widetilde{y}_{0k}^*, \widetilde{y}_{1k}^*, \cdots, \widetilde{y}_{mk}^*)\} \subset \left(\prod_{i=0}^{m} Y_i^*\right)^2
\end{aligned}$$

such that

$$\begin{aligned}
&(x_i, y_{ik}) \in \mathrm{gph}(G_i) \cap ((\bar{x}, \bar{y}_i) + \varepsilon\mathbb{B}) \ (0 \le i \le m),\\
&\widetilde{x}_k \in \Omega, \widetilde{y}_{0k} \in \Theta, \widetilde{y}_{ik} \in -\Lambda_i \ (1 \le i \le m),\\
&(\widetilde{x}_k, \bar{y}_0 - \widetilde{y}_{0k}, \widetilde{y}_{1k}, \cdots, \widetilde{y}_{mk}) \in (\bar{x}, \bar{y}_0, \cdots, \bar{y}_k) + \varepsilon\mathbb{B},\\
&(x_{ik}^*, y_{ik}^*) \in \hat{N}((x_{ik}, y_{ik}); \mathrm{gph}(G_i)) \ (0 \le i \le m),\\
&\widetilde{x}_k^* \in N(\widetilde{x}_k; \Omega), \widetilde{y}_{0k}^* \in -\hat{N}(\widetilde{y}_{0k}; \Theta) \text{ and } \widetilde{y}_{ik}^* \in -\hat{N}(-\widetilde{y}_{ik}; \Lambda_i) \ (1 \le i \le m)
\end{aligned} \tag{5.45}$$

with

$$\max\left\{\left\|\left(\sum_{i=0}^{m} x_{ik}^*\right) + \widetilde{x}_k^*\right\|, \max_{0 \le i \le m} \|y_{ik}^* + \widetilde{y}_{ik}^*\|\right\} \le \frac{1}{k} \quad \text{and} \tag{5.46}$$

$$1 - \frac{1}{k} \le \left(\sum_{i=0}^{m} \|(x_{ik}^*, y_{ik}^*)\|\right) + \|(\widetilde{x}_k^*, \widetilde{y}_{0k}^*, \cdots, \widetilde{y}_{mk}^*)\| \le 1 + \frac{1}{k}. \tag{5.47}$$

It follows from (5.47) that $\{(x_{ik}^*, y_{ik}^*)\}$ $(0 \le i \le m)$ and $\{(\widetilde{x}_k^*, \widetilde{y}_{0k}^*, \cdots, \widetilde{y}_{mk}^*)\}$ are bounded sequences, therefore by Theorem 2.9, we may assume that there exist $x_i^*, \widetilde{x}^* \in X^*$ and $y_i^*, \widetilde{y}^* \in Y^*$ $(0 \le i \le m)$ such that

$$x_{ik}^* \xrightarrow{w^*} x_i^*, \ \widetilde{x}_k^* \xrightarrow{w^*} \widetilde{x}^* \ y_{ik}^* \xrightarrow{w^*} y_i^*, \ \widetilde{y}_{ik}^* \xrightarrow{w^*} \widetilde{y}_i^* \text{ as } k \to \infty. \tag{5.48}$$

Passing to limit $k \to \infty$ in (5.46) we get

$$\left(\sum_{i=0}^{m} x_i^*\right) + \widetilde{x}^* = 0 \text{ and } y_i^* = -\widetilde{y}_i^* \text{ for all } i = 0, \cdots, m. \tag{5.49}$$

Therefore, by (5.49) and passing to limit $k \to \infty$ in (5.45), we obtain

$$\begin{aligned} & x_i^* \in D^*G_i(\bar{x}, \bar{y}_i)(\widetilde{y}_i^*) \ (0 \le i \le m), -\sum_{i=0}^{m} x_i^* \in N(\bar{x}; \Omega), \\ & \widetilde{y}_0^* \in -N(0; \Theta), \widetilde{y}_i^* \in -N(-\bar{y}_i; \Lambda_i) \ (1 \le i \le m). \end{aligned} \tag{5.50}$$

It remains to show that $(x_0^*, x_1^*, \cdots, x_m^*, \widetilde{y}_0^*, \widetilde{y}_1^*, \cdots, \widetilde{y}_m^*) \neq 0$. Suppose the contrary that $(x_0^*, x_1^*, \cdots, x_m^*, \widetilde{y}_0^*, \widetilde{y}_1^*, \cdots, \widetilde{y}_m^*) = 0$. Then $\widetilde{x}^* = 0$ and $y_i^* = 0$ for $i = 0, \cdots, m$ by (5.49). Using (5.46), we see that either (I) or (II) implies $\|x_{ik}^*\| \to 0$, $\|\widetilde{x}^*\| \to 0$, $\|y_{ik}^*\| \to 0$ and $\|\widetilde{y}_{ik}^*\| \to 0$ as $k \to \infty$ $(0 \le i \le m)$. This contradicts (5.47). We complete the proof of the Fritz John version.

For proving the Lagrange version, firstly, under the same notations as the proof of the Fritz John version, we obtain from (5.50) that

$$0 \in D^*F(\bar{x}, \bar{y}_0)(\widetilde{y}_0^*) + \sum_{i=1}^{m} D^*G_i(\bar{x}, \bar{y}_i)(\widetilde{y}_i^*) + N(\bar{x}; \Omega)$$

where $\widetilde{y}_0^* \in -N(0; \Theta)$ and $\widetilde{y}_i^* \in -N(-\bar{y}_i; \Lambda_i)$ $(i = 1, \cdots, m)$. It suffices to show that $(\widetilde{y}_0^*, \cdots, \widetilde{y}_m^*) \neq 0$. Suppose that $(\widetilde{y}_0^*, \cdots, \widetilde{y}_m^*) = 0$. Then by (5.46), we see that (I) or (II) implies $\|\widetilde{y}_{ik}^*\| \to 0$ and $\|y_{ik}^*\| \to 0$ as $k \to \infty$ $(0 \le i \le m)$. Hence by the definition of mixed coderivative we have $x_i^* \in D_M^*G_i(\bar{x}, \bar{y}_i)(0)$ for all $i = 0, \cdots, m$. Therefore by (5.45), (5.50), the first part of (5.49) and the mixed qualification condition (5.39) we know that

$$-\widetilde{x}^* \in \left(\sum_{i=0}^{m} D_M^*G_i(\bar{x}, \bar{y}_i)(0)\right) \cap (-N(\bar{x}; \Omega)) = \{0\}$$

and so $\widetilde{x}^* = 0$. Hence $\sum_{i=0}^{m} x_i^* = 0$. It follows from another mixed qualification condition (5.40) that $x_i^* = 0$ for all $i = 0, \cdots, m$. Then $(x_0^*, x_1^*, \cdots, x_m^*, \widetilde{y}_0^*, \widetilde{y}_1^*, \cdots, \widetilde{y}_m^*) = 0$ which is a contradiction. The Lagrange version follows.

□

5.3 Comparisons with known necessary conditions

Theorem 5.12 and Theorem 5.21 are the main results of this chapter. They are new at least in the following four aspects (some points are suggested in [3, Remark 4.11]):

(1) We consider fully local minimizers which are more general than partially local minimizers (or local minimizers) seen in previous literature.

(2) We use the LAC assumptions instead of closedness/ convexity/ pointedness of the ordering cones. Hence some special orderings (e.g. lexicographical order, preference relation in economics) can be taken into our consideration.

(3) The assumptions on the functions, constraint sets and ordering cones are local assumptions (e.g. LAC property, locally closedness, SNC property, PSNC property) instead of global ones.

(4) In Theorem 5.12, the first order necessary condition can be formulated using the inverse multifunction of the objective mapping.

In this section we present some previous results in the literature. They are also quoted in [3, Remark 4.11]. We will see the above improvements from the following results, and we shall discuss each of them in detail.

5.3.1 Finite-dimensional setting

Let us begin with a simple case which is the well-known differentiable nonlinear programming problem under the Euclidean space setting:

$$\begin{aligned} &\text{minimize } f(x) \\ &\text{subject to } g_i(x) \leq 0, \quad i = 1, \cdots, p \\ &\qquad\qquad g_i(x) = 0, \quad i = p+1, \cdots, m \\ &\qquad\qquad x \in \Omega. \end{aligned} \tag{5.51}$$

where Ω is a nonempty open convex set in $\mathbb{R}^n$, $f : \mathbb{R}^n \to \mathbb{R}$ and $g_i : \mathbb{R}^n \to \mathbb{R}$ (for all i) are C^1 functions.

The renowned Fritz-John necessary conditions for optimality (by Fritz John in 1948) is a special case of our results.

Theorem 5.22. *[4, P. 146, Theorem 5.1] Let $\bar{x} \in \Omega$ be a local minimizer of the problem (5.51). Then there exist real numbers $\mu_0, \mu_1, \cdots, \mu_p, \lambda_{p+1}, \cdots, \lambda_m$ such that*

(1) $g_i(\bar{x}) \leq 0$ for $i = 1, \cdots, p$; $g_i(\bar{x}) = 0$ for $i = p+1, \cdots, m$,

(2) $(\mu_0, \mu_1, \cdots, \mu_p, \lambda_{p+1}, \cdots, \lambda_m) \neq 0$,

(3)

$$\mu_0 \nabla f(\bar{x}) + \sum_{i=1}^{p} \mu_i \nabla g_i(\bar{x}) + \sum_{i=p+1}^{m} \lambda_i \nabla g_i(\bar{x}) = 0, \tag{5.52}$$

(4) $\mu_i \geq 0$ for $i = 0, \cdots, p$,

(5) $\mu_i g_i(\bar{x}) = 0$ for $i = 1, \cdots, p$ and $\lambda_i g_i(\bar{x}) = 0$ for $i = p+1, \cdots, m$.

