January 2019

Second-Order Generalized Differentiation Of Piecewise Linear-Quadratic Functions And Its Applications

Hong Do
Wayne State University, dongochong@yahoo.com

SECOND-ORDER GENERALIZED DIFFERENTIATION OF PIECEWISE LINEAR-QUADRATIC FUNCTIONS AND ITS APPLICATIONS

by

HONG DO

DISSERTATION

Submitted to the Graduate School,

of Wayne State University,

Detroit, Michigan

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

2019

MAJOR: MATHEMATICS

Approved By:

____
Advisor Date

____

____

____

____

DEDICATION

To my parents

Nguyen Thi Nhi *and* Do Ngoc Tuynh,

and my husband

Ho Hoai Anh

ACKNOWLEDGEMENTS

First and foremost, I would like to express the deepest gratitude and appreciation to my advisor, Professor Boris Mordukhovich, for his endless support during my graduate program at Wayne State University. He has introduced the theory and applications of variational analysis to me, kindly offered me the opportunity to do research with him, and thoughtfully guided me to overcome many difficulties during my study and my life. It is my fortune and honor to have him as my advisor.

It is my pleasure to acknowledge Dr. Ebrahim Sarabi for his significant contribution in our joint work [3] with my advisor which is used in Chapter 4 and 5 of my dissertation. Ebrahim has taught me a lot during the last five years as well as shared with me his brilliant ideas. Without Professor Boris Mordukhovich and Dr. Ebrahim Sarabi, I would not be able to complete this dissertation.

I am taking this opportunity to thank Professor Peiyong Wang, Professor Alper Murat, Professor George Yin, and Professor Cathering Lebiedzik for serving on my committee. I would like to thank the entire Department of Mathematics for their kind support in a number of ways.

I want to express my gratitude to my undergraduate advisor, Professor A. A. Lodkin who taught me a lot during my time at SPBGU. His recommendation letter for me played a significant role in getting me into the PhD program at WSU.

I would like to thank my high school math teachers, Professor Nguyen Vu Luong and Professor Le Dinh Thinh, and my supervisor at Hanoi University of Science, Professor Dang Dinh Chau, for inspiring me in mathematics and encouraging me to pursue higher education in mathematics.

I also own my thanks to all of my friends at Wayne State University, especially, Richard Pineau, Tuan Hoang, Ky Tran, Dat Pham, Dang Nguyen, Trang Bui, Hussein Nasralah, Ba Nguyen and The Tran.

Most of all, I wish to thank my parents, my husband and my kids for their endless care, help, love and patience.

TABLE OF CONTENTS

Dedication ii

Acknowledgements iii

Chapter 1 Introduction **1**

Chapter 2 Preliminaries from Variational Analysis **7**

Chapter 3 Generalized Differentiation of Piecewise Linear-Quadratic Functions **10**

3.1 Calculations of the Second-order Generalized Differentiation 10

3.2 Second-order Subdifferential and Full Stability in Constrained Optimization 19

Chapter 4 Multiplier Criticality in Piecewise Linear-Quadratic Settings **22**

4.1 Equivalent Description of Criticality 22

4.2 Uniqueness of Lagrange Multipliers and Isolated Calmness 27

4.3 Characterizations of Noncritical Multipliers 30

4.4 Noncriticality in Extended Nonlinear Programming 39

Chapter 5 Noncriticality and its applications in stability **44**

5.1 Critical Multipliers and Full Stability of Minimizers in ENLPs 44

5.2 Noncriticality and Lipschitzian Stability of Solutions to ENLPs 47

References **56**

Abstract 60

Autobiographical Statement 61

CHAPTER 1 INTRODUCTION

This dissertation is devoted to the study and applications of the second-order generalized differentiation of piecewise linear-quadratic functions. Namely, we study the its subclass that consists of functions $\theta_{Y,B}(u) : \mathbb{R}^m \to \bar{\mathbb{R}}$ defined by

$$\theta_{Y,B}(u) := \sup_{y\in Y}\Big\{\langle y, u\rangle - \frac{1}{2}\langle y, By\rangle\Big\}, \tag{1.1}$$

where $Y := \{x \in \mathbb{R}^m | \langle b_i, x\rangle \leq \alpha_i, i = 1, 2, \ldots, p\}$ in $\mathbb{R}^m$ is a nonempty convex polyhedral set and B is a symmetric, positive-semidefinite matrix in $\mathbb{R}^{m\times m}$.

The first goal of this thesis is to compute the second-order subdifferential of the functions described above. The calculations in this part will be later applied in the study of the stability of composite optimization problems known as extended nonlinear programming associated with piecewise linear-quadratic functions.

The second goal of the dissertation is to study a remarkable class of optimization problems given in the following, formally unconstrained, *composite format*:

$$\text{minimize } \ \varphi(x) := \varphi_0(x) + \theta\big(\Phi(x)\big), \quad x \in \mathbb{R}^n, \tag{1.2}$$

where $\varphi_0\colon \mathbb{R}^n \to \mathbb{R}$ is an original cost function and $\Phi\colon \mathbb{R}^n \to \mathbb{R}^m$ is a constraint mapping, both are twice differentiable at the reference points unless otherwise stated, and where $\theta\colon \mathbb{R}^m \to \overline{\mathbb{R}} := (-\infty, \infty]$ is of the form (1.1) defined above.

Note that the unconstrained composite format (1.2) gives us a convenient representation of the constrained optimization problem to minimize the cost function $\varphi_0(x)$ subject to the inclusion constraint $\Phi(x) \in \Theta := \{u \in \mathbb{R}^m |\ \theta(u) < \infty\}$. In particular, conventional nonlinear programs (NLPs) with s inequality constraints and $m - s$ equality constraints described by $\mathcal{C}^2$-smooth functions can be written in the composite format (1.2), where $\theta := \delta_\Theta$ is the indicator function of the polyhedron $\Theta := \mathbb{R}^s_- \times \{0\}^{m-s}$ that is equal to 0 on Θ and to ∞ otherwise.

Problems of the ENLP type (1.2) with θ given by (1.1) were introduced by Rockafellar [36] under the name of *extended nonlinear programs* (ENLPs). It has been realized over the years that ENLPs in this form provide a suitable framework for developing both theoretical and computational aspects of optimization in broad classes of constrained problems that include stochastic programming, robust optimization, etc. The special expression (1.1) for the extended-real-valued function θ, known as the *dualizing representation* or the *piecewise linear-quadratic penalty*, is significant for the theory and applications of Lagrange multipliers in the Karush-Kuhn-Tucker (KKT) systems associated with the ENLPs under consideration.

It is not hard to check (see more details in Section 4.4) that KKT systems associated with local optimal solutions to ENLPs are included in the following more general class of *variational systems* of the *subdifferential type*

$$\Psi(x,\lambda) := f(x) + \nabla\Phi(x)^*\lambda = 0,\ \lambda \in \partial\theta\big(\Phi(x)\big) \ \text{ with } \ \theta = \theta_{Y,B}, \tag{1.3}$$

where $f\colon \mathbb{R}^n \to \mathbb{R}^n$ is a differentiable mapping while $\Phi\colon \mathbb{R}^n \to \mathbb{R}^m$ is a twice differentiable mapping in the classical sense [37, Definition 13.1(i)], where $\theta_{Y,B}$ is taken from (1.1), where * indicates the matrix transposition/adjoint operator, and where ∂ stands for the subdifferential of convex analysis.

In the second part of this thesis the multiplier criticality is studied systematically for variational systems of type (1.3) with applications to KKT systems in ENLPs.

The notions of critical and noncritical multipliers were first introduced by Izmailov [9] for the classical KKT systems corresponding to *NLPs with equality constraints* described by $\mathcal{C}^2$-smooth functions. It has been realized from the very beginning that the presence of critical multipliers plays a *negative* role in numerical optimization and is largely responsible for primal slow convergence in primal-dual algorithms of the Newtonian type. Further strong developments in this direction for NLPs and related variational inequalities have been done over the years, mainly by Izmailov, Solodov, and their collaborators; see, e.g., the book [10] and the survey paper [11], which is entirely devoted to critical multipliers. The criticality definitions in the above publications are heavily based on the specific structures of NLPs and related variational inequalities.

In [32], Mordukhovich and Sarabi suggested new definitions of critical and noncritical multipliers for a general class of *subdifferential variational systems* of type (1.3), where θ may be even a nonconvex extended-real-valued function. The given definitions in [32] are expressed via second-order generalized differential constructions of variational analysis while reduced to those from [9, 10] for the classical KKT systems corresponding to NLPs. Furthermore, for extended-real-valued *convex piecewise linear* (CPWL) functions θ in (1.3), which include (1.1) when $B = 0$, the definitions of critical and noncritical multipliers are expressed in [32] entirely in terms of the problem data with the subsequent characterizations of criticality and various applications to optimization and stability problems for such systems.

The quite recent paper of the same authors [33] contains counterparts of some major results from [32] with developing also novel issues on criticality for variational systems described by

$$f(x) + \nabla\Phi(x)^*\lambda = 0,\ \lambda \in N_\Theta\big(\Phi(x)\big), \tag{1.4}$$

where f and Φ are the same as in (1.3), and where N_Θ is the normal cone to a $\mathcal{C}^2$*-cone reducible* set $\Theta \subset \mathbb{R}^m$. This framework covers, in particular, KKT systems associated with general problems of (nonpolyhedral) *conic programming*; see, e.g., [1].

The main results in this dissertation extend those from [32], obtained for CPWL functions θ, to the case of functions $\theta_{Y,B}$ defined in (1.1), which form a major class of extended-real-valued convex *piecewise linear-quadratic* functions in variational analysis; see [37] and Section 2 below. At the same time, the new results obtained here are completely independent from those derived for the variational system (1.4) in [32] in the case of nonpolyhedral sets Θ therein.

The basic tools of first-order and second-order generalized differentiation employed in this thesis are *subdifferential* developed by Mordukhovich and *subgradient graphical derivative* developed by Rockafellar; see [20, 37] and the references therein. Using these tools allows us to establish verifiable characterizations of noncritical multipliers in the general setting of (1.3), to characterize the uniqueness of Lagrange multipliers in (1.3), to ensure noncriticality for ENLPs via a new second-order optimality condition, which is employed in turn to verify the important stability property of solutions to KKT systems that is known as robust isolated calmness and is related to noncriticality. We also reveal a relationship between the isolated calmness and Lipschitz-like properties of solution maps for canonically perturbed variational systems with the piecewise linear-quadratic term (1.1).

As mentioned above, the existence of critical multipliers is a negative factor in convergence analysis, since it seems to prevent primal superlinear convergence of major primal-dual algorithms. Thus it is crucial to find verifiable conditions, expressed entirely in terms of the problem data in question, which ensure that critical multipliers corresponding to this minimizer do not arise. It is conjectured in [21], based on preliminary results for NLPs, that *full stability* of local minimizers in the sense of [15] *rules out* the appearance of *critical multiplies*. This conjecture was verified in [32] for polyhedral problems of type (1.2) with convex piecewise linear functions θ. Now we justify this conjecture in the general case of ENLPs with piecewise linear-quadratic functions $\theta_{Y,B}$ in form (1.1).

The dissertation is organized as follows. In Chapter 2 we present some definitions and facts from variational analysis and generalized differentiation that are broadly employed throughout the whole dissertation. Other variational constructions and results are recalled in those places of the subsequent sections where they are actually used.

Chapter 3 presents calculations of the second-order subdifferential of piecewise linear-quadratic functions (1.1). The main result in this chapter will be applied later in Section 5.1.

Chapter 4 contains basic *definitions* of *critical* and *noncritical multipliers* for variational systems (1.3) involving piecewise linear-quadratic functions of type (1.1) with providing equivalent descriptions, examples, and discussions. In Section 4.2 we obtain new results on the relationship between the well-recognized *calmness* and *isolated calmness* properties of multiplier maps associated with the variational systems (1.3) with the piecewise linear-quadratic term (1.1) and the *uniqueness* of Lagrange multipliers in such systems. This is certainly of its independent interest, while the developed approach and results can be viewed as the preparation to the subsequent characterizations of noncritical multipliers in the variational systems under consideration.

Section 4.3 establishes major *characterizations* of *noncritical multipliers* for systems (1.3) with $\theta_{Y,B}$ taken from (1.1) via a novel *semi-isolated calmness* property for solution maps to canonical perturbations of (1.3) and also via two new *error bounds* that are specific for the variational systems (1.3) with the piecewise linear-quadratic term (1.1).

Section 4.4 is devoted to noncritical multipliers in *KKT systems* associated with *ENLPs* for which the results of the previous sections are automatically applied with the specification of Ψ in (1.3) as the x-partial gradient of the appropriate Lagrangian. The main new result here, that is characteristic to the optimization framework, is a novel *second-order sufficient condition* for strict local minimizers, which also ensures that all the corresponding multipliers are noncritical.

In Section 5.1 we justify, for the case of ENLPs from (1.2) and (1.1), the aforementioned conjecture on *excluding critical multipliers* corresponding to a *fully stable* local minimizer for the given ENLP. The proof of this result is based on characterizations of noncriticality via semi-isolated calmness obtained in Section 4.3.

The last Section 5.2 provides applications of the developed characterizations of noncritical multipliers for the variational systems under consideration to the study of an important stability property of solution maps to KKT systems associated with ENLPs. This property of set-valued mappings has been recently recognized as *robust isolated calmness*. The results obtained above allow us to characterize robust isolated calmness via the noncriticality and uniqueness of Lagrange multipliers on one side and via the new second-order optimality condition for ENLPs on the other. Finally, we characterize the Lipschitz-like/Aubin property of solution maps to perturbed variational systems and establish its relationship with isolated calmness.