Proof. (1) is clear. For the remaining parts, we apply the Fritz John version of Theorem 5.21 by taking $F = f$, $G_i = g_i$ $(1 \leq i \leq m)$, $\Theta = \mathbb{R}_+$, $\Lambda_i = \mathbb{R}_+$ $(1 \leq i \leq p)$ and $\Lambda_i = \{0\}$ $(p+1 \leq i \leq m)$. Notice that since Ω is open and convex, $N(\bar{x}; \Omega) = \{0\}$ (see Remark 2.26). Therefore there is a vector

$$\begin{gathered}(D^*f(\bar{x})(\mu_0), D^*g_1(\bar{x})(\mu_1), \cdots, D^*g_p(\bar{x})(\mu_p), D^*g_{p+1}(\bar{x})(\lambda_{p+1}), \cdots, D^*g_m(\bar{x})(\lambda_m), \\ \mu_0, \mu_1, \cdots, \mu_p, \lambda_{p+1}, \cdots, \lambda_m) \subset ((\mathbb{R}^n)^{m+1} \times \mathbb{R}^{m+1}) \setminus \{0\}\end{gathered} \tag{5.53}$$

such that $\mu_0 \geq 0$, $\mu_i \in -N(-g_i(\bar{x}); \mathbb{R}_+) \subset \mathbb{R}_+$ $(1 \leq i \leq p)$, $\lambda_i \in -N(0; \mathbb{R}_+) = \mathbb{R}_+$ $(p+1 \leq i \leq m)$ and

$$D^*f(\bar{x})(\mu_0) + \sum_{i=1}^{p} D^*g_i(\bar{x})(\mu_i) + \sum_{i=p+1}^{m} D^*g_i(\bar{x})(\lambda_i) = 0. \tag{5.54}$$

As f and all g_i are C^1, by Theorem 2.52 (2), (5.54) becomes (5.52). (5.53) implies that $(\mu_0, \mu_1, \cdots, \mu_p, \lambda_{p+1}, \cdots, \lambda_m) \neq 0$. We see that (2)-(4) hold. Lastly, if $i =$

$p+1, \cdots, m$, then by (1), $\lambda_i g_i(\bar{x}) = 0$. For a given $i = 1, \cdots, p$, if $g_i(\bar{x}) = 0$ then of course $\mu_i g_i(\bar{x}) = 0$. Otherwise $\mu_i = 0$ by the estimate $\mu_i \in -N(-g_i(\bar{x}); \mathbb{R}_+)$. In any case (5) holds and we finish the proof. $\square$

Under additional constraint qualifications on the feasible set near the above point $\bar{x}$, one can show $\mu_0 \neq 0$. Therefore, by dividing the equation (5.52) by μ_0 when necessary, we can assume $\mu_0 = 1$. Then the necessary conditions in the above theorem are called the Karush-Kuhn-Tucker (KKT) conditions. For details readers may refer to some texts on nonlinear programming say [4].

Through this simple situation we may understand Theorem 5.21 or a Fritz John point of problem (5.17) (see Definition 5.8 (6)) as follows. (5.21) is a general formula compared to (5.52). The dual elements $y_0^*, \cdots, y_m^*$ in (5.21) are multipliers. In the Fritz John version, it is possible that all y_i^* are zero, but is impossible after adding further restrictions or constraint qualifications, and this becomes the Lagrange version. For each i, the normal cone to Λ_i in Definition 5.8 (6) helps determine whether the multipliers vanish. Precisely, $y_i^* = 0$ if Λ_i is convex and $-\bar{y}_i$ is an interior point of Λ_i (see Remark 2.26).

Of course, our Theorem 5.21 generalizes far from the above basic result.

5.3.2 Zheng and Ng's work

Zheng and Ng [34, Definition 3.1] proposed the following property about a closed convex cone called dually compactness.

Definition 5.23. *A closed convex cone C in Y is said to be dually compact if there exists a compact subset K of Y such that*

$$C^* \subset \mathcal{W}(K)$$

where $\mathcal{W}(K) \triangleq \{y^* \in Y^* : \|y^*\| \leq \sup\{\langle y^*, y\rangle : y \in K\}\}$.

Example 5.24.

(1) If Y is finite-dimensional, then any closed convex cone C is dually compact. This is because if we take $K = \mathbb{B}$, the closed unit ball in Y, then $\mathcal{W}(K) = Y^*$ so of course $C^* \subset \mathcal{W}(K)$.

(2) If C is a solid closed convex cone, then it is dually compact. Indeed fix $c \in \text{int}(C)$, then $c + \delta\mathbb{B} \subset Y$ for some $\delta > 0$. Thus for any $c^* \in C^*$,

$$0 \leq \inf\{\langle c^*, y\rangle : y \in c + \delta\mathbb{B}\} = \langle c^*, c\rangle - \delta\|c^*\|$$

and hence

$$\|c^*\| \leq \left\langle c^*, \frac{c}{\delta}\right\rangle \leq \sup\left\{\langle c^*, y\rangle : y = \frac{c}{\delta}\right\}.$$

Therefore $C^* \subset \mathcal{W}\left(\left\{\frac{c}{\delta}\right\}\right)$.

The following result shows that dually compactness of a closed convex cone is a special case of the SNC property of a cone at the origin.

Proposition 5.25. *If a closed convex cone $C \subset Y$ is dually compact, then C is SNC at 0.*

Proof. By assumption $C^* \subset \mathcal{W}(K)$ for some compact set K in Y. Suppose $y_k \xrightarrow{C} 0$, $y_k^* \in \hat{N}(y_k; C)$ for all k and $y_k^* \xrightarrow{w^*} 0$ as $k \to \infty$. Then for each k, $z_k^* \triangleq -y_k^* \in -\hat{N}(y_k; C) \subset C^* \subset \mathcal{W}(K)$ and hence

$$0 \leq \|z_k^*\| \leq \sup\{\langle z_k^*, y\rangle : y \in K\} = \langle z_k^*, w_k\rangle \tag{5.55}$$

for some $w_k \in K$ (by the continuity of z_k^* and compactness of K). As K is compact we may assume $w_k \to w \in K$ as $k \to \infty$. Together with the convergence $z_k^* \xrightarrow{w^*} 0$ we know $\langle z_k^*, w_k\rangle \to 0$ as $k \to \infty$. If follows from (5.55) that $\|y_k^*\| = \|z_k^*\| \to 0$ as $k \to \infty$. Hence C is SNC at 0. □

Here we consider the unconstrained problem

$$\begin{aligned} &\text{minimize } F(x) \\ &\text{subject to } x \in X \end{aligned} \tag{5.56}$$

where $F : X \rightrightarrows Y$ is a multifunction between Banach spaces and Y is ordered by a nontrivial closed convex pointed cone $C \subset Y$.

Theorem 5.26. *[34, Theorem 4.1] Let X and Y be Asplund spaces and $F : X \rightrightarrows Y$ be a closed multifunction (that is,* $\mathrm{gph}(F)$ *is closed in $X \times Y$). Suppose that $(\bar{x}, \bar{y})$ is a local Pareto minimizer of (5.56). Then for any $\varepsilon > 0$ there exist $x_\varepsilon \in \bar{x} + \varepsilon\mathbb{B}$, $y_\varepsilon \in \bar{y} + \varepsilon\mathbb{B}$ and $c^* \in C^*$ with $\|c^*\| = 1$ such that*

$$0 \in D^*F(x_\varepsilon, y_\varepsilon)(c^* + \varepsilon\mathbb{B}^*) + \varepsilon\mathbb{B}^*.$$

Proof. Note that $-N(c; C) \subset C^*$ for any $c \in C$. Also as C is a closed convex pointed cone (thus is not a vector subspace), C is LAC at 0 by Proposition 5.4 (3). The conclusion follows by applying the fuzzy version of Theorem 5.12 (take $\Theta = C$ and $\Omega = X$). □

Theorem 5.27. *[34, Theorem 4.2] Let X and Y be Asplund spaces and $F : X \rightrightarrows Y$ be a closed multifunction. Suppose that $(\bar{x}, \bar{y})$ is a local Pareto minimizer of (5.56). Assume one of the following conditions is satisfied:*

(1) F is PSNC at $(\bar{x}, \bar{y})$ with respect to Y, that is, the set $\mathrm{gph}(F)$ *is PSNC at $(\bar{x}, \bar{y})$ with respect to Y.*

(2) The ordering cone C is dually compact.

(3) $\mathrm{int}(C) \neq \emptyset$ *or Y is finite dimensional.*

Then there exists $c^* \in C^*$ with $\|c^*\| = 1$ such that $0 \in D^*F(\bar{x}, \bar{y})(c^*)$.