CHAPTER 2 PRELIMINARIES FROM VARIATIONAL ANALYSIS

In this chapter we review basic notions of generalized differentiation in variational analysis (see [37]) and then recall important facts that will be used later. The notation we use are standard in variational analysis; see [22, 37].

Given a nonempty subset $\Omega \subset \mathbb{R}^d$ and a point $\bar{z} \in \Omega$, its *prenormal cone* (also called the regular or Fréchet normal cone) at $\bar{z} \in \Omega$ is defined as

$$\widehat{N}(\bar{z};\Omega) := \Big\{ v \in \mathbb{R}^m \Big| \limsup_{u \xrightarrow{\Omega} \bar{z}} \frac{\langle v, u - z \rangle}{\|u - z\|} \le 0 \Big\}, \quad z \in \Omega, \tag{2.1}$$

where the symbol $u \xrightarrow{\Omega} \bar{z}$ indicates that $u \to \bar{z}$ with $u \in \Omega$. The (Mordukhovich) *limiting normal cone* (or just *normal cone*) to Ω at z is defined via prenormal cone by

$$N(\bar{z};\Omega) = N_\Omega(\bar{z}) := \Big\{ v \in \mathbb{R}^m \Big| \text{there exist } z_k \xrightarrow{\Omega} \bar{z} \text{ and } v_k \in \widehat{N}(z_k;\Omega) \text{ with } v_k \to v \text{ as } k \to \infty \Big\}. \tag{2.2}$$

A prenormal cone is always a closed and convex cone while a normal cone is a closed cone but is usually nonconvex. If the set Ω is convex, two sets in (2.1) and (2.2) are the same and both reduce to the classical normal cone of convex analysis defined as follows:

$$N_\Omega(\bar{z}) := \big\{ v \in \mathbb{R}^n \big| \ \langle v, z - \bar{z} \rangle \le 0 \ \text{ for all } \ z \in \Omega \big\}. \tag{2.3}$$

The (Bouligand-Severi) *tangent/contingent cone* $T_\Omega(z)$ to Ω at $\bar{z}$ is defined by

$$T_\Omega(\bar{z}) := \Big\{ w \in \mathbb{R}^d \Big| \ \exists\, z_k \xrightarrow{\Omega} \bar{z}, \ \exists\, \alpha_k \ge 0 \ \text{ with } \ \alpha_k(z_k - z) \to w \ \text{ as } \ k \to \infty \Big\}, \tag{2.4}$$

where the symbol $z \xrightarrow{\Omega} \bar{z}$ indicates that $z \to \bar{z}$ with $z \in \Omega$. The *critical cone* to Ω at $\bar{z}$ for $\bar{v} \in N_\Omega(\bar{z})$ is expressed via the tangent cone (2.4) as

$$K_\Omega(\bar{z}, \bar{v}) := T_\Omega(\bar{z}) \cap \{\bar{v}\}^\perp \tag{2.5}$$

with the notation $\{\bar{v}\}^\perp := \big\{ w \in \mathbb{R}^n \big| \ \langle w, v \rangle = 0 \big\}$.

Let $\varphi : \mathbb{R}^m \to \overline{\mathbb{R}}$ be an extended-real-valued function. The *epigraph* of φ , denoted by $\mathrm{epi}\varphi$, is the set $\mathrm{epi}\varphi := \{(x, y) \in \mathbb{R}^m \times \mathbb{R} \| x \in \mathrm{dom}\varphi, y \ge \varphi(x)\}$. For $\bar{z} \in \mathrm{dom}\varphi$, the *basic* *subdifferential* and *singular subdifferential* of φ at $\bar{z}$ are given, respectively, by

$$\partial\varphi(\bar{z}) := \big\{v \in \mathbb{R}^m \big| (v,-1) \in N\big((\bar{z},\varphi(\bar{z});\text{epi}\varphi)\big)\big\}, \tag{2.6}$$

$$\partial^\infty\varphi(\bar{z}) := \big\{v \in \mathbb{R}^m \big| (v,0) \in N\big((\bar{z},\varphi(\bar{z});\text{epi}\varphi)\big)\big\}. \tag{2.7}$$

If the function φ is convex, the basic subdifferential defined in (2.6) agrees with the subdifferential of convex analysis defined as follows:

$$\partial\varphi(\bar{z}) := \big\{v \in \mathbb{R}^n \big|\ \langle v, z-\bar{z}\rangle \le \varphi(z)-\varphi(\bar{z}) \ \text{ for all } \ z \in \mathbb{R}^n\big\}, \tag{2.8}$$

and $\partial^\infty\varphi(\bar{z}) = N(\bar{z};\text{dom}\varphi)$. Besides, for any set Ω one has

$$N(\bar{z};\Omega) = \partial\delta(\bar{z};\Omega) = \partial^\infty\delta(\bar{z};\Omega), \quad \bar{z} \in \Omega,$$

where δ_Ω is the indicator function defined by $\delta_\Omega(z) = \delta(z;\Omega) := 0$ for $z \in \Omega$ and $\delta(z;\Omega) := \infty$ otherwise.

Next we consider the *second subderivative* of φ at $(\bar{x},\bar{y})$, $\bar{y} \in \mathbb{R}^n$, in the direction $\bar{w}$ is defined by

$$\mathrm{d}^2\varphi(\bar{x},\bar{y})(\bar{w}) := \liminf_{\substack{t\downarrow 0\\ w\to\bar{w}}} \frac{\varphi(\bar{x}+tw)-\varphi(\bar{x})-t\langle\bar{y},\, w\rangle}{\frac{1}{2}t^2}. \tag{2.9}$$

For a set-valued mapping $F\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^p$, define its *domain* and *graph* by, respectively,

$$\text{dom}\, F := \big\{x \in \mathbb{R}^n \big|\ F(x) \ne \emptyset\big\} \ \text{ and } \ \text{gph}\, F := \big\{(x,y) \in \mathbb{R}^n \times \mathbb{R}^p \big|\ y \in F(x)\big\}.$$

The *limiting coderivative* of F at $\bar{x},\bar{y}) \in \text{gph}\, F$ is given by

$$D^*F(\bar{x},\bar{y})(v) := \big\{u \in \mathbb{R}^n \big| (u,-v) \in N\big((\bar{x},\bar{y});\text{gph}\, F\big)\big\}, \quad v \in \mathbb{R}^p, \tag{2.10}$$

and the *graphical derivative* of F at $(\bar{x},\bar{y}) \in \text{gph}\, F$ is given by

$$DF(\bar{x},\bar{y})(u) := \big\{v \in \mathbb{R}^p \big|\ (u,v) \in T_{\text{gph}\, F}(\bar{x},\bar{y})\big\}, \quad u \in \mathbb{R}^n. \tag{2.11}$$

In this thesis we also use another second-order generalized derivative of an extended-real-valued convex function $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ at $\bar{z} \in \text{dom}\,\varphi$ for $\bar{v} \in \partial\varphi(\bar{z})$ that is defined via the graphical derivative (2.11) of the subgradient mapping $\partial\varphi\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ under the name of the *subgradient graphical derivative* by

$$D\partial\varphi(\bar{x},\bar{z})(u) := D\big(\partial\varphi\big)(\bar{z},\bar{v})(u), \quad u \in \mathbb{R}^n. \tag{2.12}$$

We now formulate the basic facts about the functions $\theta_{Y,B}$ taken from (1.1) that are systematically exploited in this work. The proofs of these facts can be found in [37, Examples 11.18, 13.23 and Theorem 13.40]. Recall that the *horizon cone* of a nonempty set $Y \subset \mathbb{R}^m$ used below is defined by

$$Y^\infty := \big\{y \in \mathbb{R}^m \big|\ \exists\, y_k \in Y,\ \exists\, \lambda_k \downarrow 0 \ \text{ with } \ \lambda_k y_k \to y\big\}.$$

Recall also [37, Definition 10.20] that a function $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ is *piecewise linear-quadratic* if its domain $\operatorname{dom}\varphi$ can be represented as the union of finitely many convex polyhedral sets, relative to each of which $\varphi(x)$ is given by an expression of the form $\frac{1}{2}\langle x, Ax\rangle + \langle a, x\rangle + \alpha$ for some scalar $\alpha \in \mathbb{R}$, vector $a \in \mathbb{R}^n$, and $n \times n$ symmetric matrix A.

Theorem 2.1. (properties of piecewise linear-quadratic penalties). *Let $\theta_{Y,B}$ be defined by* (1.1)*. Then the following properties hold:*

(i) *The function $\theta_{Y,B}$ is a proper and convex piecewise linear-quadratic with the domain*

$$\operatorname{dom}\theta_{Y,B} = \big(Y^\infty \cap \ker B\big)^*.$$

(ii) *The subdifferential* (2.8) *of $\theta_{Y,B}$ is calculated by*

$$\partial\theta_{Y,B}(u) = \mathop{\arg\max}_{y\in Y}\Big\{\langle y, u\rangle - \frac{1}{2}\langle y, By\rangle\Big\} = (N_Y + B)^{-1}(u), \quad u \in \mathbb{R}^m. \tag{2.13}$$

(iii) *Given any $(\bar{z},\bar{\lambda}) \in \operatorname{gph}\partial\theta_{Y,B}$, the second subderivative* (2.9) *is calculated by*

$$\mathrm{d}^2\theta_{Y,B}(\bar{z},\bar{\lambda})(u) = 2\theta_{\mathcal{K},B}(u) := \sup_{w\in\mathcal{K}}\big\{2\langle w, u\rangle - \langle w, Bw\rangle\big\}, \quad u \in \mathbb{R}^m, \tag{2.14}$$

in the same form $\theta_{\mathcal{K},B}(u)$ as in (1.1) *with the replacement of Y by critical cone $\mathcal{K} := K_Y(\bar{\lambda}, \bar{z} - B\bar{\lambda})$ defined via* (2.5)*. Furthermore, the subgradient graphical derivative* (2.12) *of $\theta_{Y,B}$ at $\bar{z}$ for $\bar{\lambda}$ is represented as*

$$D\partial\theta_{Y,B}(\bar{z},\bar{\lambda})(u) = \partial\theta_{\mathcal{K},B}(u), \quad u \in \mathbb{R}^m. \tag{2.15}$$

CHAPTER 3 GENERALIZED DIFFERENTIATION OF PIECEWISE LINEAR-QUADRATIC FUNCTIONS

This chapter consists of two sections. In the first section we compute the second-order subdifferential of functions of the form (1.1). A calculation of its second-order subdifferential can be found in [[37], Lemma 4.4], but the result is not formulated in terms of the problem initial data. In this chapter we will compute its second-order subdifferential completely in terms of the initial data, i.e. the polyhedral Y and the matrix B. The second section of the chapter presents an application of the calculations in the first section.

3.1 Calculations of the Second-order Generalized Differentiation

For brevity in what follows we denote by θ the function $\theta_{Y,B}$. Assume that $\bar{z} \in dom\theta$. We will compute the *second-order subdifferential* of θ of θ at $\bar{z}$ relative to $\bar{v} \in \partial\theta(\bar{z})$ defined by

$$\partial^2\theta(\bar{z},\bar{v})(w) := D^*\partial\theta(\bar{z},\bar{v})(w). \tag{3.1}$$

We will compute the cones $\widehat{N}((\bar{z},\bar{v}); gph\partial\theta)$ and $N((\bar{z},\bar{v}); gph\partial\theta)$ and then apply the definition (2.10) to get the formula for the second-order subdifferential of θ at $\bar{z}$ relative to $\bar{v}$.

Denote by $\mathfrak{K}$ the set $\big\{J \subset I(\bar{v}) | \bar{z} - B\bar{v} \in cone\{b_i,\, i \in J\}\big\}$, where $I(\bar{v})$ is the set of active constraints of the set Y given by $I(\bar{v}) := \big\{i \in \{1,2,\ldots,p\} \big| \langle \bar{v}, b_i\rangle = \alpha_i\big\}$. Since $\mathfrak{K}$ is a finite set, we may enumerate its elements and rewrite $\mathcal{K}$ as $\mathfrak{K} = \{J_1, J_2, \ldots, J_l\}$, where $J_l := I(\bar{v})$. The following proposition gives a formula for the prenormal cone to the graph $\mathrm{gph}\partial\theta$ at the point $(\bar{z},\bar{v})$.

Proposition 3.1. The normal cone $\widehat{N}\big((\bar{z},\bar{v}); \mathrm{gph}\partial\theta\big)$ is calculated by the formula

$$\widehat{N}\big((\bar{z},\bar{v}); \mathrm{gph}\partial\theta\big) = \big\{(u,w) \in \mathbb{R}^m \times \mathbb{R}^m | (u, Bu + w) \in K_Y(\bar{v}, \bar{z} - B\bar{v}) \times \mathcal{A}(\bar{z},\bar{v})\big\}, \tag{3.2}$$

where $\mathcal{A}(\bar{z},\bar{v}) := \bigcap_{k=1}^{l} \big[cone\{b_i, i \in I(\bar{v}) \setminus J_k\} + span\{b_i, i \in J_k\}\big]$.

Proof. By the definition of active constraints, $\langle \bar{v}, b_i\rangle = \alpha_i$ if $i \in I(\bar{v})$ and $\langle \bar{v}, b_i\rangle < \alpha_i$ if

$i \notin I(\bar{v})$. Therefore there is a neighborhood U_0 of $\bar{v}$ such that $\langle v, b_i\rangle < \alpha_i$ for all $i \notin I(\bar{v})$ and for all $v \in U_0$, which means $i \notin I(v)$. It follows from here that $I(v) \subset I(\bar{v})$. Since $\partial\theta(z) = (N_Y + B)^{-1}(z)$ by the formula (2.13), it is deduced from $(z,v) \in \text{gph}\partial\theta$ that $z - Bv \in N_Y(v)$. If $\bar{z} - B\bar{v} \notin N_Y(v)$ for some v close to $\bar{v}$, we can choose a neighborhood $U(\bar{z},\bar{v})$ sufficiently small such that $z - Bv \notin N_Y(v)$ for all $(z,v) \in U(\bar{z},\bar{v})$. Since there is a finite number of normal cones $N_Y(v)$ to Y at v, for all points (z,v) close to $(\bar{z},\bar{v})$ it belongs to $\text{gph}\partial\theta$ if and only if $\bar{z} - B\bar{v} \in N_Y(v)$. Furthermore, for all vectors v sufficiently close to $\bar{v}$, we have that $I(v) \subset I(\bar{v})$, and therefore $N_Y(v)$ is one of the cones $cone\{b_i,\, i \in J\}$, where $J \subset I(\bar{v})$. It follows from the argument above that for (z,v) close enough to $(\bar{z},\bar{v})$, the inclusion $(z,v) \in U(\bar{z},\bar{v}) \cap \text{gph}\partial\theta$ is equivalent to the system of inclusions $I(v) \in \mathfrak{K} = \{J_1, J_2, \ldots, J_l\}$ and $z - Bv \in N_Y(v)$. The definition of prenormal cone (2.1) yields

$$\widehat{N}\big((\bar{z},\bar{v}); \text{gph}\partial\theta\big) = \left\{ (u,w) \in \mathbb{R}^m \times \mathbb{R}^m \,\middle|\, \limsup_{(z,v)\xrightarrow{\text{gph}\partial\theta}(\bar{z},\bar{v})} \frac{\langle (u,w), (z,v) - (\bar{z} - \bar{v})\rangle}{\|z - \bar{z}\| + \|v - \bar{v}\|} \le 0 \right\}.$$

Denote by Y_{J_k} the set $\{x \in Y | I(x) = J_k\}$, $k = 1, 2, \ldots, l$. For any $(u,w) \in \widehat{N}\big((\bar{z},\bar{v}); \text{gph}\partial\theta\big)$ one has

$$\limsup_{(z,v)\xrightarrow{\text{gph}\partial\theta}(\bar{z},\bar{v})} \frac{\langle u, z - \bar{z}\rangle + \langle w, v - \bar{v}\rangle}{\|z - \bar{z}\| + \|v - \bar{v}\|} \le 0,$$

which is equivalent to

$$\limsup_{\substack{v \xrightarrow{Y_{J_1}\cup\cdots\cup Y_{J_l}} \bar{v} \\ z - Bv \to \bar{z} - B\bar{v} \\ z - Bv \in N_Y(v)}} \frac{\langle u, z - Bv - \bar{z} + B\bar{v}\rangle + \langle u, Bv - B\bar{v}\rangle + \langle w, v - \bar{v}\rangle}{\|z - Bv - \bar{z} + B\bar{v}\| + \|Bv - B\bar{v}\| + \|v - \bar{v}\|} \le 0.$$

Denoting $\bar{y} := \bar{z} - B\bar{v}$, $y := z - Bv$ and noticing that $Y_{J_k} \bigcap Y_{J_i} = \emptyset$, we deduce from the above inequality that

$$\limsup_{\substack{v \xrightarrow{Y_{J_k}} \bar{v} \\ y \to \bar{y} \\ y \in N_Y(v)}} \frac{\langle u, y - \bar{y}\rangle + \langle Bu + w, v - \bar{v}\rangle}{\|y - \bar{y}\| + \|Bv - B\bar{v}\| + \|v - \bar{v}\|} \le 0 \quad \text{for all } k \in \{1, 2, \ldots, l\}.$$

Let $N_Y^{J_k} := cone\{b_i, i \in J_k\}$. It is not hard to see that $N_Y(v) = N_Y^{J_k}$ for all $v \in Y_{J_k}$. Furthermore, denoting by Y'_{J_k} the closure of the set Y^{J_k}, the latter inequalities can be rewritten as

$$\limsup_{\substack{v \xrightarrow{Y_{J_k}} \bar{v} \\ y \xrightarrow{N_Y^{J_k}} \bar{y}}} \frac{\langle u, y - \bar{y}\rangle + \langle Bu + w, v - \bar{v}\rangle}{\|y - \bar{y}\| + \|Bv - B\bar{v}\| + \|v - \bar{v}\|} \leq 0 \quad \text{for all } k \in \{1, 2, \ldots, l\}$$

or

$$(u, Bu + w) \in \widehat{N}\big((\bar{y}, \bar{v}); N_Y^{J_k} \times Y'_{J_k}\big) \text{ for all } k \in \{1, 2, \ldots, l\}. \tag{3.3}$$

Applying the Proposition 6.41 in [37] and noticing that all the sets under consideration are convex, we have

$$\widehat{N}\big((\bar{y}, \bar{v}); N_Y^{J_k} \times Y'_{J_k}\big) = \widehat{N}\big(\bar{y}; N_Y^{J_k}\big) \times \widehat{N}\big(\bar{v}; Y'_{J_k}\big) = N\big(\bar{y}; N_Y^{J_k}\big) \times N\big(\bar{v}; Y'_{J_k}\big)$$

It is deduced then that the inclusion $(u, w) \in \widehat{N}((\bar{z}, \bar{v}); \mathrm{gph}\partial\theta)$ is equivalent to the following inclusion

$$(u, Bu + w) \in \bigcap_{k=1}^{l} \Big[N(\bar{y}; N_Y^{J_k}) \times N(\bar{v}; Y'_{J_k})\Big]. \tag{3.4}$$

Let us first find the normal cone $N(\bar{y}; N_Y^{J_k})$. Since all sets $N_Y^{J_k}$ are cones, one has

$$\begin{aligned} N(\bar{y}; N_Y^{J_k}) &= N(\bar{z} - B\bar{v}; N_Y^{J_k}) \\ &= (N_Y^{J_k})^* \cap \{\bar{z} - B\bar{v}\}^\perp \\ &= \{\langle b_i, x\rangle \leq 0, i \in J_k\} \cap \{\bar{z} - B\bar{v}\}^\perp \end{aligned} \tag{3.5}$$

Now to find $N(\bar{v}; Y'_{J_k})$, we first notice that $Y'_{J_k} = \mathrm{cl}Y_{J_k}$, so the set Y'_{J_k} is locally closed around $\bar{v}$. Since $Y_{J_k} = \{x \in Y | I(x) = J_k\} = \{x \in Y | \langle b_i, x\rangle = \alpha_i \text{ if } i \in J_k, \langle b_i, x\rangle < \alpha_i \text{ if } i \notin J_k\}$, it follows that $Y'_{J_k} = \mathrm{cl}\{x \in Y | I(x) = J_k\} = \{x \in Y | \langle b_i, x\rangle = \alpha_i \text{ if } i \in J_k, \langle b_i, x\rangle \leq \alpha_i \text{ if } i \notin J_k\}$ and thus

$$N(\bar{v}; Y'_{J_k}) = cone\{b_i, i \in I(\bar{v}) \setminus J_k\} + span\{b_i, i \in J_k\}. \tag{3.6}$$

Combining (3.5) and (3.6) yields the following calculations

$$\bigcap_{k=1}^{l}\Big[N(\bar{y};N_Y^{J_k})\times N(\bar{v};Y'_{J_k})\Big]$$

$$=\bigcap_{k=1}^{l}\Big[N(\bar{z}-B\bar{v};N_Y^{J_k})\Big]\times\Big[\bigcap_{k=1}^{l}N(\bar{v};Y'_{J_k})\Big]$$

$$=\Big[\bigcap_{k=1}^{l}\big\{\langle b_i,x\rangle\le 0, i\in J_k\big\}\cap\{\bar{z}-B\bar{v}\}^{\perp}\Big]\times\Big[\bigcap_{k=1}^{l}N(\bar{v};Y'_{J_k})\Big]$$

$$=\Big[\big\{\langle b_i,x\rangle\le 0, i\in I(\bar{v})\big\}\cap\big\{\bar{z}-B\bar{v}\big\}^{\perp}\Big]\times\bigcap_{k=1}^{l}N(\bar{v};Y'_{J_k})$$

$$=\Big[T_Y(\bar{v})\cap\{\bar{z}-B\bar{v}\}^{\perp}\Big]\times\bigcap_{k=1}^{l}\Big[cone\{b_i, i\in I(\bar{v})\setminus J_k\}+span\{b_i, i\in J_k\}\Big]. \tag{3.7}$$

Remembering that $T_Y(\bar{v})\cap\{\bar{z}-B\bar{v}\}^{\perp}$ is the critical cone to the set Y at $\bar{v}$ for $\bar{z}-B\bar{v}$, the formula (3.1) follows from (3.4) and (3.7) and the proposition is proved. □

The next proposition establishes a formula for the tangent cone to the graph of the subgradient mapping $\partial\varphi$ at a point $(\bar{z},\bar{v})$ in terms of the initial data, which are the set Y and the matrix B.

Proposition 3.2. The tangent cone $T\big((\bar{z},\bar{v});\text{gph}\partial\theta\big)$ is given by

$$T\big((\bar{z},\bar{v});\text{gph}\partial\theta\big)=\bigcup_{k=1}^{l}\Big\{(u,w)\Big|\ \ \langle b_i,w\rangle\le 0\,\forall\, i\in I(\bar{v})\backslash J_k,\ \langle b_i,w\rangle=0\,\forall\, i\in J_k,\ u-Bw\in cone\{b_i, i\in J_k\}+span\{\bar{z}-B\bar{v}\}\Big\}. \tag{3.8}$$

Proof. By the definition of the tangent cone (2.4), we have

$$T\big((\bar{z},\bar{v});\text{gph}\partial\theta\big)=\Big\{(u,w)\in\mathbb{R}^m\times\mathbb{R}^m\Big|\text{there are } (z_j,v_j)\xrightarrow{\text{gph}\partial\theta}(\bar{z},\bar{v}) \text{ such that } \alpha_j\ge 0 \text{ and } \alpha_j\big((z_j,v_j)-(\bar{z},\bar{v})\big)\to(u,w) \text{ as } j\to\infty\Big\}$$

It follows that a vector $(u,w)\in\mathbb{R}^m\times\mathbb{R}^m$ belongs to $T\big((\bar{z},\bar{v});\text{gph}\partial\theta\big)$ if and only if there exists a sequence $\{(z_j, v_j)\} \in \mathbb{R}^m \times \mathbb{R}^m$ that satisfies the system

$$\begin{cases} (z_j, v_{)} \xrightarrow{\text{gph}\partial\theta} (\bar{z}, \bar{v}) \\ \alpha_j(z_j - \bar{z}) \to u \text{ as } j \to \infty \\ \alpha_j(v_j - \bar{v}) \to w \text{ as } j \to \infty \end{cases}$$

Using the same notations as in the proof of the Propositon (3.1), it is not hard to see that the above system can be rewritten as

$$\begin{cases} v_j \xrightarrow{Y_{J_1} \bigcup \ldots \bigcup Y_{J_l}} \bar{v} \text{ as } j \to \infty \\ z_j - Bv_j \to \bar{z} - B\bar{v} \text{ as } j \to \infty \\ z_j - Bv_j \in N_Y(v_j) \text{ for all } j \in \mathbb{N} \\ \alpha_j(z_j - \bar{z}) \to u \text{ as } j \to \infty \\ \alpha_j(v_j - \bar{v}) \to w \text{ as } j \to \infty \end{cases}$$

which is equivalent to the following assertion:

$$\begin{cases} v_j \xrightarrow{Y_{J_k}} \bar{v} \\ z_j - Bv_j \xrightarrow{N_Y^{J_k}} \bar{z} - B\bar{v} \text{ as } j \to \infty \\ \alpha_j(v_j - \bar{v}) \to w \text{ as } j \to \infty \\ \alpha_j(z_j - \bar{z}) - \alpha_j B(v_j - \bar{v}) \to u - Bw \text{ as } j \to \infty \end{cases} \qquad \text{for some } k \in \{1, \ldots, l\}.$$

It is followed from the above that for some $k \in \{1, \ldots, l\}$ we have

$$\begin{cases} v_j \xrightarrow{Y_{J_k}} \bar{v} \\ z_j - Bv_j \xrightarrow{N_Y^{J_k}} \bar{z} - B\bar{v} \\ \alpha_j\big[(z_j - Bv_j, v_j) - (\bar{z} - B\bar{v}, \bar{v})\big] \to (u - Bw, w) \text{ as } j \to \infty \end{cases}$$

This implies that $(u, w) \in T\big((\bar{z}, \bar{v}); \text{gph}\partial\theta\big)$ if and only if

$$(u - Bw, w) \in \bigcup_{k=1}^{l} T\big((\bar{z} - B\bar{v}, \bar{v}); N_Y^{J_k} \times Y'_{J_k}\big). \tag{3.9}$$

Since for all $k \in \{1, 2, \ldots, l\}$ both sets $N_Y^{J_k}$ and Y'_{J_k} are convex, we may apply the formula in [[37], Proposition 6.41] and obtain that $(u, w) \in T\big((\bar{z}, \bar{v}); \text{gph}\partial\theta\big)$ is equivalent to the inclusion

$$(u - Bw, w) \in \bigcup_{k=1}^{l} \big[T(\bar{z} - B\bar{v}; N_Y^{J_k}) \times T(\bar{v}; Y'_{J_k})\big]. \tag{3.10}$$