Proof. By Example 5.24, (3) implies (2). It suffices to consider the situation that either (1) or (2) holds. (1) is equivalent to the PSNC property of F^{-1} at $(\bar{y}, \bar{x})$ while by Proposition 5.25, (2) implies C is SNC at 0. Also as before, C is LAC at 0. Hence the conclusion follows from Corollary 5.13 (taking $\Theta = C$) . □

Next we study the following constrained vector optimization problem:

$$\begin{aligned} &\text{minimize } F_0(x) \\ &\text{subject to } F_i(x) \cap (-C_i) \neq \emptyset, \quad i = 1, \cdots, m \\ &\qquad x \in \Omega \end{aligned} \tag{5.57}$$

where $X, Y_0, Y_1, \cdots, Y_m$ be Banach spaces, Ω be a closed subset of X, $F_i : X \rightrightarrows Y_i$ $(i = 0, 1, \cdots, m)$ be closed multifunctions, Y_0 is ordered by a closed convex cone C_0 which is not a vector subspace, and Y_i is ordered by a closed convex cone C_i $(i = 1, \cdots, m)$.

Theorem 5.28. *[35, Theorem 4.1] Consider the problem (5.57) where $X, Y_0, Y_1, \cdots, Y_m$ are all Asplund spaces, each F_i is a closed multifunction, and $(\bar{x}, \bar{y}_0) \in \mathrm{gph}(F_0)$ is a local Pareto minimizer of the problem (5.57) and $\bar{y}_i \in F_i(\bar{x}) \cap (-C_i)$ $(1 \leq i \leq m)$. Suppose that each C_i is dually compact and that each F_i is PSNC at $(\bar{x}, \bar{y}_i)$. Then one of the following assertion holds.*

(1) There exist $c_i^* \in C_i^*$ $(0 \leq i \leq m)$ such that

$$\sum_{i=0}^{m} \|c_i^*\| = 1 \text{ and } 0 \in \sum_{i=0}^{m} D^* F_i(\bar{x}, \bar{y}_i)(c_i^*) + N(\bar{x}; \Omega).$$

(2) There exist $x_i^* \in D^* F_i(\bar{x}, \bar{y}_i)(0)$ $(0 \leq i \leq m)$ and $w^* \in N(\bar{x}; \Omega)$ such that

$$\|w^*\| + \sum_{i=0}^{m} \|x_i^*\| = 1 \text{ and } w^* + \sum_{i=0}^{m} x_i^* = 0.$$

Proof. It follows from Proposition 5.25 that each C_i is SNC at 0. Since C_0 is a closed convex cone which is not a vector subspace, it is LAC at 0 by Proposition 5.4 (3). We recall that $-N(-\bar{y}_i; C_i) \subset C_i^*$ and $D_M^* F_i(\bar{x}, \bar{y})(0) \subset D^* F_i(\bar{x}, \bar{y})(0)$ $(0 \leq i \leq m)$. Therefore the conclusion follows from applying the Lagrange version of Theorem 5.21, with taking $F, G_1, \cdots, G_m, \Theta, \Lambda_1, \cdots, \Lambda_m$ in the setting of that theorem to be $F_0, F_1, \cdots, F_m, C_0, C_1, \cdots, C_m$ respectively. □

Theorem 5.29. *[35, Theorem 4.2] Consider the problem (5.57) where $X, Y_0, Y_1, \cdots, Y_m$ are all Asplund spaces, each F_i is a closed multifunction, and $(\bar{x}, \bar{y}_0) \in \mathrm{gph}(F_0)$ is a local Pareto minimizer of the problem (5.57) and $\bar{y}_i \in F_i(\bar{x}) \cap (-C_i)$ $(1 \leq i \leq m)$. Suppose that each C_i is dually compact and that each F_i is Lipschitz-like at $(\bar{x}, \bar{y}_i)$. Then there exist $c_i^* \in C_i^*$ $(0 \leq i \leq m)$ such that*

$$\sum_{i=0}^{m} \|c_i^*\| = 1 \text{ and } 0 \in \sum_{i=0}^{m} D^* F_i(\bar{x}, \bar{y}_i)(c_i^*) + N(\bar{x}; \Omega).$$

Proof. Since all F_i are Lipschitz-like at $(\bar{x}, \bar{y}_i)$, by Theorem 2.53, we know $D^*_M F_i(\bar{x}, \bar{y}_i)(0) = \{0\}$. Also each F_i is PSNC at the corresponding point $(\bar{x}, \bar{y}_i)$ by Theorem 2.48 (4). Hence we apply the Lagrange version of Theorem 5.21 to the setting of this theorem just as the last proof with noting that two qualification conditions (5.39) and (5.40) in Theorem 5.21 are satisfied. The desired result is then obtained. □

Our main results improve Zheng and Ng's work in the following aspects. First, in practice there may not exist local Pareto minimizers, while our results include the necessary conditions for weaker notions like local weak/ relative minimizers. Second, the dually compactness assumption on the ordering cones can be replaced by a weaker assumption: SNC property at the origin. The book [25] mentions more criteria to check the SNC property. We remark that in the paper [35], Zheng and Ng established the above necessary conditions by a separation theorem which generalizes the approximate extremal principle.

5.3.3 Dutta and Tammer's work

Dutta and Tammer formulated the following necessary condition for weak minimality by scalarization. Roughly speaking, they consider a continuous convex functional g on $\mathbb{R}^l$ such that if $\bar{x}$ is a weak minimizer of f, then it is a minimizer of $g \circ f$ which is real-valued. The desired result can be obtained by studying the subdifferential of $g \circ f$. Besides, this result is also a consequence of Theorem 5.12.

Theorem 5.30. *[7, Theorem 3.2] Given a single-valued locally Lipschitz mapping $f : X \to \mathbb{R}^l$ where X is an Asplund space and $\mathbb{R}^l$ is ordered by a closed convex pointed cone K. Assume S is a closed subset of X, and let $\bar{x}$ be a weak minimizer of f. Then there exists $v^* \in K^* \setminus \{0\}$ such that*

$$0 \in \partial(v^* \circ f)(\bar{x}) + N(\bar{x}; S).$$

Further if f is strictly differentiable then one has

$$0 \in (\nabla f(\bar{x}))^* v^* + N(\bar{x}; S).$$

Proof. As $\bar{x}$ is a weak minimizer of f, we must have $\text{int}(K) \neq \emptyset$ and $0 \notin \text{int}(K)$. Since K is a closed convex pointed solid cone, which must not be a vector subspace, Proposition 5.4 (1) tells us that $\text{int}(K) \cup \{0\}$ is LAC at 0. As K is a convex subset in $\mathbb{R}^l$, K is SNC at $\bar{x}$ by Theorem 2.46 (1). By Theorem 2.48 (2), f is PSNC at $(\bar{x}, f(\bar{x}))$. Also $D^*_M f(\bar{x})(0) = \{0\}$ by Theorem 2.53. The desired results can be obtained by applying the Lagrange version of Theorem 5.12 with taking $F = f$, $\Omega = S$, $\Theta = \text{int}(K) \cup \{0\}$ there, and using Theorem 2.56. □

The main drawback of the abovementioned result is that it requires the ordering cone to have nonempty interior which is not a must. We have explained this in Chapter 1. Also the objective mappings are restricted to locally Lipschitz ones.

5.3.4 Bao and Mordukhovich's previous work

Bao and Mordukhovich [2] established necessary conditions for optimality which are similar to our results, but they do not discuss the unified notion given in Chapter 3 and set-valued optimization problems with operator constraints done in this chapter.

For a multifunction $F : X \rightrightarrows Y$, we define the indicator multifunction of $\Omega \subset X$, $\Delta(\cdot;\Omega) : X \rightrightarrows Y$ by

$$\Delta(x;\Omega) \triangleq \begin{cases} \{0\} & \text{if } x \in \text{Dom}(F) \\ \emptyset & \text{otherwise} \end{cases}$$

Also define the mapping $F_\Omega : X \rightrightarrows Y$ by

$$F_\Omega(x) \triangleq F(x) + \Delta(x;\Omega) \text{ for all } x \in X.$$

The following two results are proved in [2, Theorem 5.3, Corollary 5.4].

Theorem 5.31. *Consider the problem (5.16) where X and Y are Asplund spaces, $\Theta = C \subset Y$ is a closed convex cone with $C \setminus (-C) \neq \emptyset$. Fix $(\bar{x}, \bar{y}) \in \text{gph}(F)$.*

Assume that the constraint set $\Omega \subset X$ is locally closed at $\bar{x}$, gph(F) *is locally closed at $(\bar{x}, \bar{y})$ and the qualification condition*

$$D^*_M F(\bar{x}, \bar{y})(0) \cap (-N(\bar{x}; \Omega)) = \{0\} \tag{5.58}$$

is satisfied. Further assume either F is PSNC at $(\bar{x}, \bar{y})$ or Ω is SNC at $\bar{x}$. Suppose one of the following cases is satisfied:

(1) $(\bar{x}, \bar{y})$ is a local Pareto minimizer of the problem (5.16) and either C is SNC at 0 or F_Ω^{-1} is PSNC at $(\bar{y}, \bar{x})$.

(2) $(\bar{x}, \bar{y})$ is a local quasi relative minimizer of the problem (5.16) and either C is SNC at 0 or F_Ω^{-1} is PSNC at $(\bar{y}, \bar{x})$.

(3) $(\bar{x}, \bar{y})$ is a local intrinsic relative minimizer of the problem (5.16) and either C is SNC at 0 or F_Ω^{-1} is PSNC at $(\bar{y}, \bar{x})$.