Since $T(\bar{v}; Y'_{J_k}) = N(\bar{v}; Y'_{J_k})^*$, where $N(\bar{v}; Y'_{J_k}) = cone\{b_i, i \in I(\bar{v}) \setminus J_k\} + span\{b_i, i \in J_k\}$ by (3.6), one has

$$T(\bar{v}; Y'_{J_k}) = N(\bar{v}; Y'_{J_k})^* = \big\{y | \langle b_i, y\rangle \le 0 \text{ if } i \in I(\bar{v}) \backslash J_k, \langle b_i, y\rangle = 0 \text{ if } i \in J_k\big\}. \tag{3.11}$$

Furthermore, it is deduced from $N_Y^{J_k} = \text{cone}\{b_i, i \in J_k\}$ that

$$\begin{aligned} T(\bar{z} - B\bar{v}; N_Y^{J_k}) &= N(\bar{z} - B\bar{v}; N_Y^{J_k})^* \\ &= \big((N_Y^{J_k})^* \cap \{\bar{z} - B\bar{v}\}^{\perp}\big)^* \\ &= \big(\{y | \langle b_i, y\rangle \le 0, i \in J_k\} \cap \{\bar{z} - B\bar{v}\}^{\perp}\big)^* \\ &= \big(\{y | \langle b_i, y\rangle \le 0, i \in J_k,\ \langle \bar{z} - B\bar{v}, y\rangle = 0\}\big)^* \\ &= cone\{b_i, i \in J_k\} + span\{\bar{z} - B\bar{v}\} \end{aligned} \tag{3.12}$$

We can see from (3.10) that

$$T\big((\bar{z}, \bar{v}); \text{gph}\partial\theta\big) = \bigcup_{k=1}^{l} \Big\{(u, w) | w \in T(\bar{v}; Y'_{J_k}),\ u - Bw \in T(\bar{z} - B\bar{v}; N_Y^{J_k})\Big\},$$

which justifies (3.8) by taking into account (3.11) and (3.12). This completes the proof of the propositon. □

Now we prove the main result in this chapter, which is the formula for the normal cone to the graph of the subgradient mapping. We will use all notations in the proofs of the Propositions (3.1) and (3.8). Also, denote $T_Y^{J_k} := \{x \in \mathbb{R}^m | \langle b_i, x\rangle \le 0, i \in J_k\}$.

Theorem 3.3. *The normal cone* $N\big((\bar{z}, \bar{v}); \text{gph}\partial\theta\big)$ *is calculated by the formula*

$$N((\bar{z}, \bar{v}); \text{gph}\partial\theta) = \bigcup_{k=1}^{l} \Big\{(u, w) \in \mathbb{R}^m \times \mathbb{R}^m | (u, Bu + w) \in \big[T_Y^{J_k} \cap \{\bar{z} - B\bar{v}\}^{\perp}\big] \times span\{b_i, i \in J_k\}\Big\}. \tag{3.13}$$

Proof. We will use the definition

$$N((\bar{z}, \bar{v}); \text{gph}\partial\theta) = \operatorname*{Limsup}_{(z,v) \xrightarrow{\text{gph}\partial\theta} (\bar{z}, \bar{v})} \widehat{N}((z, v); \text{gph}\partial\theta)$$

to compute the normal cone to the graph of the subdifferential $\partial\theta$ at $(\bar{z}, \bar{v})$.

For each (z, v) denote by $\mathfrak{K}(z, v)$ the set $\big\{J \subset I(v) | z - Bv \in cone\{b_i, i \in J\}\big\}$. It follows from the formula (3.1) that

$$\widehat{N}((z, v); \text{gph}\partial\theta) = \big\{(u, w) \in \mathbb{R}^m \times \mathbb{R}^m | (u, Bu + w) \in K_Y(v, z - Bv) \times \mathcal{A}(z, v)\big\},$$

where $K_Y(v, z - Bv) = T_Y(v) \cap \{z - Bv\}^\perp$ and

$$\mathcal{A}(z, v) = \bigcap_{J \in \mathfrak{K}(z,v)} \Big(cone\{b_i, i \in I(v) \setminus J\} + span\{b_i, i \in J\}\Big).$$

Recall that in the proof of the Propositon (3.1) we deduce that $(z, v) \in \text{gph}\partial\theta \cap U(\bar{z}, \bar{v})$ if and only if $I(v) \in \mathfrak{K} = \{J_1, J_2, \ldots, J_l\}$ and $z - Bv \in N_Y(v)$. Hence, the normal cone to the graph of the subdifferential $\partial\theta$ at $(\bar{z}, \bar{v})$ can be rewritten as follows:

$$N((\bar{z}, \bar{v}); \text{gph}\partial\theta) = \bigcup_{k=1}^{l} \Big[\operatorname*{Limsup}_{\substack{(z,v) \to (\bar{z}, \bar{v}) \\ I(v) = J_k \\ z - Bv \in N_Y(v)}} \widehat{N}((z, v); \text{gph}\partial\theta) \Big] \tag{3.14}$$

or

$$N((\bar{z}, \bar{v}); \text{gph}\partial\theta) = \bigcup_{k=1}^{l} \Big[\operatorname*{Limsup}_{\substack{(z,v) \to (\bar{z}, \bar{v}) \\ v \in Y_{J_k} \\ z - Bv \in N_Y^{J_k}}} \widehat{N}((z, v); \text{gph}\partial\theta) \Big]. \tag{3.15}$$

Since $T_Y(v) = T_Y^{J_k}$ for all $v \in Y_{J_k}$, applying the formula (3.1) gives us

$$N((\bar{z}, \bar{v}); \text{gph}\partial\theta) = \bigcup_{k=1}^{l} \Bigg(\operatorname*{Limsup}_{\substack{(z,v) \to (\bar{z}, \bar{v}) \\ v \in Y_{J_k} \\ z - Bv \in N_Y^{J_k}}} \Big\{(u, w) | (u, Bu + w) \in \big[T_Y^{J_k} \cap \{z - Bv\}^\perp\big] \times \mathcal{A}(z, v)\Big\}\Bigg). \tag{3.16}$$

We now prove that

$$\operatorname*{Limsup}_{\substack{(z,v)\to(\bar z,\bar v)\\ v\in Y_{J_k}\\ z-Bv\in N_Y^{J_k}}} \Big\{(u,w)|(u,Bu+w)\in\big[T_Y^{J_k}\cap\{z-Bv\}^\perp\big]\times\mathcal{A}(z,v)\Big\}$$
$$=\Big\{(u,w)\,\big|\,(u,Bu+w)\in\big[T_Y^{J_k}\cap\{\bar z-B\bar v\}^\perp\big]\times span\{b_i,i\in J_k\}\Big\}. \tag{3.17}$$

The "$\subseteq$" inclusion is obvious since for any (z,v) with $I(v)=J_k$ and $z-Bv\in N_Y^{J_k}$ the set $\mathcal{A}(z,v)\subset span\{b_i,i\in J_k\}$. To prove the "$\supseteq$" inclusion, pick any (u_0,w_0) from the right hand side of (3.17). For any set $J_k\in\mathfrak{K}$ we can find a set of vectors $\{b_1,b_2,\ldots,b_q\}\subset\{b_i,i\in J_k\}$ such that $cone\{b_1,b_2,\ldots,b_q\}=cone\{b_i,i\in J_k\}=N_Y^{J_k}$ and $cone\{b_i,i\in I\}\subsetneqq cone\{b_1,b_2,\ldots,b_q\}$ for any subset $I\subsetneqq\{1,2,\ldots,q\}$. We may choose vectors z_j and v_j, $j\in\mathbb{N}$, so that $v_j\in Y_{J_k}$, $z_j-Bv_j\in cone\{b_1,b_2,\ldots,b_q\}$ and $z_j-Bv_j\notin cone\{b_i,i\in I\}$ for any $I\subsetneqq\{1,2,\ldots,q\}$ and in addition $z_j-Bv_j\to\bar z-B\bar v$ as $j\to\infty$. It follows from the choice of z_j and v_j that for any $J\in\mathfrak{K}(z_j,v_j)$ one has

$$cone\{b_i,i\in J\}=cone\{b_1,b_2,\ldots,b_q\}=cone\{b_i,i\in J_k\},$$

and therefore

$$span\{b_i,i\in J\}=span\{b_1,b_2,\ldots,b_q\}=span\{b_i,i\in J_k\}.$$

Thus

$$\mathcal{A}(z_j,v_j)=\bigcap_{J\in\mathfrak{K}(z_j,v_j)}\big(cone\{b_i,i\in J_k\backslash J\}+span\{b_i,i\in J\}\big)=span\{b_i,i\in J_k\}$$

The latter equation yields

$$\Big\{(u,w)\,\big|\,(u,Bu+w)\in\big[T_Y^{J_k}\cap\{z_j-Bv_j\}^\perp\big]\times\mathcal{A}(z_j,v_j)\Big\}$$
$$=\Big\{(u,w)\,\big|\,(u,Bu+w)\in\big[T_Y^{J_k}\cap\{z_j-Bv_j\}^\perp\big]\times span\{b_i,i\in J_k\}\Big\}. \tag{3.18}$$

Since $(u_0,Bu_0+w_0)\in\big[T_Y^{J_k}\cap\{\bar z-B\bar v\}^\perp\big]\times span\{b_i,i\in J_k\}$ and $(z_j,v_j)\to(\bar z,\bar v)$ as $j\to\infty$, it is easy to find (u_j,w_j) with $u_j\in T_Y^{J_k}\cap\{z_j-Bv_j\}^\perp$, $Bu_j+w_j\in span\{b_i,i\in J_k\}$ and $(u_j,w_j)\to(u_0,w_0)$ as $j\to\infty$. By the definition of Limsup and the result in (3.18), we deduce that (u_0,w_0) belongs to the left hand side of (3.17) and thus this formula is proved.

Combining (3.17) with (3.16) justifies the formula (3.13) and thus completes the proof of the theorem. □

It is not hard to see that the result in the Theorem 3.3 together with the definition of the second-order subdifferential gives us the following.

Theorem 3.4. *The second-order subdifferential $\partial^2\theta(\bar z,\bar v)(w)$ is calculated by*

$$\partial^2\theta(\bar z,\bar v)(w)=\bigcup_{k=1}^{l}\Big\{u\in\mathbb{R}^m|(u,Bu-w)\in\big[T_Y^{J_k}\cap\{\bar z-B\bar v\}^{\perp}\big]\times span\{b_i,i\in J_k\}\Big\}. \quad (3.19)$$

To end this chapter, we give two examples to illustrate the application of the formula (3.19) in calculating the second-order subdifferential of functions θ of the form (1.1). All results obtained are expressed in terms of the initial data, i.e. the set Y, the matrix B and the given point $(\bar z,\bar v)$.

Example 3.5. Consider the function θ from (1.1) where $B:=I$ is the 2×2 identity matrix and the set Y is the nonnegative orthant in $\mathbb{R}^2$, i.e.

$$Y=\mathbb{R}^2_+:=\big\{y=(y_1,y_2)\in\mathbb{R}^2\,\big|\,y_1\geq 0,y_2\geq 0\big\}.$$

It is easy to see that Y can be written in the form $\{y\in\mathbb{R}^2\,|\langle b_i,y\rangle\leq 0,i=1,2\}$, where $b_1:=(0,-1)$ and $b_2:=(-1,0)$. The function θ now is as follows:

$$\theta=\theta_{\mathbb{R}^2_+,I}(u)=\sup_{y\in\mathbb{R}^2_+}\Big\{\langle y,u\rangle-\frac{1}{2}\langle y,y\rangle\Big\},\quad u\in\mathbb{R}^2.$$

Let $\bar z:=0\in\mathbb{R}^2$. By Theorem 2.1(ii) we have that $\bar v\in\partial\theta(\bar z)$ if and only if $\bar z-B\bar v\in N_{\mathbb{R}^2_+}(\bar v)=\mathbb{R}^2_-\cap\bar v^{\perp}$, which yields that $\bar v=0$. It follows that the set of active constraints $I(\bar v)=\{1,2\}$ and $\mathcal{K}=\{J_1,J_2,J_3\}$ with $J_1:=\{1\}$, $J_2=\{2\}$ and $J_3=I(\bar v)=\{1,2\}$. It is not hard to check the following:

$$T_Y^{J_1}=\{x\in\mathbb{R}^2\,|\,\langle b_1,x\rangle\leq 0\}=\mathbb{R}\times\mathbb{R}_+,\qquad \text{span}\{b_1\}=\{0\}\times\mathbb{R},$$

$$T_Y^{J_2}=\{x\in\mathbb{R}^2\,|\,\langle b_2,x\rangle\leq 0\}=\mathbb{R}_+\times\mathbb{R},\qquad \text{span}\{b_1\}=\mathbb{R}\times\{0\},$$

$$T_Y^{J_3}=\{x\in\mathbb{R}^2\,|\,\langle b_i,x\rangle\leq 0,i=1,2\}=\mathbb{R}\times\mathbb{R}_+,\qquad \text{span}\{b_1,b_2\}=Y\times\mathbb{R}^2.$$

Applying the formula (3.19) gives

$\partial^2\theta(0,0)(w) = \{u \in \mathbb{R}^2 \,|\, (u, u-w) \in (\mathbb{R}\times\mathbb{R}_+)\times(\{0\}\times\mathbb{R}) \cup (\mathbb{R}_+\times\mathbb{R})\times(\mathbb{R}\times\{0\}) \cup Y\times\mathbb{R}^2\}$,

which in the case $w = 0$ reduces to the set Y. Therefore, we have $\partial^2\theta(0,0)(0) = Y$.