(4) $(\bar{x}, \bar{y})$ is a local relative minimizer of the problem (5.16) and either $\overline{\text{aff}}(C)$ is finite codimensional or F_Ω^{-1} is PSNC at $(\bar{y}, \bar{x})$.

(5) $(\bar{x}, \bar{y})$ is a local weak minimizer of the problem (5.16).

Then we have

$$0 \in D^* F(\bar{x}, \bar{y})(y^*) + N(\bar{x}; \Omega) \text{ for some } y^* \in -N(0; C) \text{ with } \|y^*\| = 1.$$

Corollary 5.32. *The same conclusion of Theorem 5.31 can be obtained if the original assumptions are kept except the following modifications: The qualification condition (5.58) is replaced by*

$$D^* F(\bar{x}, \bar{y})(0) \cap (-N(\bar{x}; \Omega)) = \{0\}. \tag{5.59}$$

Also the assumption "F_Ω^{-1} is PSNC at $(\bar{y}, \bar{x})$" is replaced by either

(1) "F^{-1} is PSNC at $(\bar{y}, \bar{x})$ and Ω is SNC at $\bar{x}$", or

(2) "F is SNC at $(\bar{x}, \bar{y})$".

Corollary 5.32 is a consequence of Theorem 5.12. To show this, first, by Theorem 2.46 (2), C must be SNC at 0 whenever $\text{int}(C) \neq \emptyset$. Hence if assumption (5) in Theorem 5.31 holds, then C is SNC at 0. Second, when applying Theorem 5.12, we need to ensure that the relation set is LAC at 0. This is guaranteed by Proposition 5.4. Third, (5.59) is stronger than (5.58). Taking these three points into consideration and applying the Lagrange version of Theorem 5.12 under the assumptions given in Corollary 5.32, we get the desired conclusion.

Chapter 6

A weak notion: approximate efficiency

In formulating necessary optimality conditions in last chapter we need to assume a Θ-minimal point, say, ideal/Pareto/weak minimal point, exists. Nevertheless in applications this cannot be guaranteed. To tackle this, one direction is to consider weaker notions like relative/intrinsic/quasi minimal points as discussed in the previous chapters. For example, Bao and Mordukhovich discussed the existence of intrinsic minimal points in [2]. In this chapter we are going to investigate on another common weak notion called approximate minimality.

Suppose g is a function from a nonempty set Ω to the real line. We call $\bar{x} \in \Omega$ an ε-minimizer of g ($\varepsilon \geq 0$) if $g(\bar{x}) - \varepsilon \leq g(x)$ for all $x \in \Omega$. Note that if $\bar{x}$ is a global minimizer of g, then it is an ε-minimizer of g for any $\varepsilon \geq 0$. Gutiérrez, Jiménez and Novo [16] developed an interesting generalization for this and many types of approximate solutions. In this chapter we are going to discuss their way of generalization and scalarization results provided by them. Originally they discuss the notion on single-valued mappings, but we will discuss it on multifunctions.

For convenience, we mention here that unless otherwise specified, the definitions and the results in this chapter where the objective map is single-valued can be found in [16].

6.1 Approximate minimality

Given two topological vector spaces X and Y, a multifunction $F : X \rightrightarrows Y$ and a nonempty subset Ω of X. Also let $C \subset Y$ be a nonempty set. Denote $C(\varepsilon)$ to be εC if $\varepsilon > 0$ and take

$$C(0) \triangleq \bigcup_{\varepsilon>0} C(\varepsilon).$$

Definition 6.1. *Let $\varepsilon \geq 0$, $\bar{x} \in \Omega$ and $(\bar{x}, \bar{y}) \in \mathrm{gph}(F)$. We say that $(\bar{x}, \bar{y})$ is an* <u>*ε-minimizer*</u> *of the problem (3.1) if*

$$F(\Omega) \cap (\bar{y} - C(\varepsilon)) \subset \{\bar{y}\}.$$

Denote $\mathrm{AE}(F, \Omega, C; \varepsilon)$ be the collection of such minimizers. If $F = f$ is single-valued, then we use $\mathrm{AE}(f, \Omega, C; \varepsilon)$ to denote the collection of all points $\bar{x} \in \Omega$ such that $(\bar{x}, f(\bar{x})) \in \mathrm{AE}(F, \Omega, C; \varepsilon)$. It follows from the above definitions that

$$\mathrm{AE}(F, \Omega, C; 0) = \bigcap_{\varepsilon>0} \mathrm{AE}(F, \Omega, C; \varepsilon). \tag{6.1}$$

Remark 6.2. *Here are some special cases:*

(1) If C is a convex cone, then $C(\varepsilon) = C$ for all $\varepsilon > 0$ and so the set $\mathrm{AE}(F, \Omega, C; 0)$ is precisely the collection of Pareto minimizers $\mathrm{Min}(F, \Omega, C)$.

(2) If C is a proper solid convex cone, then the set $\mathrm{AE}(F, \Omega, \mathrm{int}(C); 0)$ (i.e., the set C that we consider is replaced by $\mathrm{int}(C)$) is precisely the collection of weak minimizers $\mathrm{WMin}(F, \Omega, C)$.

(3) If g is an extended real-valued function on Y, we let

$$\begin{aligned}\mathrm{AE}(g, \Omega; \varepsilon) \triangleq \mathrm{AE}(g, \Omega, \mathbb{R}_+; \varepsilon) &= \{y \in \Omega : g(y) - \varepsilon \leq g(z) \text{ for all } z \in \Omega\} \\ &= \text{the set of all } \varepsilon\text{-minimizers of } g.\end{aligned}$$

The first concept of approximate minimality which is also the most popular one was introduced by Kutateladze [22]. It was generalized a little bit to become the following notion introduced by Németh [28] (cf. [16, Section 4.1]).

Example 6.3. *Let $K \subset Y$ be a pointed convex cone and H be a subset of $K \setminus \{0\}$. Consider $C \triangleq H + K$. Then for all $\varepsilon > 0$,*

$$C(\varepsilon) = \varepsilon(H + K) = \varepsilon H + \varepsilon K = \varepsilon H + K.$$

We claim $0 \notin C$. If $0 \in C$, then $0 = h + k$ for some $h \in H$, $k \in K$. Then $K \setminus \{0\} \ni h = -k \in -K$ contradicting K is pointed. Thus $0 \notin C$, implying $0 \notin C(\varepsilon)$ for any $\varepsilon > 0$. Therefore for each $\varepsilon > 0$, as K is a cone,

$$\begin{aligned}(\bar{x}, \bar{y}) \in \mathrm{AE}(F, \Omega, C; \varepsilon) &\Leftrightarrow F(\Omega) \cap (\bar{y} - C(\varepsilon)) = \emptyset \\ &\Leftrightarrow F(\Omega) \cap (\bar{y} - \varepsilon H - K) = \emptyset.\end{aligned} \tag{6.2}$$

In particular when H is a singleton, that is, $H = \{\xi\}$ where $\xi \in K \setminus \{0\}$, (6.2) reduces to

$$(\bar{x}, \bar{y})) \in \mathrm{AE}(F, \Omega, C; \varepsilon) \Leftrightarrow F(\Omega) \cap (\bar{y} - \varepsilon\xi - K) = \emptyset.$$

We call $(\bar{x}, \bar{y})$ an (ε, ξ)-minimizer. When F is single-valued, the notion of (ε, ξ)-minimizers was introduced by Kutateladze [22]. Meanwhile, one can check the following statement easily: If $(\bar{x}, \bar{y}) \in \mathrm{Min}(F, \Omega, K)$, then $(\bar{x}, \bar{y})$ is an (ε, ξ)-minimizer. for any $\varepsilon > 0$ and $\xi \in K \setminus \{0\}$. This explains why approximate minimality is a weaker notion than Pareto minimality.

6.2 A scalarization result

We can establish the necessary conditions for approximate minimality of the problem (3.1) through scalarization. We find out a nice real-valued (or extended real-valued) function G such that $\bar{x}$ is an approximate solution to $G \circ F$ whenever $(\bar{x}, \bar{y})$ is an approximate solution to the problem (3.1). Now we discuss some properties of the set C given before which become the assumptions of the main result.

Definition 6.4. *A set $C \subset Y$ is called*

(1) proper if $\emptyset \neq C \neq Y$.

(2) solid if $\text{int}(C) \neq \emptyset$.

(3) pointed if $C \cap (-C) \subset \{0\}$, that is, $C \cap (-C) = \{0\}$ if $0 \in C$ and $C \cap (-C) = \emptyset$ otherwise.

(4) coradiant if "$tc \in C$ whenever $t > 1$ and $c \in C$".

Example 6.5.

(1) A solid pointed cone is a solid pointed coradiant set.

(2) For each convex set $A \subset Y$ containing 0, A^c is a coradiant set. Indeed for all $y \in A^c$ and $t > 1$, if $ty \in A$, then $y = \left(1 - \frac{1}{t}\right) 0 + \frac{1}{t}(ty) \in A$ by convexity. A contradiction happens. Thus $ty \in A^c$.

The next proposition [16, Lemma 3.1] lists some basic properties of $C(\varepsilon)$ for a nice set C.

Proposition 6.6. *Suppose C is a solid pointed coradiant set. Then we have:*

(1) $C(\varepsilon) = \varepsilon C$ is a solid pointed coradiant set for all $\varepsilon > 0$.

(2) $C(\varepsilon_2) \subset C(\varepsilon_1)$ whenever $0 < \varepsilon_1 < \varepsilon_2$.