The next example considers a simple case where the matrix B is singular.

Example 3.6. Consider the initial data of (1.1) as follows:

$$Y := \mathbb{R}^2_+, \quad B := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

By Theorem 2.1(i), we obtain $\operatorname{dom}\theta_{Y,B} = \mathbb{R}\times\mathbb{R}_-$. Since $\partial\theta_{Y,B}(u) = (N_Y + B)^{-1}(u)$ by Theorem 2.1(iii), it is not hard to see $\partial\theta = \partial\theta_{Y,B}(0) = \{0\}\times\mathbb{R}_+$. Let $\bar{v} := (0,1) \in \partial\theta(0)$. It is clear that $I(\bar{v}) = \{2\}$, therefore $\mathcal{K} = \{J_1\}$, where $J_1 = I(\bar{v}) = \{2\}$. It then yields that the set $T_Y^{J_1} = \{x \,|\, \langle b_2, x\rangle \le 0\}$ and hence

$$\partial^2\theta(\bar{z},\bar{v})(w) = \partial^2\theta(0,\bar{v})(w) = \{u \in \mathbb{R}^2 \,|\, (u, Bu - w) \in (\mathbb{R}_+\times\mathbb{R})\times(\mathbb{R}\times\{0\})\}.$$

By the structure of the matrix B we see that for any $w = (w_1, w_2) \in \mathbb{R}^2$, the second-order subdifferential $\partial^2\theta(\bar{z},\bar{v})(w) \neq \emptyset$ only if $w_2 = 0$ and in this case $\partial^2\theta(\bar{z},\bar{v})(w) = \mathbb{R}_+\times\mathbb{R}$.

3.2 Second-order Subdifferential and Full Stability in Constrained Optimization

This section concerns the two-parametric unconstrained optimization problems studied in [[15]]. Namely, for a proper extended-real-valued function $\varphi : \mathbb{R}^m \times \mathbb{R}^d \to \bar{\mathbb{R}}$ consider the following minimization problem

$$\text{minimize } \varphi(x,w) - \langle p, x\rangle \text{ over } x \in \mathbb{R}^m. \tag{3.20}$$

Label the above problem as $P(w,p)$. Recall that the point $x \in \mathbb{R}^m$ is a feasible solution to $P(w,p)$ if the value $\varphi(x,w)$ is finite. Given $(\bar{w},\bar{p})$, let $\bar{x}$ be a feasible solution to the problem $P(\bar{w},\bar{p})$ and let $\gamma > 0$. Define the functions

$$m_\gamma(w,p) := \inf_{\|x-\bar{x}\|\le\gamma} \{\varphi(x,w) - \langle p, x\rangle\}, \quad (w,p) \in \mathbb{R}^d \times \mathbb{R}^m \tag{3.21}$$

and

$$M_\gamma(w,p) := \mathop{\arg\min}_{\|x-\bar{x}\|\le\gamma} \big\{\varphi(x,w) - \langle p,x\rangle\big\}, \quad (w,p)\in\mathbb{R}^d\times\mathbb{R}^m. \tag{3.22}$$

We say that $\bar{x}$ is a locally optimal soltion to $P(\bar{w},\bar{p})$ if $\bar{x}\in M_\gamma(\bar{w},\bar{p})$ for some small $\gamma > 0$. Furthermore, we say $\bar{x}$ is a *fully stable* locally optimal solution to problem $P(\bar{w},\bar{p})$ if there is a number $\gamma > 0$ and neighborhoods W of $\bar{w}$ and P of $\bar{p}$ such that the mapping $(w,p)\mapsto M_\gamma(w,p)$ is single-valued and Lipschitz continuous with $M_\gamma(\bar{w},\bar{p}) = \{\bar{x}\}$ and the function $(w,p)\mapsto m_\gamma(w,p)$ is likewise Lipschitz continuous on $W\times P$. This section of the dissertation deals with the case the function $\varphi(x,w)$ has the form

$$\varphi(x,w) := \varphi_0(x,w) + \theta(\Phi(x,w)), \tag{3.23}$$

where θ is of the form (1.1) and $\Phi(x,w) := (\varphi_1(x,w),\ldots,\varphi_m(x,w))$. Assume that all the functions $\varphi_0,\varphi_1,\ldots,\varphi_m$ are twice continuously differentiable around the point $(\bar{x},\bar{w})$. We also assume the LICQ condition at $(\bar{x},\bar{w})$, i.e. the vectors $\nabla_x\varphi_1(\bar{x},\bar{w}),\ldots,\nabla_x\varphi_m(\bar{x},\bar{w})$ are linearly independent. Let $\bar{p}$ satisfy the stationarity condition

$$\bar{p}\in\nabla_x\varphi_0(\bar{x},\bar{w}) + \nabla_x\Phi(\bar{x},\bar{w})^*\partial\theta(\Phi(\bar{x},\bar{w})). \tag{3.24}$$

Define the extended Lagrangian function for the problem $P(w,p)$ when φ has the form (3.23) as follows:

$$L(x,w,v) := \varphi_0(x,w) + \Phi(x,w)^*v - \frac{1}{2}\langle v,Bv\rangle \text{ with } v\in\mathbb{R}^m. \tag{3.25}$$

The full stability of the problem $P(w,p)$ defined in (3.20) has been studied in detail in the paper [[31]]. The Theorem 7.3(ii) in this paper provides a sufficient condition for the full stability of $\bar{x}$ under the condition that the second-order derivative of the function ϑ is 0. This result can not be applied to our case when the function φ is of the form (3.23) because of the structure (1.1) of the function θ. In what follows we deduce a necessary and sufficient condition for the full stability of $\bar{x}$ in the problem under consideration. Following [[31], Theorem 7.3], determine the unique vector $\bar{v}\in\mathbb{R}^m$ from the equation

$$\nabla_x\Phi(\bar{x},\bar{w})^*\bar{v} = \bar{p} - \nabla_x\varphi_0(\bar{x},\bar{w}). \tag{3.26}$$

Denote $\bar{z} := \Phi(\bar{x}, \bar{w})$. It follows from the proof of this theorem that $\bar{x}$ is a fully stable locally optimal solution to $P(\bar{w}, \bar{p})$ if and only if the following holds

$$\langle \xi, \nabla^2_{xx} L(\bar{x}, \bar{w}, \bar{v})\xi\rangle + \langle \eta, \nabla_x \Phi(\bar{x}, \bar{w})\xi\rangle > 0 \text{ if } \eta \in \partial^2\theta(\bar{z}, \bar{v})(\nabla_x \Phi(\bar{x}, \bar{w})\xi),\ \xi \neq 0. \tag{3.27}$$

Employing the formula for the second-order subdifferential in (3.19) the above condition can be restated as follows:

$$\langle \xi, \nabla^2_{xx} L(\bar{x}, \bar{w}, \bar{v})\xi\rangle + \langle \eta, \nabla_x \Phi(\bar{x}, \bar{w})\xi\rangle > 0 \tag{3.28}$$

for all (η, ξ) satisfying the following:

(i) $\xi \neq 0$ and $\langle \eta, \bar{z} - B\bar{v}\rangle = 0$,

(ii) there exists $J \subset I(\bar{v})$ such that $\bar{z} - B\bar{v} \in \operatorname{cone}\{b_i, i \in J\}$, $\langle b_i, \eta\rangle \leq 0$ for all $i \in J$ and $B\eta - \nabla\Phi(\bar{x}, \bar{w})\xi \in \operatorname{span}\{b_i, i \in J\}$.

The condition (3.28) obtained above is the necessary and sufficient condition for the full stability of $\bar{x}$. It is formulated in a form that can be checked using the initial data of the problem, which makes the condition more applicable in practical.

CHAPTER 4 MULTIPLIER CRITICALITY IN PIECEWISE LINEAR-QUADRATIC SETTINGS

4.1 Equivalent Description of Criticality

In this section we formulate the definitions of critical and noncritical multipliers corresponding to stationary points of the variational system (1.3) with the piecewise linear-quadratic term (1.1), establish an equivalent description of criticality entirely via the given data of (1.3), and then present two examples illustrating the calculation of critical and noncritical multipliers for this setting.

Given a point $\bar{x} \in \mathbb{R}^n$, define the set of *Lagrange multipliers* associated with $\bar{x}$ by

$$\Lambda(\bar{x}) := \big\{\lambda \in \mathbb{R}^m \big|\ \Psi(\bar{x}, \lambda) = 0,\ \lambda \in \partial\theta_{Y,B}\big(\Phi(\bar{x})\big)\big\}. \tag{4.1}$$

If $(\bar{x}, \bar{\lambda})$ is a solution to the variational system (1.3), we clearly get $\bar{\lambda} \in \Lambda(\bar{x})$. Furthermore, it is not hard to check that the inclusion $\bar{\lambda} \in \Lambda(\bar{x})$ ensures that $\bar{x}$ is a *stationary point* of (1.3) in the sense that it satisfies the condition

$$0 \in f(\bar{x}) + \partial\big(\theta_{Y,B} \circ \Phi\big)(\bar{x}). \tag{4.2}$$

Suppose from now on that $\Lambda(\bar{x}) \neq \emptyset$, which is ensured, e.g., by any constraint qualification condition in problems of constrained optimization. The following definitions of critical and noncritical multipliers for (1.3), are just specifications of those from [32], given there for general variational systems with the subsequent implementation for the case of a convex piecewise linear function θ. It is worth noticing that the function θ from (1.1) with $B = 0$ is convex piecewise linear, namely its epigraph is a convex polyhedral set, and so can be covered by the results already established in [32]; however, when $B \neq 0$, it is a convex piecewise linear-quadratic function and requires different techniques to achieve similar results.

Definition 4.1. (critical and noncritical multiplies in variational systems). Let $(\bar{x}, \bar{\lambda})$ be a solution to the variational system (1.3). We say that $\bar{\lambda} \in \Lambda(\bar{x})$ is a CRITICAL LAGRANGE

MULTIPLIER for (1.3) corresponding to $\bar{x}$ if there exists a nonzero vector $\xi \in \mathbb{R}^n$ such that

$$0 \in \nabla_x \Psi(\bar{x}, \bar{\lambda})\xi + \nabla\Phi(\bar{x})^* D\partial\theta_{Y,B}\big(\Phi(\bar{x}), \bar{\lambda}\big)\big(\nabla\Phi(\bar{x})\xi\big). \tag{4.3}$$

A given multiplier $\bar{\lambda} \in \Lambda(\bar{x})$ is NONCRITICAL for (1.3) corresponding to $\bar{x}$ if the generalized equation (4.3) admits only the trivial solution $\xi = 0$.

Applying the representations of Theorem 2.1 for the graphical derivative in (4.3) gives us an equivalent description of critical and noncritical multipliers from Definition 4.1, expressed entirely in terms of the initial data of (1.3).

Theorem 4.2. (equivalent description of criticality via piecewise linear-quadratic penalties). *Let $(\bar{x}, \bar{\lambda})$ be a solution to the variational system* (1.3) *with the term $\theta_{Y,B}$ taken from* (1.1). *Denoting $\bar{z} := \Phi(\bar{x})$ and $\mathcal{K} := K_Y(\bar{\lambda}, \bar{z} - B\bar{\lambda})$ via the critical cone* (2.5)*, we have that the multiplier $\bar{\lambda}$ corresponding to $\bar{x}$ is critical for* (1.3) *if and only if the system*

$$\begin{cases} \nabla_x \Psi(\bar{x}, \bar{\lambda})\xi + \nabla\Phi(\bar{x})^*\eta = 0, \quad \langle \nabla\Phi(\bar{x})\xi - B\eta, \eta\rangle = 0, \\ \nabla\Phi(\bar{x})\xi - B\eta \in \mathcal{K}^*, \ \ \text{and} \ \ \eta \in \mathcal{K} \end{cases} \tag{4.4}$$

admits a solution $(\xi, \eta) \in \mathbb{R}^n \times \mathbb{R}^m$ with $\xi \neq 0$. Accordingly, $\bar{\lambda}$ is a noncritical multiplier in this setting if and only if we have $\xi = 0$ for any solution (ξ, η) to (4.4).

Proof. To achieve the claimed equivalencies, we require to calculate the graphical derivative $D\partial\theta_{Y,B}$ in (4.3) for the function $\theta_{Y,B}$ given in (1.1). First we use formula (2.15) from Theorem 2.1(iii), which yields

$$D\partial\theta_{Y,B}(\bar{z}, \bar{\lambda})\big(\nabla\Phi(\bar{x})\xi\big) = \partial\theta_{\mathcal{K},B}\big(\nabla\Phi(\bar{x})\xi\big).$$

On the other hand, the second expression of $\partial\theta_{\mathcal{K},B}$ in (2.13) of Theorem 2.1(ii) shows that

$$\partial\theta_{\mathcal{K},B}\big(\nabla\Phi(\bar{x})\xi\big) = \big(N_{\mathcal{K}} + B\big)^{-1}\big(\nabla\Phi(\bar{x})\xi\big).$$

Putting these representations together, we arrive at

$$D\partial\theta_{Y,B}(\bar{z}, \bar{\lambda})\big(\nabla\Phi(\bar{x})\xi\big) = \big(N_{\mathcal{K}} + B\big)^{-1}\big(\nabla\Phi(\bar{x})\xi\big). \tag{4.5}$$

Picking further any vector η from the set on the left-hand side of (4.5) gives us therefore that $\eta \in \big(N_{\mathcal{K}} + B\big)^{-1}\big(\nabla\Phi(\bar{x})\xi\big)$ and so $\nabla\Phi(\bar{x})\xi - B\eta \in N_{\mathcal{K}}(\eta)$. Since $\mathcal{K}$ is a convex cone, the latter inclusion is equivalent to the conditions

$$\langle \nabla\Phi(\bar{x})\xi - B\eta, \eta\rangle = 0, \quad \nabla\Phi(\bar{x})\xi - B\eta \in \mathcal{K}^*, \quad \eta \in \mathcal{K}.$$

Finally, we substitute the obtained descriptions of $\eta \in D\partial\theta_{Y,B}(\bar{z}, \bar{\lambda})(\nabla\Phi(\bar{x})\xi)$ into (4.3) and thus clearly verify both assertions of the theorem. □

Next we present two examples, which demonstrate how to use the descriptions of Theorem 4.2 to explicitly determine critical and noncritical multipliers and illustrate in this way some characteristic features of multiplier criticality.

Example 4.3. (calculating critical and noncritical multipliers). Consider the multidimensional case of (1.3) with $\theta_{Y,B}$ from (1.1), where $B = I_m =: I$ is the $m \times m$ identity matrix, and where the convex polyhedral set Y is the nonnegative orthant in $\mathbb{R}^m$, i.e.,

$$Y = \mathbb{R}^m_+ := \big\{y = (y_1, \ldots, y_m) \in \mathbb{R}^m \big|\ y_i \geq 0 \ \text{ for all } \ i = 1, \ldots, m\big\}.$$

Thus the function $\theta_{Y,B}$ from (1.1) reduces in this case to

$$\theta_{\mathbb{R}^m_+, I}(u) = \sup_{y \in \mathbb{R}^m_+} \Big\{\langle y, u\rangle - \frac{1}{2}\langle y, y\rangle\Big\}, \quad u \in \mathbb{R}^m.$$

For any $\bar{x} \in \mathbb{R}^n$ and $\bar{z} := \Phi(\bar{x})$, by Theorem 2.1(ii) we have that $\lambda \in \partial\theta_{\mathbb{R}^m_+, I}(\bar{z})$ if and only if $\bar{z} - B\lambda \in N_{\mathbb{R}^m_+}(\lambda) = \mathbb{R}^m_- \cap \lambda^\perp$. Denoting $\bar{z} - \lambda$ by $\widehat{\lambda}$, the latter inclusion is equivalent to the following system of equations and inclusions:

$$\begin{cases} \lambda + \widehat{\lambda} = \bar{z} \\ \langle \lambda, \widehat{\lambda}\rangle = 0 \\ \lambda \in \mathbb{R}^m_+ \\ \widehat{\lambda} \in \mathbb{R}^m_- \end{cases} \tag{4.6}$$

It is not hard to see that for each fixed $\bar{x}$ and $\bar{z} = \Phi(\bar{x})$ this system has only one solution, which implies that the set of Lagrange multipliers has at most one element.

We now give two specific examples of mappings f and Φ, where one has a noncritical multiplier and the other has a critical multiplier. First, let $f(x) := x$ and $\Phi(x) := (x_1, 0, \ldots, 0) \in \mathbb{R}^m$ for all $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, and let $\bar{x} := 0 \in \mathbb{R}^n$. Combining (4.6) with the fact that $\Psi(\bar{x}, \lambda) = (\lambda_1, 0, \ldots, 0) \in \mathbb{R}^n$ implies that the unique Lagrange multiplier is $\bar{\lambda} = 0$. Then we calculate the critical cone $\mathcal{K} = K_Y(0, \bar{z})$ in Theorem 4.2 with $\bar{z} = \Phi(\bar{x}) = 0$ and its dual cone $\mathcal{K}^*$ by, respectively,

$$\mathcal{K} = T_{\mathbb{R}^m_+}(0) \cap \{\bar{z}\}^\perp = \mathbb{R}^m_+ \ \text{ and } \ \mathcal{K}^* = \text{span}\{\bar{z}\} + N_{\mathbb{R}^m_+}(0) = \mathbb{R}^m_-.$$

It follows from Theorem 4.2 that the unique Lagrange multiplier $\bar{\lambda} = 0$ is noncritical if and only if the system of equations and inclusions

$$\begin{cases} \nabla_x \Psi(\bar{x}, \bar{\lambda})\xi + \nabla\Phi(\bar{x})^*\eta = 0 \\ \langle \nabla\Phi(\bar{x})\xi - \eta, \eta \rangle = 0 \\ \nabla\Phi(\bar{x})\xi - \eta \in \mathbb{R}^m_- \\ \eta \in \mathbb{R}^m_+ \end{cases}$$

admits the only solution pairs $(\xi, \eta) \in \mathbb{R}^n \times \mathbb{R}^m$ with $\xi = 0$. Denoting $\zeta := \nabla\Phi(\bar{x})\xi - \eta$, the above system can be equivalently rewritten as

$$\begin{cases} \nabla_x \Psi(\bar{x}, \bar{\lambda})\xi + \nabla\Phi(\bar{x})^*\eta = 0 \\ \nabla\Phi(\bar{x})\xi - \eta - \zeta = 0 \\ \langle \zeta, \eta \rangle = 0 \\ \zeta \in \mathbb{R}^m_- \\ \eta \in \mathbb{R}^m_+. \end{cases} \tag{4.7}$$

Since $\nabla_x \Psi(\bar{x}, \bar{\lambda})\xi = \xi$, $\nabla\Phi(\bar{x})\xi = (\xi_1, 0, \ldots, 0) \in \mathbb{R}^m$, and $\nabla\Phi(\bar{x})^*\eta = (\eta_1, 0, \ldots, 0) \in \mathbb{R}^n$ for any $\eta = (\eta_1, \ldots, \eta_m) \in \mathbb{R}^m$, it can be easily checked that the latter system has the unique solution pair $(\xi, \eta) = (0, 0)$. This tells us that $\bar{\lambda} = 0$ is a noncritical multiplier.

Next we consider the case where $\Phi(x) := (x_1, 0, \ldots, 0) \in \mathbb{R}^m$ as before while $f(x) := (x_1, \ldots, x_{n-1}, 0) \in \mathbb{R}^n$ for all $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$. Proceeding similarly to the previous case shows that $\bar{\lambda} = 0$ is the unique Lagrange multiplier with the same critical cone $\mathcal{K}$. In this setting we have $\nabla_x \Psi(\bar{x}, \bar{\lambda})\xi = (\xi_1, \ldots, \xi_{n-1}, 0) \in \mathbb{R}^n$, and therefore system (4.7) reduces to

$$\begin{cases} (\xi_1, \ldots, \xi_{n-1}, 0) + (n_1, 0, \ldots, 0) = 0 \\ \nabla\Phi(\bar{x})\xi - \eta - \zeta = 0 \\ \langle \zeta, \eta \rangle = 0 \\ \zeta \in \mathbb{R}^m_- \\ \eta \in \mathbb{R}^m_+. \end{cases}$$

It shows that all the pairs (ξ, η) with $\eta = 0$ and $\xi = (0, \ldots, 0, \xi_n)$ for $\xi_n \in \mathbb{R}$ are solutions to the above system. Thus the multiplier $\bar{\lambda} = 0$ is critical. In Section 4.4 we revisit this example in the optimization framework; see Example 4.9.

The next two-dimensional example presents a simple linear-quadratic variational system of type (1.3) with $\theta_{Y,B}$ from (1.1) such that a stationary point therein is associated with both critical and noncritical Lagrange multipliers.

Example 4.4. (variational systems with both critical and noncritical multipliers corresponding to a given stationary point). Specify the data of (1.1) and (1.3) as follows:

$$Y := \mathbb{R}^2_+, \quad B := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad f(x) := -x, \text{ and } \Phi(x) := (0, x^2) \text{ for } x \in \mathbb{R}. \tag{4.8}$$

Thus we have in (1.3) that $\Psi(x, \lambda) = f(x) + \nabla\Phi(x)^*\lambda = -x + 2x\lambda_2$ for any $x \in \mathbb{R}$ and $\lambda = (\lambda_1, \lambda_2) \in \mathbb{R}^2$. It follows from the Example 3.6 that $\partial\theta_{Y,B}(0) = \{0\} \times \mathbb{R}_+$, and so $\Lambda(\bar{x}) = \{0\} \times \mathbb{R}_+$ with $\bar{x} := 0$. Then for any $\lambda = (\lambda_1, \lambda_2) \in \Lambda(\bar{x})$ we get $\lambda_1 = 0$ and $\lambda_2 \geq 0$. On the other hand, conditions (4.1) from Theorem 4.2 read now as

$$(2\lambda_2 - 1)\xi = 0, \quad \langle -B\eta, \eta \rangle = 0, \quad -B\eta \in \mathcal{K}^*, \quad \eta \in \mathcal{K}.$$

This tells us that if $\lambda_2 \neq \frac{1}{2}$, the latter system admits only the solution $\xi = 0$, and thus the obtained Lagrange multiplier λ is noncritical. In the case where $\lambda_2 = \frac{1}{2}$, this system admits nontrivial solutions ξ, and so the Lagrange multiplier $\lambda = (0, \frac{1}{2})$ is critical.

4.2 Uniqueness of Lagrange Multipliers and Isolated Calmness

This section is devoted to the study of uniqueness of Lagrange multipliers corresponding to given stationary points of the variational systems (1.3) with piecewise linear-quadratic penalties (1.1). This issue is definitely of its own interest while seems to be independent of multiplier criticality. However, the methods we develop for the uniqueness study and the obtained conditions for it occur to be closely related to the subsequent characterizations of noncritical multiplies as well as their deeper understanding and specification.

First we recall some “at-point" (vs. “around/neighborhood") stability properties of set-valued mappings that have been recognized in variational analysis; see, e.g., [34, 22, 37] with the references and commentaries therein.

It is said that a mapping $F\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is *calm* at $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ if there exist a constant $\ell \geq 0$ and neighborhoods U of $\bar{x}$ and V of $\bar{y}$ such that

$$F(x) \cap V \subset F(\bar{x}) + \ell\|x - \bar{x}\|\mathbb{B} \;\text{ for all }\; x \in U, \tag{4.9}$$

where $\mathbb{B}$ stands for the closed unit ball of the space in question. If (4.9) is replaced by

$$F(x) \cap V \subset \big\{\bar{y}\big\} + \ell\,\|x - \bar{x}\|\,\mathbb{B} \;\text{ for all }\; x \in U, \tag{4.10}$$

then the corresponding property is known as *isolated calmness* of F at $(\bar{x}, \bar{y})$. If the $\operatorname{gph} F$ is locally closed at $(\bar{x}, \bar{y})$, the latter property admits the graphical derivative characterization

$$DF(\bar{x}, \bar{y})(0) = \{0\} \tag{4.11}$$

known as the *Levy-Rockafellar criterion*; see the commentaries to [34, Theorem 4E.1].

Finally, F enjoys the *robust isolated calmness* property at $(\bar{x}, \bar{y})$ if in addition to (4.10) we have $F(x) \cap V \neq \emptyset$. This name is coined quite recently [2], while the property itself has been actually used in optimization over the years; see the discussions in [2, 32].

In this section we employ the calmness and isolated calmness properties for characterizations of uniqueness of Lagrange multipliers in (1.3) with the piecewise linear-quadratic term (1.1). Robust isolated calmness is used in the last section of this dissertation.

Using the data of (1.3), consider the set-valued mapping $G\colon \mathbb{R}^n \times \mathbb{R}^m \rightrightarrows \mathbb{R}^n \times \mathbb{R}^m$ given by

$$G(x,\lambda) := \begin{pmatrix} \Psi(x,\lambda) \\ -\Phi(x) \end{pmatrix} + \begin{pmatrix} 0 \\ (\partial\theta_{Y,B})^{-1}(\lambda) \end{pmatrix} \quad \text{for all } (x,\lambda) \in \mathbb{R}^n \times \mathbb{R}^m. \tag{4.12}$$

Then fix a point $\bar{x} \in \mathbb{R}^n$ and define the parameterized *multiplier map* $M_{\bar{x}}\colon \mathbb{R}^n \times \mathbb{R}^m \rightrightarrows \mathbb{R}^m$ associated with $\bar{x}$ by

$$M_{\bar{x}}(p_1,p_2) := \big\{\lambda \in \mathbb{R}^m \big|\ (p_2,p_2) \in G(\bar{x},\lambda)\big\}, \quad (p_1,p_2) \in \mathbb{R}^n \times \mathbb{R}^m. \tag{4.13}$$

We have $M_{\bar{x}}(0,0) = \Lambda(\bar{x})$ for the Lagrange multiplier set (4.1) of the unperturbed system (1.3).

The next theorem characterizes uniqueness of Lagrange multipliers in variational systems (1.3) with the term $\theta_{Y,B}$ from (1.1) via both calmness and isolated calmness properties of the multiplier map (4.13), which are equivalent to each other in this case and are characterized in turn by a novel *dual qualification condition*.

Theorem 4.5. (characterizations of uniqueness of Lagrange multipliers in variational systems). *Let $(\bar{x},\bar{\lambda})$ be a solution to the variational system* (1.3) *with $\theta_{Y,B}$ taken from* (1.1). *Then the following properties are equivalent:*

(i) $\Lambda(\bar{x}) = \{\bar{\lambda}\}$.

(ii) *$M_{\bar{x}}$ is calm at $\big((0,0),\bar{\lambda}\big)$ and $\Lambda(\bar{x}) = \{\bar{\lambda}\}$.*

(iii) *$M_{\bar{x}}$ is isolatedly calm at $\big((0,0),\bar{\lambda}\big)$.*

(iv) *We have the dual qualification condition*

$$D\partial\theta_{Y,B}(\bar{z},\bar{\lambda})(0) \cap \ker \nabla\Phi(\bar{x})^* = \{0\}, \tag{4.14}$$

where $D\partial\theta_{Y,B}(\bar{z},\bar{\lambda})$ is calculated by (4.5).