(3) $C(0) \cup \{0\} = \text{cone}(C)$ is a solid pointed cone.

Proof.

(1) For $\varepsilon > 0$, we have that εC is solid as $\text{int}(\varepsilon C) = \varepsilon \text{int}(C)$. If $y \in (\varepsilon C) \cap (-\varepsilon C)$, then $\frac{1}{ve} y \in C \cap (-C) \subset \{0\}$ and thus $\frac{1}{\varepsilon} y = 0$ or $y = 0$. Therefore εC is pointed. Let $y \in \varepsilon C$. Then $y = \varepsilon c$ for some $c \in C$. For any $t > 1$, $ty = t(\varepsilon c) = \varepsilon tc \in \varepsilon C$ as C is coradiant. Hence εC is also coradiant.

(2) Suppose $0 < \varepsilon_1 < \varepsilon_2$ and $y \in C(\varepsilon_2) = \varepsilon_2 C$. Then $y = \varepsilon_2 c$ for some $c \in C$. Let $t \triangleq \frac{\varepsilon_2}{\varepsilon_1} > 1$. We see that

$$y = \varepsilon_2 c = \frac{\varepsilon_2}{\varepsilon_1}(\varepsilon_1 c) = t(\varepsilon_1 c) \in tC(\varepsilon_1) \subset C(\varepsilon_1)$$

as $C(\varepsilon_1)$ is coradiant by (1). Thus $C(\varepsilon_2) \subset C(\varepsilon_1)$.

(3) It is clear that cone(C) is a solid cone. Suppose $y \in \text{cone}(C) \cap (-\text{cone}(C))$. Then $y \in C(\varepsilon_1)$ and $-y \in C(\varepsilon_2)$ for some $\varepsilon_1, \varepsilon_2 > 0$. Let $\varepsilon_0 = \min\{\varepsilon_1, \varepsilon_2\} > 0$. As $C(\varepsilon_1) \subset C(\varepsilon_0)$ and $C(\varepsilon_2) \subset C(\varepsilon_0)$ by (2), $y \in C(\varepsilon_0) \cap (-C(\varepsilon_0)) \subset \{0\}$ by (1). Therefore $y = 0$. This implies cone(C) is pointed. We arrive at the conclusion.

□

Definition 6.7. *A subset C of Y is said to be* <u>*star-shaped*</u> *if there exists $q \in C$ such that*

$$\alpha q + (1-\alpha)y \in C \text{ for all } y \in C, \alpha \in (0,1). \tag{6.3}$$

Denote by $\ker(C)$ *the* <u>*kernel*</u> *of C which is the collection of all points $q \in C$ such that (6.3) holds.*

Lemma 6.8. *We have the following properties.*

(1) If $C \subset Y$ is coradiant, then $\ker(C)$ *is also coradiant.*

(2) Further if $\ker(C)$ *is solid, then* $\text{int}(\ker(C))$ *is also coradiant.*

Proof.

(1) Let $q \in \ker(C)$ and $\lambda > 0$. We want to show $(\lambda+1)q \in \ker(C)$. For any $y \in C$, $\alpha \in (0,1)$, $\frac{\alpha\lambda+\alpha}{\alpha\lambda+1}q + \left(1 - \frac{\alpha\lambda+\alpha}{\alpha\lambda+1}\right)y \in C$ as $q \in \ker(C)$. Therefore

$$\begin{aligned}\alpha(\lambda+1)q + (1-\alpha)y &= (\alpha\lambda+1)\left(\frac{\alpha(\lambda+1)}{\alpha\lambda+1}q + \frac{1-\alpha}{\alpha\lambda+1}y\right)\\ &= (\alpha\lambda+1)\left(\frac{\alpha\lambda+\alpha}{\alpha\lambda+1}q + \left(1-\frac{\alpha\lambda+\alpha}{\alpha\lambda+1}\right)y\right) \in C\end{aligned}$$

by the assumption that C is coradiant. We are done.

(2) If $q \in \text{int}(\ker(C))$ and $\lambda > 0$, then $q + U \subset \ker(C)$ for some neighborhood U of 0. Hence $(\lambda+1)q + (\lambda+1)U \subset (\lambda+1)\ker(C) \subset \ker(C)$ since $\ker(C)$ is coradiant by (1). This implies $(\lambda+1)q \in \text{int}(\ker(C))$.

□

Lemma 6.9. *Suppose C is a proper star-shaped coradiant set such that $\ker(C)$ is solid. Then the following statements hold:*

(1) $c + tq \in C$ whenever $c \in C$, $q \in \ker(C)$ and $t > 0$.

(2) $c + tq \in C(\varepsilon)$ whenever $\varepsilon > 0$, $c \in C(\varepsilon)$, $q \in \ker(C)$ and $t > 0$.

(3) $\text{cl}(C(\varepsilon)) + (0,\infty)q \subset \text{int}(C(\varepsilon))$ whenever $\varepsilon > 0$ and $q \in \ker(C)$.

(4) $Y = \mathbb{R}q - C(\varepsilon)$ whenever $q \in \text{int}(C)$ and $\varepsilon > 0$.

(5) $Y = \mathbb{R}q - \varepsilon\text{int}(\ker(C))$ whenever $q \in \text{int}(\ker(C))$ and $\varepsilon > 0$.

(6) For all $q \in \text{int}(\ker(C))$, $\varepsilon > 0$, and $y \in Y$, there exists $t \in \mathbb{R}$ such that $y + tq \notin -\text{cl}(C(\varepsilon))$.

(7) $\text{int}(\text{cl}(C(\varepsilon))) = \text{int}(C(\varepsilon))$ whenever $\varepsilon > 0$.

Proof.

(1) Let $c \in C$, $q \in \ker(C)$ and $t > 0$. Since C is a star-shaped coradiant set and $q \in \ker(C)$,

$$c + tq = (t+1)\left(\frac{1}{t+1}c + \left(1 - \frac{1}{t+1}\right)q\right) \in C.$$

(2) Given $\varepsilon > 0$, $c \in C(\varepsilon)$, $q \in \ker(C)$ and $t > 0$. As $\frac{1}{\varepsilon}c \in C$, we apply (1) that $\frac{1}{\varepsilon}c + \frac{t}{\varepsilon}q \in C$ implying $c + tq \in \varepsilon C = C(\varepsilon)$.

(3) Let $\varepsilon > 0$ and $q \in \ker(C)$. We first show

$$C(\varepsilon) + (0,\infty)q \subset \text{int}(C(\varepsilon)). \tag{6.4}$$

As $q \in \text{int}(\ker(C))$, there is a neighborhood U of 0 such that $q+U \subset \ker(C)$. If $c \in C(\varepsilon)$ and $t > 0$, then by (2), $(c+tq)+tU = c+t(q+U) \subset C(\varepsilon)$. We get $c+tq \in \text{int}(C(\varepsilon))$. Hence (6.4) is justified. Next we prove the inclusion in (3). Let $c \in \text{cl}(C(\varepsilon))$ and $t > 0$. We have $c = \lim_\alpha v_\alpha$ for some net $\{v_\alpha\} \subset C(\varepsilon)$. Therefore $\lim_\alpha \left(q + \frac{1}{t}(c - v_\alpha)\right) = q \in \text{int}(\ker(C))$. Hence there exists $v \in C(\varepsilon)$ such that $q + \frac{1}{t}(c-v) \in \text{int}(\ker(C))$. Together with (6.4) (where q is replaced by $q + \frac{1}{t}(c-v) \in \text{int}(\ker(C))$) we obtain

$$c + tq = v + t\left(q + \frac{1}{t}(c-v)\right) \in \text{int}(C(\varepsilon)).$$

The proof is complete.

(4) Let $q \in \text{int}(C)$ and $\varepsilon > 0$. For each $y \in Y$, take $\delta \in (0,1)$ such that $q - \frac{\delta}{\varepsilon}y \in C$. As $C(\varepsilon)$ is a coradiant set by Proposition 6.6 (1), $y \in \frac{\varepsilon}{\delta}q - \frac{1}{\delta}(\varepsilon C) \subset \mathbb{R}q - C(\varepsilon)$. Hence $Y \subset \mathbb{R}q - C(\varepsilon)$ and of course equality holds.

(5) Let $q \in \text{int}(\ker(C))$ and $\varepsilon > 0$. Following the idea of the proof of (4) and invoking Lemma 6.8 (1), there exists $\delta \in (0,1)$ such that

$$y \in \frac{\varepsilon}{\delta} - \varepsilon\left(\frac{1}{\delta}\text{int}(\ker(C))\right) \subset \mathbb{R}q - \varepsilon\text{int}(\ker(C)).$$

The last inclusion holds because $\text{int}(\ker(C))$ is a coradiant set by Lemma 6.8 (2). Thus $Y = \mathbb{R}q - \varepsilon\text{int}(\ker(C))$.

(6) Suppose to the contrary that there exist $q \in \text{int}(\ker(C))$, $\varepsilon > 0$ and $y \in Y$ such that $y + \mathbb{R}q \subset -\text{cl}(C(\varepsilon))$. By this, (5) and (3), we obtain:

$$\begin{aligned} Y &= \mathbb{R}q - \varepsilon\text{int}(\ker(C)) = (y + \mathbb{R}q) - \varepsilon\text{int}(\ker(C)) - y \\ &\subset -\text{cl}(C(\varepsilon)) - \varepsilon\text{int}(\ker(C)) - y \\ &\subset -\text{int}(C(\varepsilon)) - y = -\text{int}(\varepsilon C) - y = -\varepsilon\text{int}(C) - y \end{aligned}$$

implying $Y \subset \text{int}(C)$ contradicting the fact that C is proper. Thus (6) holds.