Proof. Denoting $\bar{z} := \Phi(\bar{x})$ as above, we begin with proving the equivalence (iii)$\Longleftrightarrow$(iv). To proceed, observe that the graph of $M_{\bar{x}}$ is closed and deduce from (4.11) that $M_{\bar{x}}$ is isolatedly calm at $((0,0),\bar{\lambda})$ if and only if $DM_{\bar{x}}\big((0,0),\bar{\lambda}\big)(0,0)=\{0\}$. It is not hard to check that $\eta\in DM_{\bar{x}}\big((0,0),\bar{\lambda}\big)(0,0)$ amounts to saying that η is a solution to the system

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} \in \begin{bmatrix} \nabla\Phi(\bar{x})^*\eta \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ D(\theta_{Y,B})^{-1}(\bar{\lambda},\bar{z})(\eta) \end{bmatrix}.$$

This tells us that η is a solution to the above system if and only if

$$\eta\in D\partial\theta_{Y,B}(\bar{z},\bar{\lambda})(0)\cap\ker\nabla\Phi(\bar{x})^*.$$

Combining these facts verifies the equivalence between conditions (iii) and (iv).

Next we show that (i)$\Longrightarrow$(iv). Assume on the contrary that the dual qualification condition (4.14) fails while (i) holds, and so find an element

$$\eta\in D\partial\theta_{Y,B}(\bar{z},\bar{\lambda})(0)\cap\ker\nabla\Phi(\bar{x})^* \ \text{ such that } \ \eta\neq 0.$$

Since $\Psi(\bar{x},\bar{\lambda}+t\eta)=0$ for any $t>0$, we get from $\eta\in D\partial\theta_{Y,B}(\bar{z},\bar{\lambda})(0)$ and (2.15) that $\eta\in\partial\theta_{\mathcal{K},B}(0)$, and hence $-B\eta\in N_{\mathcal{K}}(\eta)$ by Theorem 2.1(ii). Choosing t to be sufficiently small and employing the Reduction Lemma from [34, Lemma 2E.4] ensure the existence of a neighbored U of $(0,0)\in\mathbb{R}^m\times\mathbb{R}^m$ such that

$$t(\eta,-B\eta)\in[\text{gph}\,N_{\mathcal{K}}]\cap U=\big[\text{gph}\,N_Y-(\bar{\lambda},\bar{z}-B\bar{\lambda})\big]\cap U.$$

This in turn results in $\bar{z}-B\bar{\lambda}-tB\eta\in N_Y(\bar{\lambda}+t\eta)$, which yields by (2.13) the inclusion $\bar{\lambda}+t\eta\in\partial\theta_{Y,B}(\bar{z})$. Combining the latter with $\Psi(\bar{x},\bar{\lambda}+t\eta)=0$ results in $\bar{\lambda}+t\eta\in\Lambda(\bar{x})$. However, we have $\eta\neq 0$ thus $\bar{\lambda}+t\eta\neq\bar{\lambda}$ for any $t>0$, which contradicts (i) and so verifies the claimed implication (i)$\Longrightarrow$(iv).

To show further that the isolated calmness of $M_{\bar{x}}$ at $\big((0,0),\bar{\lambda}\big)$ imposed in (iii) yields (ii), it suffices to check that $\Lambda(\bar{x})=\{\bar{\lambda}\}$. Indeed, the assumed isolated calmness allows us to find a neighborhood O of $\bar{\lambda}$ such that $M_{\bar{x}}(0,0)\cap O=\{\bar{\lambda}\}$, which tells us by the convex-valuedness of $M_{\bar{x}}$ that $M_{\bar{x}}(0,0)=\{\bar{\lambda}\}$. Combining the latter with $M_{\bar{x}}(0,0)=\Lambda(\bar{x})$ verifies (ii). Since (ii) obviously implies (i), we complete the proof of the theorem. $\square$

The next example reveals that the dual qualification condition (4.14) is essential for the uniqueness of Lagrange multipliers in Theorem 4.5.

Example 4.6. (nonuniqueness of Lagrange multipliers under failure of the dual qualification condition). Consider the variational system (1.3) with term (1.1), where Y and B are taken from (4.8), while $\Phi\colon \mathbb{R}^2 \to \mathbb{R}^2$ is defined by $\Phi(x_1, x_2) := (x_1, 0)$ and $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ is defined by $f(x) = 0$ for all $x \in \mathbb{R}^2$. It is shown in Example 4.4 that $\operatorname{dom}\theta_{Y,B} = \mathbb{R} \times \mathbb{R}_-$. Letting $\bar{x} := (0, 0)$, we get by the direct calculation that

$$\partial\theta_{Y,B}(\bar{x}) = \{0\} \times \mathbb{R}_+ \ \text{ and } \ \Psi(\bar{x}, \lambda) = \nabla\Phi(x)^*\lambda = (\lambda_1, 0),$$

and so $\Lambda(\bar{x}) = \{0\} \times \mathbb{R}_+$, which is not a singleton.

Let us now show that the dual qualification condition fails in this setting. Having $\ker\nabla\Phi(\bar{x})^* = \{0\} \times \mathbb{R}$ and choosing $\bar{\lambda} := (0, 0)$ give us the critical cone

$$\mathcal{K} = T_Y(\bar{\lambda}) \cap \big\{\Phi(\bar{x}) - B\bar{\lambda}\big\}^{\perp} = Y,$$

and so $\partial\theta_{\mathcal{K},B}(0, 0) = \{0\} \times \mathbb{R}_+$. Combining it with (2.15), we arrive at

$$\partial\theta_{\mathcal{K},B}(0, 0) \cap \ker\nabla\Phi(\bar{x})^* = D\partial\theta_{Y,B}(\bar{z}, \bar{\lambda})(0, 0) \cap \ker\nabla\Phi(\bar{x})^* = \{0\} \times \mathbb{R}_+ \neq \{(0, 0)\},$$

which demonstrates the failure of the dual qualification condition (4.14).

4.3 Characterizations of Noncritical Multipliers

In this section we derive major characterizations of noncritical multipliers for the piecewise linear-quadratic variational systems (1.3) in terms of semi-isolated calmness and error bounds.

Using the mapping G from (4.12), define the *solution map* $S\colon \mathbb{R}^n \times \mathbb{R}^m \rightrightarrows \mathbb{R}^n \times \mathbb{R}^m$ for the *canonical perturbation* of system (1.3) by

$$S(p_1, p_2) := \big\{(x, \lambda) \in \mathbb{R}^n \times \mathbb{R}^m \big|\ (p_1, p_2) \in G(x, \lambda)\big\}. \tag{4.15}$$

The property of *semi-isolated calmness* used in (4.17) was introduced in [32] for solution maps to general variational systems with a product structure of values as in (4.15). The reader can see that for such mappings the semi-isolated calmness of the variational systems of type (1.3) occupies an intermediate position between the calmness and isolated calmness.

In what follows we use the notation $\mathrm{dist}(x;\Omega)$ for the distance between a point $x \in \mathbb{R}^n$ and a set $\Omega \subset \mathbb{R}^n$, $\mathbb{B}_\varepsilon(x)$ for the closed ball centered at $x \in \mathbb{R}^n$ with radius $\varepsilon > 0$, and

$$P\varphi(x) := \mathrm{argmin}\Big\{\varphi(u) + \frac{1}{2}\|x-u\|^2 \Big| \ u \in \mathbb{R}^n\Big\}, \quad x \in \mathbb{R}^n, \tag{4.16}$$

for the *proximal mapping* $P\varphi\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ associated with a function $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$.

Theorem 4.7. (major characterizations of noncritical multipliers in variational systems). *Let* $(\bar{x},\bar{\lambda})$ *be a solution to the variational system* (1.3) *with the piecewise linear-quadratic term* (1.1). *Then the following conditions are equivalent:*

(i) *The Lagrange multiplier* $\bar{\lambda}$ *is noncritical for* (1.3) *corresponding to* $\bar{x}$.

(ii) *There exist numbers* $\varepsilon > 0$, $\ell \ge 0$ *and neighborhoods* U *of* $0 \in \mathbb{R}^n$ *and* W *of* $0 \in \mathbb{R}^m$ *such that for any* $(p_1,p_2) \in U \times W$ *the following inclusion holds:*

$$S(p_1,p_2) \cap \mathbb{B}_\varepsilon(\bar{x},\bar{\lambda}) \subset \big[\{\bar{x}\} \times \Lambda(\bar{x})\big] + \ell\big(\|p_1\| + \|p_2\|\big)\mathbb{B}. \tag{4.17}$$

(iii) *There exist numbers* $\varepsilon > 0$ *and* $\ell \ge 0$ *such that the error bound estimate*

$$\|x-\bar{x}\| + \mathrm{dist}\big(\lambda;\Lambda(\bar{x})\big) \le \ell\Big(\|\Psi(x,\lambda)\| + \mathrm{dist}\big(\Phi(x);(\partial\theta_{Y,B})^{-1}(\lambda)\big)\Big)$$

holds for any $(x,\lambda) \in \mathbb{B}_\varepsilon(\bar{x},\bar{\lambda})$ *in terms of the inverse subdifferential of* $\theta_{Y,B}$.

(iv) *There are numbers* $\varepsilon > 0$ *and* $\ell \ge 0$ *such that the error bound estimate*

$$\|x-\bar{x}\| + \mathrm{dist}\big(\lambda;\Lambda(\bar{x})\big) \le \ell\Big(\|\Psi(x,\lambda)\| + \|\Phi(x) - (P\theta_{Y,B})(\lambda + \Phi(x))\|\Big) \tag{4.18}$$

holds for any $(x,\lambda) \in \mathbb{B}_\varepsilon(\bar{x},\bar{\lambda})$ *in terms of the proximal mapping* $P\theta_{Y,B}$ *from* (4.16).

Proof. Let us first verify that (ii) implies (i). Theorem 4.2 reduces it to proving that the semi-isolated calmness property in (ii) ensures that for any solution $(\xi,\eta) \in \mathbb{R}^n \times \mathbb{R}^m$ to the system (4.4) we have $\xi = 0$. Define $(x_t,\lambda_t) := (\bar{x}+t\xi,\bar{\lambda}+t\eta)$ for all $t > 0$ and observe that

$$\begin{aligned}\Psi(x_t,\lambda_t)-\Psi(\bar{x},\bar{\lambda}) &= \big(f(x_t)-f(\bar{x})\big)+\big(\nabla\Phi(x_t)-\nabla\Phi(\bar{x})\big)^*\bar{\lambda}+t\nabla\Phi(x_t)^*\eta\\ &= t\nabla f(\bar{x})\xi+o(t)+t\big(\nabla^2\Phi(\bar{x})\xi\big)^*\bar{\lambda}+t\nabla\Phi(\bar{x})^*\eta+o(t)\\ &= t\big(\nabla_x\Psi(\bar{x},\bar{\lambda})\xi+\nabla\Phi(\bar{x})^*\eta\big)+o(t)=o(t)\end{aligned}$$

whenever t is sufficiently small. Letting $p_{1t}:=\Psi(x_t,\lambda_t)$ and using $\Psi(\bar{x},\bar{\lambda})=0$, we deduce from the last equality above that $p_{1t}=o(t)$. It follows in the similar way that

$$\Phi(x_t)=\Phi(\bar{x})+t\nabla\Phi(\bar{x})\xi+o(t)\ \text{ for all small }\ t>0.$$

Denoting further $z_t:=\Phi(\bar{x})+t\nabla\Phi(\bar{x})\xi$ implies that

$$z_t-\Phi(x_t)=o(t)\ \text{ as }\ t>0,$$

and therefore we get $p_{2t}=o(t)$ for $p_{2t}:=z_t-\Phi(x_t)$.