(7) For a given $\varepsilon > 0$, it is clear that $\text{int}(C(\varepsilon)) \subset \text{int}(\text{cl}(C(\varepsilon)))$. For the reverse inclusion, fix $q \in \text{int}(\ker(C))$ and let $c \in \text{int}(\text{cl}(C(\varepsilon)))$. Then there exists $\delta > 0$ such that $c - \delta q \in \text{cl}(C(\varepsilon))$. Therefore by (3), $c \in \text{cl}(C(\varepsilon)) + \delta q \in \text{int}(C(\varepsilon))$. We finish the proof.

□

Under the same setting as Lemma 6.9, given $q \in \text{int}(\ker(C))$ and $\varepsilon > 0$, we now consider the generalized Gerstewitz function (see [12]) $\Phi : 2^Y \to \overline{\mathbb{R}}$ defined by

$$\Phi(A) \triangleq \inf\{t \in \mathbb{R} : tq \in A + \text{cl}(C(\varepsilon))\} \text{ for any set } A \subset Y.$$

Note that Φ is determined by C, q and ε.

Proposition 6.10. *[12, P. 892] Let A be a nonempty subset of Y and $r \in \mathbb{R}$. Then we have:*

(1) $\Phi(A) < +\infty$.

(2) If there exists $t_0 \in \mathbb{R}$ such that $t_0 q \in A + \text{cl}(C(\varepsilon))$, then $tq \in A + \text{int}(C(\varepsilon))$ for all $t > t_0$.

(3) $\Phi(A) < r$ *if and only if* $rq \in A + \text{int}(C(\varepsilon))$.

Proof.

(1) Let A be a nonempty set with $a \in A$. Since $q \in \text{int}(C)$, by Lemma 6.9 (4), $a \in t_0 q - C(\varepsilon)$ for some $t_0 \in \mathbb{R}$. Then the set $\{t \in \mathbb{R} : tq \in A + \text{cl}(C(\varepsilon))\}$ is nonempty and $\Phi(A) \neq +\infty$.

(2) If $t_0 \in \mathbb{R}$ is the number given in the assumption, then for $t > t_0$, we have $tq = t_0 q + (t - t_0)q \in A + \text{cl}(C(\varepsilon)) + (0, \infty)q \subset A + \text{int}(C(\varepsilon))$ by Lemma 6.9 (3).

(3) If $\Phi(A) < r$, then $tk \in A+\text{cl}(C(\varepsilon))$ for some $t < r$. Proposition 6.10 (2) shows $rq \in A + \text{int}(C(\varepsilon))$. If $rq \in A + \text{int}(C(\varepsilon))$, then since $A + \text{int}(C(\varepsilon))$ is open, there exists $\delta > 0$ such that $(r-\delta)q = rq-\delta q \in A+\text{int}(C(\varepsilon)) \subset A+\text{cl}(C(\varepsilon))$. Thus $\Phi(A) \leq r - \delta < r$.

□

From the previous definition we can define the well-known Gerstewitz function (see [13, 14, 23]) $\varphi : Y \to \overline{\mathbb{R}}$ by $\varphi(y) \triangleq \Phi(\{y\}) \triangleq \inf\{t \in \mathbb{R} : y \in tq - \text{cl}(C(\varepsilon))\}$ for any $y \in Y$. Then we can show that φ must take real values. Indeed, by Proposition 6.10 (1) we know $\varphi < +\infty$. If there is $y \in Y$ such that $\varphi(y) = -\infty$, then by Proposition 6.10 (2), $y - tq \in -\text{cl}(C(\varepsilon))$ for all $t \in \mathbb{R}$, contradicting Lemma 6.9 (6).

The following result formulates the necessary condition of approximate minimality through scalarization.

Theorem 6.11. *Under the same setting as as Lemma 6.9, or more precisely, let C be a proper star-shaped coradiant set such that* $\ker(C)$ *is solid. Given* $(\bar{x}, \bar{y}) \in \text{gph}(F)$, $q \in \text{int}(\ker(C))$ *and* $\varepsilon > 0$. *If* $(\bar{x}, \bar{y}) \in \text{AE}(F, \Omega, \text{int}(C); \varepsilon)$, *then* $\bar{x} \in \text{AE}(\widetilde{\Phi} \circ F, \Omega; \varepsilon)$, *where* $\widetilde{\Phi} : 2^Y \to \overline{\mathbb{R}}$ *is given by* $\widetilde{\Phi}(A) = \Phi(A - \bar{y})$.

Proof. Notice that by Proposition 6.10 (1), $\Phi \not\equiv +\infty$. If $(\bar{x}, \bar{y}) \in \text{AE}(F, \Omega, \text{int}(C); \varepsilon)$, then by the definition,

$$(F(\Omega) - \bar{y}) \cap (-\text{int}(C(\varepsilon))) = \emptyset.$$

Hence by Proposition 6.10 (2) we get $(\widetilde{\Phi} \circ F)(x) \geq 0$ for all $x \in \Omega$. Furthermore

$$(\widetilde{\Phi} \circ F)(\bar{x}) = \Phi(F(\bar{x}) - \bar{y}) = \inf\{t \in \mathbb{R} : tq \in F(\bar{x}) - \bar{y} + \text{cl}(C(\varepsilon))\} \leq \varepsilon$$

as $\bar{y} \in F(\bar{x})$ and $\varepsilon q \in C(\varepsilon)$. Thus we have

$$(\widetilde{\Phi} \circ F)(\bar{x}) - \varepsilon \leq 0 \leq (\widetilde{\Phi} \circ F)(x) \text{ for all } x \in \Omega$$

implying $\bar{x} \in \text{AE}(\widetilde{\Phi} \circ F, \Omega; \varepsilon)$. □

Remark 6.12. *Theorem 6.11 also provides a necessary condition for* $(\bar{x}, \bar{y}) \in \mathrm{AE}(F, \Omega, C; \varepsilon) \subset \mathrm{AE}(F, \Omega, \mathrm{int}(C); \varepsilon)$*; and* $(\bar{x}, \bar{y}) \in \mathrm{WMin}(F, \Omega, C)$ *(cf. (6.1) and Remark 6.2 (2)).*

Example 6.13. *Concerning the special case discussed in Example 6.3, when can we apply the previous theorem for this situation? This is possible if we add the assumptions that* K *is a solid convex cone and* H *is a convex set. It is because under these additional assumptions, we have* $\mathrm{int}(C) = \mathrm{int}\left(\bigcup_{q \in H}(q + K)\right) \supset \bigcup_{q \in H}(q + \mathrm{int}(K)) \neq \emptyset$ *so* C *is solid. As* C *is convex, the set* $\ker(C) = C$ *is also solid. Moreover, we know* $C = H + K \subset K \setminus \{0\} + K \subset K$ *as* K *is a convex cone. Thus* C *is pointed since* K *is pointed. In particular,* C *is proper. For any* $q \in H$ *and* $t > 1$,

$$t(q + K) = q + ((t - 1)q + tK) \subset q + ((t - 1)K + tK) = q + (K + K) = q + K.$$

Thus

$$tC = t(H + K) = \bigcup_{q \in H} t(q + K) \subset \bigcup_{q \in H}(q + K) = H + K = C$$

implying C *is coradiant. We conclude that* C *is a proper convex coradiant set such that* $\ker(C)$ *is solid. Therefore we can apply the previous theorem, particularly for the most standard case of which* H *is a singleton.*

Finally we provide a condition [16, Proposition 5.7] to ensure that the set $\ker(C)$ is solid. Recall the recession cone of C,

$$0^+C \triangleq \{d \in Y : y + rd \in C \text{ for any } y \in Y, r \geq 0\}.$$

Proposition 6.14. *Let* C *be a star-shaped set and* $q \in \ker(C)$*. Then* $q + 0^+C \subset \ker(C)$*, and* $\ker(C)$ *is a solid set when* 0^+C *is a solid cone.*

Proof. To prove the first assertion, consider $d \in 0^+(C)$, $y \in C$ and $\alpha \in (0, 1)$. Then

$$\alpha(q + d) + (1 - \alpha)y = \alpha q + (1 - \alpha)y + \alpha d \in C + \alpha 0^+C \subset C$$

where the last inclusion follows from the definition of 0^+C. Hence $q+d \in \ker(C)$. Thus $q + 0^+C \subset \ker(C)$ is true. Also $q + \mathrm{int}(0^+C) \subset \mathrm{int}(\ker(C))$. Therefore if 0^+C is a solid cone, then $\ker(C)$ is solid. □

6.3 Variational approach

Throughout this section we let $X, Y_0, \cdots, Y_m$ be Banach spaces, $G_i : X \rightrightarrows Y_i$ $(i = 0, 1, \cdots, m)$ be multifunctions, Ω be a subset of X, K_0 be a pointed convex cone and K_i be a convex cone in Y_i $(i = 1, \cdots, m)$. We study the following constrained vector optimization problem (cf. Problem (5.17))

$$\begin{aligned} &\text{minimize } G_0(x) \\ &\text{subject to } G_i(x) \cap (-K_i) \neq \emptyset, \quad i = 1, \cdots, m \\ &\qquad x \in \Omega. \end{aligned} \tag{6.5}$$

Following the idea of Definition 5.15, given $\varepsilon > 0$ and $\xi \in K \setminus \{0\}$ we say that a feasible pair $(\bar{x}, \bar{y}_0) \in \mathrm{gph}(G_0)$ is an (ε, ξ)-minimizer of the problem (6.5) if

$$G_0(\widetilde{\Omega}) \cap (\bar{y}_0 - \varepsilon\xi - K_0) = \emptyset \tag{6.6}$$

where $\widetilde{\Omega} \subset X$ is given by $\widetilde{\Omega} \triangleq \Omega \cap \bigcap_{i=1}^{m} G_i^{-1}(-K_i)$.

By applying Theorem 4.1, we can show the following fuzzy necessary condition for approximate efficiency.

Theorem 6.15. *[36, Theorem 4.3] Consider the problem (6.5) where $X, Y_0, \cdots, Y_m$ are all Asplund spaces. Assume* $\mathrm{gph}(G_i)$, K_i *$(0 \leq i \leq m)$ are all closed sets. Let $(\bar{x}, \bar{y}_0) \in \mathrm{gph}(G_0)$ and assume $\bar{x} \in \widetilde{\Omega}$ ($\widetilde{\Omega}$ is given in (6.6)). Given $\varepsilon > 0$ and $\xi \in K \setminus \{0\}$ with $\|\xi\| < 1$. Suppose $(\bar{x}, \bar{y}_0)$ is an (ε, ξ)-minimizer of the problem (6.5). Let $\bar{y}_i \in G_i(\bar{x}) \cap (-K_i)$ for each $i = 1, \cdots, m$. Then for any $\lambda > 0$ there exist $x_i \in D(\bar{x}, \lambda)$, $y_i \in G_i(x_i) \cap D(\bar{y}_i, \lambda)$, $x_{m+1} \in \Omega \cap D(\bar{x}, \lambda)$, $c_i^* \in K_i^*$, $x_i^* \in \hat{D}^* G_i(x_i, y_i)\left(c_i^* + \frac{\varepsilon}{\lambda}\mathbb{D}^*\right) + \frac{\varepsilon}{\lambda}\mathbb{D}^*$, $0 \leq i \leq m$ and $x_{m+1}^* \in \hat{N}(x_{m+1}; \Omega) + \frac{\varepsilon}{\lambda}\mathbb{D}^*$*

such that

$$\sum_{i=0}^{m+1} x_i^* = 0 \text{ and } \frac{1}{2} - \frac{\varepsilon}{\lambda} < \sum_{i=0}^{m} (\|x_i^*\| + \|c_i^*\|) < 1 + \frac{\varepsilon}{\lambda}. \tag{6.7}$$

Proof. Equip the product space $X \times \prod_{i=0}^{m} Y_i$ with the following norm:

$$\|(x, y_0, \cdots, y_m)\| \triangleq \max\left\{\|x\|, \max_{0\le i\le m}\|y_i\|\right\} \text{ for any } (x, y_0, \cdots, y_m) \in X \times \prod_{i=0}^{m} Y_i.$$

We let

$$A_i \triangleq \left\{(x, y_0, \cdots, y_m) \in X \times \prod_{i=0}^{m} Y_i : (x, y_i) \in \mathrm{gph}(G_i)\right\} \quad (i = 0, 1, \cdots, m) \text{ and}$$

$$A_{m+1} \triangleq \Omega \times (\bar{y}_0 - \varepsilon\xi - K_0) \times \prod_{i=1}^{m} (\bar{y}_i - K_i).$$

We claim that $\bigcap_{i=0}^{m+1} A_i = \emptyset$. To show this, suppose to the contrary that there exist $\widetilde{x} \in \Omega$ and $\widetilde{y}_i \in G_i(\bar{x})$ such that

$$\widetilde{y}_0 \in \bar{y}_0 - \varepsilon\xi - K_0 \text{ and } \widetilde{y}_i \in \bar{y}_i - K_i \subset -K_i - K_i = -K_i \ (i = 1, \cdots, m).$$

It turns out that $\widetilde{x} \in \widetilde{\Omega}$ and so $\widetilde{y}_0 \in G_0(\widetilde{\Omega}) \cap (\bar{y}_0 - \varepsilon\xi - K_0)$ contradicting (6.6). We have proved our claim. Now we let $a_0 = \cdots = a_m = (\bar{x}, \bar{y}_0, \cdots, \bar{y}_m)$ and $a_{m+1} = (\bar{x}, \bar{y}_0 - \varepsilon\xi, \bar{y}_1, \cdots, \bar{y}_m)$. It follows that

$$\max_{0\le i\le m} \|a_i - a_{m+1}\| = \|\varepsilon\xi\| < \varepsilon \le \gamma_\infty(A_0, A_1, \cdots, A_{m+1}) + \varepsilon.$$

By Theorem 4.1, there exist $\widetilde{a}_i = (x_i, y_{i,0}, \cdots, y_{i,m}) \in A_i$ and $(x_i^*, y_{i,0}^*, \cdots, y_{i,m}^*) \in X^* \times \prod_{j=0}^{m} Y_j^*$ $(0 \le i \le m+1)$ such that

$$\sum_{i=0}^{m+1} d((x_i^*, y_{i,0}^*, \cdots, y_{i,m}^*), \hat{N}(\widetilde{a}_i; A_i)) < \frac{\varepsilon}{\lambda}, \tag{6.8}$$

$$\begin{aligned} \max_{0\le i\le m+1} \|\widetilde{a}_i - a_i\| = \max_{0\le i\le m} \{&\max\{\|x_i - \bar{x}\|, \max_{0\le k\le m} \|y_{i,k} - \bar{y}_k\|\}, \\ &\max\{\|x_{m+1} - \bar{x}\|, \|y_{m+1,0} - \bar{y}_0 + \varepsilon\xi\|, \max_{1\le k\le m} \|y_{m+1,k} - \bar{y}_k\|\} < \lambda, \end{aligned} \tag{6.9}$$

$$\sum_{i=0}^{m+1}\left(\|x_i^*\|+\sum_{k=0}^{m}\|y_{i,k}^*\|\right)=1 \text{ and} \tag{6.10}$$

$$\sum_{i=0}^{m+1}(x_i^*, y_{i,0}^*, \cdots, y_{i,m}^*)=0. \tag{6.11}$$

Since all K_i are convex cones, using Proposition 2.41 and Proposition 2.27 we obtain

$$\hat{N}(\widetilde{a}_{m+1}; A_{m+1}) \subset \hat{N}(x_{m+1}; \Omega) \times \prod_{i=0}^{m} K_i^* \text{ and}$$

$$\hat{N}(\widetilde{a}_i; A_i) = \{(x^*, y_0^*, \cdots, y_m^*) : (x^*, y_i^*) \in \hat{N}((x_i, y_{i,i}); \text{gph}(G_i)) \text{ and } y_k^* = 0 \text{ for all } k \neq i\}$$

for $0 \le i \le m$. This and (6.8) imply that there exist

$$(\widetilde{x}_i^*, \widetilde{y}_i^*) \in \hat{N}((x_i, y_{i,i}); \text{gph}(G_i)) \quad (0 \le i \le m), \tag{6.12}$$

$$\widetilde{x}_{m+1}^* \in \hat{N}(x_{m+1}; \Omega) \text{ and } (c_0^*, \cdots, c_m^*) \in \prod_{i=0}^{m} K_i^* \tag{6.13}$$

such that

$$\sum_{i=0}^{m+1}\|\widetilde{x}_i^* - x_i^*\| + \sum_{i=0}^{m}\|\widetilde{y}_i^* - y_{i,i}^*\| + \sum_{i,k=0,k\neq i}^{m}\|y_{i,k}^*\| + \sum_{k=0}^{m}\|y_{m+1,k}^* - c_k^*\| < \frac{\varepsilon}{\lambda}. \tag{6.14}$$

It follows from (6.11) and (6.14) that for all $k = 0, \cdots, m$,

$$\begin{aligned}
-\widetilde{y}_k^* &= -\widetilde{y}_k^* + 0 \\
&= -\widetilde{y}_k^* + (c_k^* - c_k^*) + \left(y_{k,k}^* + y_{m+1,k}^* + \sum_{i=0,i\neq k}^{m} y_{i,k}^*\right) \\
&= c_k^* + \left[(y_{k,k}^* - \widetilde{y}_k^*) + \sum_{i=0,i\neq k}^{m} y_{i,k}^* + (y_{m+1,k}^* - c_k^*)\right] \in c_k^* + \frac{\varepsilon}{\lambda}\mathbb{D}^*.
\end{aligned}$$

By (6.11)-(6.14), one has

$$x_k^* \in \hat{D}^* G_k(x_k, y_{k,k})\left(c_k^* + \frac{\varepsilon}{\lambda}\mathbb{D}^*\right) + \frac{\varepsilon}{\lambda}\mathbb{D}^* \ (k = 0, 1, \cdots, m),$$

$$x_{m+1}^* \in \hat{N}(x_{m+1}; \Omega) + \frac{\varepsilon}{\lambda}\mathbb{D}^* \text{ and } \sum_{i=0}^{m+1} x_i^* = 0.$$

It remains to show (6.7). Indeed, since $\sum_{i=0}^{m} x_i^* = -x_{m+1}^*$ by (6.11), we have from the triangle inequality that

$$\sum_{i=0}^{m+1} \|x_i^*\| \leq 2\sum_{i=0}^{m} \|x_i^*\|. \tag{6.15}$$

Again by (6.11), for all $k = 0, \cdots, m$, one has $-y_{k,k}^* = y_{m+1,k}^* + \sum_{i=0, i\neq k}^{m} y_{i,k}^*$ and so

$$\sum_{k=0}^{m} \|y_{k,k}^*\| \leq \sum_{k=0}^{m} \|y_{m+1,k}^*\| + \sum_{i,k=0, i\neq k}^{m} \|y_{i,k}^*\|.$$

Consequently

$$\begin{aligned}\sum_{i=0}^{m+1}\sum_{k=0}^{m} \|y_{i,k}^*\| &= \sum_{k=0}^{m} \|y_{m+1,k}^*\| + \sum_{i,k=0,i\neq k}^{m} \|y_{i,k}^*\| + \sum_{k=0}^{m} \|y_{k,k}^*\| \\ &\leq 2\left(\sum_{k=0}^{m} \|y_{m+1,k}^*\| + \sum_{i,k=0,i\neq k}^{m} \|y_{i,k}^*\|\right).\end{aligned} \tag{6.16}$$

By adding up the estimates (6.15) and (6.16) and making use of (6.10), we get

$$1 \leq 2\left(\sum_{i=0}^{m} \|x_i^*\| + \sum_{k=0}^{m} \|y_{m+1,k}^*\| + \sum_{i,k=0,i\neq k}^{m} \|y_{i,k}^*\|\right)$$

and hence (by triangle inequality and (6.14))

$$\begin{aligned}\frac{1}{2} &\leq \sum_{i=0}^{m} \|x_i^*\| + \sum_{k=0}^{m} \|y_{m+1,k}^*\| + \sum_{i,k=0,i\neq k}^{m} \|y_{i,k}^*\| \\ &\leq \sum_{i=0}^{m} \|x_i^*\| + \sum_{k=0}^{m} (\|y_{m+1,k}^* - c_k^*\| + \|c_k^*\|) + \sum_{i,k=0,i\neq k}^{m} \|y_{i,k}^*\| \\ &\leq \sum_{i=0}^{m} (\|x_i^*\| + \|c_i^*\|) + \sum_{k=0}^{m} \|y_{m+1,k}^* - c_k^*\| + \sum_{i,k=0,i\neq k}^{m} \|y_{i,k}^*\| \\ &< \sum_{i=0}^{m} (\|x_i^*\| + \|c_i^*\|) + \frac{\varepsilon}{\lambda}.\end{aligned}$$

Thus the first inequality in (6.7) is true. Moreover by (6.10) and (6.14) we know that

$$\sum_{i=0}^{m} (\|x_i^*\| + \|y_{m+1,i}^*\|) \leq 1 \text{ and } \sum_{i=0}^{m} \|c_i^* - y_{m+1,i}^*\| < \frac{\varepsilon}{\lambda}.$$

The triangle inequality guarantees the second inequality in (6.7) holds. We thus complete the proof. □

Bibliography

[1] E. Asplund, *Fréchet differentiability of convex functions*, Acta Mathematica, 121 (1968), pp. 31-47.

[2] T. Q. Bao and B. S. Mordukhovich, *Relative Pareto minimizers for multiobjective problems: existence and optimality conditions*, Mathematical Programming, 122 (2010), pp. 301-347.

[3] T. Q. Bao and B. S. Mordukhovich, *Extended Pareto optimality in multiobjective problems*, Chapter of the book Recent Advances in Vector Optimization (Q. H. Ansari and J. - C. Yao, eds.), Research Report #11 (2010), 46 pp.; Springer, Berlin.

[4] L. D. Berkovitz, *Convexity and Optimization in* $\mathbb{R}^n$, John Wiley and Sons, Inc., New York (2002).

[5] J. M. Borwein and R. Goebel, *Notions of relative interior in Banach spaces*, Journal of Mathematics Sciences, 115 (2003), pp. 2542-2553.

[6] J. M. Borwein and Q. J. Zhu, *Techniques of Variational Analysis*, CMS Books in Mathematics, Springer, New York (2005).

[7] J. Dutta and C. Tammer, *Lagrangian conditions for vector optimization in Banach spaces*, Mathematical Methods of Operations Research, 64 (2006), pp. 521-540.

[8] M. Durea and C. Tammer, *Fuzzy necessary optimality conditions for vector optimization problems*, Optimization, 58 (2009), pp. 449-467.

[9] I. Ekeland, *Remarques sur les problems variationels 1*, CRAS, 275 (1972), pp. 1057-1059.

[10] I. Ekeland, *On the variational principle*, Journal of Mathematical Analysis and Applications, 47 (1974), pp. 324-353.

[11] M. Fabian, *Subdifferentiability and trustworthiness in the light of a new variational principle of Borwein and Preiss*, Acta Universitatis Carolinae - Mathematica et Physica, 30 (1989), pp. 51-56.

[12] F. Flores-Bazán and B. Jiménez, *Strict Efficiency in Set-valued Optimization*, SIAM Journal of Control and Optimization, 48 (2009), pp. 881-908.

[13] C. Gerstewitz, *Nichtkonvexe Dualität in der Vektoroptimierung*, Wissenschaftliche Zeitschrift den Technischen Hochschule "Carl Schorlemmer" Leuna-Merseburg, 25 (1983), pp. 357-364.

[14] C. Gerstewitz and E. Iwanow, *Dualität für nichtkonvexe Vektoroptimierungsprobleme*, Workshop on Vector Optimization (Plauen, 1984), Wissenschaftliche Zeitschrift der Technischen Hochschule Ilmenau, 31 (1985), pp. 61-81.

[15] A. Göpfert, H. Riahi, C. Tammer and C. Zălinescu, *Variational Methods in Partially Ordered Spaces*, CMS Books in Mathematics, 17, Springer, New York (2003).

[16] C. Gutiérrez ,B. Jiménez and V. Novo, *A unified approach and optimality conditions for approximate solutions of vector optimization problems*, SIAM Journal of Optimization, 3 (2006), pp. 688-710.

[17] J.-B. HIRIART-URRUTY, *New Concepts in Nondifferentiable Programming, Analyse non convexe*, Bull. Soc. Math. France, 60 (1979), pp. 57-85.

[18] J.-B. HIRIART-URRUTY, *Tangent Cones, Generalized Gradients and Mathematical Programming in Banach Spaces*, Mathematics of Operations Research, 4 (1979), pp. 79-97.

[19] J.-B. HIRIART-URRUTY AND C. LEMARÉCHAL, *Convex Analysis and Minimization Algorithms I*, Grundlehren der mathematischen Wissenschaften 305, Springer-Verlag, Berlin (1993).

[20] J.-B. HIRIART-URRUTY AND C. LEMARÉCHAL, *Fundamentals of Convex Analysis*, Springer-Verlag, Berlin (2001).

[21] J. JAHN, *Vector Optimization: Theory, Applications and Extensions*, Springer, Berlin (2004).

[22] S. S. KUTATELADZE, *Convex ε-programming*, Soviet Mathematics - Doklady, 20 (1979), pp. 391-393.

[23] D. T. LUC, *Theory of Vector Optimization*, Springer-Verlag, Berlin (1989).

[24] R. E. MEGGINSON, *An Introduction to Banach Space Theory*, Springer-Verlag (1998).

[25] B. S. MORDUKHOVICH, *Variational Analysis and Generalized Differentiation, I: Basic Theory*, Grundlehren Series (Fundamental Principles of Mathematical Sciences) 330. Springer, Berlin (2006).

[26] B. S. MORDUKHOVICH, *Variational Analysis and Generalized Differentiation, II: Applications*, Grundlehren Series (Fundamental Principles of Mathematical Sciences) 330. Springer, Berlin (2006).

[27] B. S. Mordukhovich and Y. H. Shao, *Extremal characterization of Asplund spaces*, Proceedings of the American Mathematical Society, 124 (1996), pp. 197-205.

[28] A. B. Németh, *A nonconvex vector minimization problem*, Nonlinear Analysis, 10 (1986), pp. 669-678.

[29] R. R. Phelps, *Convex Functions, Monotone Operators and Differentiability*, Lecture Notes in Mathematics 1364, Springer-Verlag (1989).

[30] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, Springer, Berlin (1988).

[31] W. Schirotzek, *Nonsmooth Analysis*, Springer, Berlin (2007).

[32] A. Zaffaroni, *Degrees of efficiency and degrees of minimality*, SIAM Journal on Control and Optimization, 42 (2003), pp. 1071-1086.

[33] C. Zălinescu, *Convex Analysis in General Vector Spaces*, World Scientific, New Jersey (2002).

[34] X. Y. Zheng and K. F. Ng, *The Fermat rule for multifunctions on Banach spaces*, Math. Program. Ser. A, 104 (2005), pp. 69-90.

[35] X. Y. Zheng and K. F. Ng, *The Lagrange multiplier rule for multifunction in Banach spaces*, SIAM Journal of Optimization, 17 (2006), pp. 1154-1175.

[36] X. Y. Zheng and K. F. Ng, *A unified separation theorem for closed sets in a Banach space and optimality conditions for vector optimization*, unpublished (2011).

Index