Let us now prove that $(x_t,\lambda_t)\in S(p_{1t},p_{2t})$ for $t>0$ sufficiently small. Since $p_{1t}=\Psi(x_t,\lambda_t)$, we only need to verify by Theorem 2.1(ii) that

$$\lambda_t\in\partial\theta_{Y,B}(z_t)=(N_Y+B)^{-1}(z_t),\ \text{ or equivalently }\ z_t-B\lambda_t\in N_Y(\lambda_t). \tag{4.19}$$

To proceed with checking (4.19), deduce from (4.4) that

$$\eta\in\mathcal{K}=K_Y(\bar{\lambda},\bar{z}-B\bar{\lambda})=T_Y(\bar{v})\cap\{\bar{z}-B\bar{\lambda}\}^{\perp}.$$

Denoting $\lambda_t:=\bar{\lambda}+t\eta$ and remembering that Y is a convex polyhedral set, we conclude that $\lambda_t\in Y$ for all $t>0$ sufficiently small. Furthermore, it follows from (4.4) that

$$\nabla\Phi(\bar{x})\xi-B\eta\in\mathcal{K}^*=N_Y(\bar{\lambda})+\mathbb{R}(\bar{z}-B\bar{\lambda}).$$

Thus there exist $\alpha\in\mathbb{R}$ and $w\in N_Y(\bar{\lambda})$ such that $\nabla\Phi(\bar{x})\xi-B\eta=\alpha(\bar{z}-B\bar{\lambda})+w$. Using this together with (4.4) gives us the equalities

$$0=\langle\nabla\Phi(\bar{x})\xi-B\eta,\eta\rangle=\alpha\langle\bar{z}-B\bar{\lambda},\eta\rangle+\langle w,\eta\rangle=\langle w,\eta\rangle.$$

Recall that $N_Y(\bar{\lambda})=\{\sum_{i\in I(\bar{\lambda})}\beta_i b_i|\ \beta_i\ge 0\}$, where $I(\bar{\lambda})$ stands for the set of active constraints in Y at $\bar{\lambda}$. It allows us to deduce from the inclusion $w\in N_Y(\bar{\lambda})$ that there are numbers $\beta_i \geq 0$ as $i \in I(\bar{\lambda})$ such that $w = \sum_{i\in I(\bar{\lambda})} \beta_i b_i$, and therefore

$$\sum_{i\in I(\bar{\lambda})} \beta_i \langle b_i, \eta\rangle = \langle w, \eta\rangle = 0.$$

Observe furthermore the relationships

$$z_t - B\lambda_t = \Phi(\bar{x}) + t\nabla\Phi(\bar{x})\xi - B\bar{\lambda} - tB\eta = \bar{z} - B\bar{\lambda} + t(\nabla\Phi(\bar{x})\xi - B\eta) = (1+t\alpha)(\bar{z} - B\bar{\lambda}) + tw,$$

where $1 + t\alpha > 0$ for small $t > 0$. Since both $\bar{z} - B\bar{\lambda}$ and w belong to $N_Y(\bar{\lambda})$, it follows that $(1 + t\alpha)(\bar{z} - B\bar{\lambda}) + tw \in N_Y(\bar{\lambda})$, and thus there is $\tau_{it} \geq 0$ for $i \in I(\bar{\lambda})$ such that $z_t - B\lambda_t = \sum_{i\in I(\bar{\lambda})} \tau_{it} b_i$. Noting that $\langle z_t - B\lambda_t, \eta\rangle = 0$ and $\langle b_i, \eta\rangle \leq 0$ for all $i \in I(\bar{\lambda})$, we deduce that

$$\langle b_i, \eta\rangle = 0 \text{ for all } i \in I(\bar{\lambda}) \text{ with } \tau_{it} > 0. \tag{4.20}$$

Let us now show that

$$\tau_{it} = 0 \ \text{ if } \ i \in I(\bar{\lambda}) \setminus I(\lambda_t).$$

Suppose on the contrary that there is an index $i_0 \in I(\bar{\lambda}) \setminus I(\lambda_t)$ for which $\tau_{i_0 t} > 0$. This means that $\langle b_{i_0}, \bar{\lambda}\rangle = \alpha_{i_0}$ and $\langle b_{i_0}, \lambda_t\rangle < \alpha_{i_0}$. Therefore

$$\langle b_{i_0}, \bar{\lambda}\rangle + t\langle b_{i_0}, \eta\rangle = \langle b_{i_0}, \lambda_t\rangle < \alpha_{i_0},$$

which in turn yields $\langle b_{i_0}, \eta\rangle < 0$, a contradiction with (4.20). Thus for all $i \in I(\bar{\lambda}) \setminus I(\lambda_t)$ we get $\tau_{it} = 0$ and hence arrive at

$$z_t - B\lambda_t = \sum_{i\in I(\lambda_t)} \tau_{it} b_i \in N_Y(\lambda_t).$$

This verifies (4.19) and thus implies that $(x_t, \lambda_t) \in S(p_{1t}, p_{2t})$. It now follows from the assumed semi-isolated calmness (4.17) in (ii) that

$$\|\xi\| = \frac{\|x_t - \bar{x}\|}{t} \leq \frac{\ell\big(\,\|p_{1t}\| + \|p_{2t}\|\,\big)}{t},$$

which results in $\xi = 0$ by letting $t \downarrow 0$. It tells us $\bar{\lambda}$ is noncritical and hence justify the implication (ii) $\Longrightarrow$ (i) of the theorem.

Next we prove the opposite implication (i)$\Longrightarrow$(ii). Assuming that the multiplier noncriticality in (i) holds, let us first verify the following statement.

Claim: *There exist numbers $\varepsilon > 0$ and $\ell \geq 0$ and neighborhoods U of $0 \in \mathbb{R}^n$ and W of $0 \in \mathbb{R}^m$ such that for any $(p_1, p_2) \in U \times W$ and $(x_{p_1p_2}, \lambda_{p_1p_2}) \in S(p_1, p_2) \cap \mathbb{B}_\varepsilon(\bar{x}, \bar{\lambda})$ we have*

$$\|x_{p_1p_2} - \bar{x}\| \leq \ell\big(\,\|p_1\| + \|p_2\|\,\big). \tag{4.21}$$

To justify this claim, suppose on the contrary that (4.21) fails and thus for any $k \in \mathbb{N}$ find $(p_{1k}, p_{2k}) \in \mathbb{B}_{1/k}(0) \times \mathbb{B}_{1/k}(0)$, $k \in \mathbb{N}$, and $(x_k, \lambda_k) \in S(p_{1k}, p_{2k}) \cap \mathbb{B}_{1/k}(\bar{x}, \bar{\lambda})$ such that

$$\frac{\|p_{1k}\| + \|p_{2k}\|}{\|x_k - \bar{x}\|} \to 0 \ \text{ as } \ k \to \infty.$$

Denote $t_k := \|x_k - \bar{x}\|$ and deduce from the convergence above that $p_{1k} = o(t_k)$ and $p_{2k} = o(t_k)$. Since $\theta_{Y,B}$ is a convex piecewise linear-quadratic function, it follows from the proof of [37, Theorem 11.14(b)] that $\operatorname{gph} \partial\theta_{Y,B}$ is a union of finitely many convex polyhedral sets. This together with [34, Theorem 3D.1] and $\bar{z} := \Phi(\bar{x}) \in \operatorname{dom} \partial\theta_{Y,B}$ ensures the existence of a number $\ell' \geq 0$ and a neighborhood O of $\bar{z}$ such that for all $z \in O \cap \operatorname{dom} \partial\theta_{Y,B}$ we have

$$\partial\theta_{Y,B}(z) \subset \partial\theta_{Y,B}(\bar{z}) + \ell'\, \|z - \bar{z}\|\, \mathbb{B}. \tag{4.22}$$

Suppose without loss of generality that $z_k := p_{2k} + \Phi(x_k) \in O$ for all $k \in \mathbb{N}$. Since $\lambda_k \in \partial\theta_{Y,B}(z_k)$, there exist $\lambda \in \partial\theta_{Y,B}(\bar{z})$ and $b \in \mathbb{B}$ such that $\lambda_k = \lambda + \ell'\, \|z_k - \bar{z}\|\, b$. Using this along with the classical Hoffman lemma, we find a number $M \geq 0$ such that

$$\begin{aligned} \operatorname{dist}\big(\lambda_k; \Lambda(\bar{x})\big) \ &\leq M\, \big(\|\Psi(\bar{x}, \lambda_k)\| + \operatorname{dist}\big(\lambda_k; \partial\theta_{Y,B}(\bar{z})\big)\big) \\ &\leq M\, \|\Psi(\bar{x}, \lambda_k) - \Psi(x_k, \lambda_k)\| + M\, \|\Psi(x_k, \lambda_k)\| + \ell'\, \|z_k - \bar{z}\| \\ &\leq M\rho(1 + \|\lambda_k\|)\, \|x_k - \bar{x}\| + M\, \|p_{1k}\| + \ell'\rho\, \|x_k - \bar{x}\| + \ell'\|p_{2k}\|, \end{aligned} \tag{4.23}$$

where ρ is a common calmness constant for the mappings f, Φ, and $\nabla\Phi$ at $\bar{x}$. Since $\Lambda(\bar{x})$ is closed and convex, for each $k \in \mathbb{N}$ there exists a vector $\mu_k \in \Lambda(\bar{x})$ for which

$$\frac{\|\lambda_k - \mu_k\|}{t_k} \leq M\rho(1 + \|\lambda_k\|) + M\frac{\|p_{1k}\|}{t_k} + \ell'\rho + \ell'\frac{\|p_{2k}\|}{t_k}, \quad k \in \mathbb{N}.$$

Thus we can assume without loss of generality that

$$\frac{\lambda_k-\mu_k}{t_k}\to\widetilde{\eta}\ \text{ for some }\ \widetilde{\eta}\in\mathbb{R}^m.$$

By passing to a subsequence if necessary, it follows that

$$\frac{x_k-\bar{x}}{t_k}\to\xi\ \text{ as }\ k\to\infty\ \text{ with some }\ 0\neq\xi\in\mathbb{R}^n.$$

Due to $\mu_k\in\Lambda(\bar{x})$ and the discussions above we get the equalities

$$\begin{aligned} o(t_k)=p_{1k} &= \Psi(x_k,\mu_k)=\Psi(x_k,\mu_k)-\Psi(\bar{x},\mu_k)+\nabla\Phi(x_k)^*(\lambda_k-\mu_k)\\ &=\nabla_x\Psi(\bar{x},\mu_k)(x_k-\bar{x})+\nabla\Phi(x_k)^*(\lambda_k-\mu_k)+o(t_k), \end{aligned}$$

which lead us as $k\to\infty$ to the limiting condition

$$\nabla_x\Psi(\bar{x},\bar{\lambda})\xi+\nabla\Phi(\bar{x})^*\widetilde{\eta}=0, \tag{4.24}$$

It further follows from $(x_k,\lambda_k)\in S(p_{1k},p_{2k})$ that $\lambda_k\in\partial\theta_{Y,B}(z_k)$, which is equivalent to the inclusion $z_k-B\lambda_k\in N_Y(\lambda_k)$ for each $k\in\mathbb{N}$ by Theorem 2.1(ii). Since Y is a convex polyhedral set, the Reduction Lemma from [34, Lemma 2E.4]) tells us that

$$z_k-B\lambda_k-(\bar{z}-B\bar{\lambda})\in N_{\mathcal{K}}(\lambda_k-\bar{\lambda})$$

for all $k\in\mathbb{N}$ sufficiently large, where $\mathcal{K}$ is the critical cone to Y at $\bar{z}$ for $\bar{z}-B\bar{\lambda}$ taken from Theorem 2.1(iii). This along with Theorem 2.1(iii) brings us to the conclusions

$$\begin{aligned} &\lambda_k-\bar{\lambda}\in\partial\theta_{\mathcal{K},B}(z_k-\bar{z})=D\partial\theta_{Y,B}(\bar{z},\bar{\lambda})(z_k-\bar{z}),\ \text{ and so}\\ &\frac{\lambda_k-\bar{\lambda}}{t_k}\in D\partial\theta_{Y,B}(\bar{z},\bar{\lambda})\Big(\frac{z_k-\bar{z}}{t_k}\Big)=\partial\theta_{\mathcal{K},B}\Big(\frac{z_k-\bar{z}}{t_k}\Big), \end{aligned} \tag{4.25}$$

which imply in turn that $\dfrac{z_k-\bar{z}}{t_k}\in\operatorname{dom}\partial\theta_{\mathcal{K},B}$. Since $\mathcal{K}$ is a convex polyhedral set, it follows from Theorem 2.1(i) that $\theta_{\mathcal{K},B}$ is a convex piecewise linear-quadratic function. Thus [37, Proposition 10.21] tells us that $\operatorname{dom}\partial\theta_{\mathcal{K},B}=\operatorname{dom}\theta_{\mathcal{K},B}$. Employing Theorem 2.1(i) ensures that $\operatorname{dom}\theta_{\mathcal{K},B}$ is a closed set. Combining it with the convergence $\dfrac{z_k-\bar{z}}{t_k}\to\nabla\Phi(\bar{x})\xi$ as $k\to\infty$ yields

$$\nabla\Phi(\bar{x})\xi\in\operatorname{dom}\partial\theta_{\mathcal{K},B}. \tag{4.26}$$

Since $\mu_k \in \Lambda(\bar{x})$, we get $\mu_k \in \partial\theta_{Y,B}(\bar{z})$ and, proceeding similarly to the proof of (4.25), arrive at

$$\frac{\mu_k - \bar{\lambda}}{t_k} \in \partial\theta_{\mathcal{K},B}(0).$$

Furthermore, it follows from $\bar{\lambda} \in \Lambda(\bar{x})$ and $\mu_k \in \Lambda(\bar{x})$ that $\bar{\lambda} - \mu_k \in \ker \nabla\Phi(\bar{x})^*$. Using (4.26) and arguing as in the proof of (4.22), we find $\ell' \geq 0$ and a neighborhood O of $\nabla\Phi(\bar{x})\xi$ such that

$$\partial\theta_{\mathcal{K},B}(u) \subset \partial\theta_{\mathcal{K},B}\big(\nabla\Phi(\bar{x})\xi\big) + \ell' \, \|u - \nabla\Phi(\bar{x})\xi\| \, \mathbb{B}$$

for all $u \in O \cap \operatorname{dom} \partial\theta_{\mathcal{K},B}$. Employing the latter together with (4.25) leads us to the relationships

$$\begin{aligned} \frac{\lambda_k - \mu_k}{t_k} &= \frac{\lambda_k - \bar{\lambda}}{t_k} + \frac{\bar{\lambda} - \mu_k}{t_k} \\ &\in \partial\theta_{\mathcal{K},B}\Big(\frac{z_k - \bar{z}}{t_k}\Big) - \big[\ker \nabla\Phi(\bar{x})^* \cap \partial\theta_{\mathcal{K},B}(0)\big] \\ &\subset \partial\theta_{\mathcal{K},B}(\nabla\Phi(\bar{x})\xi) + \ell' \Big\| \frac{z_k - \bar{z}}{t_k} - \nabla\Phi(\bar{x})\xi \Big\| \mathbb{B} - \big[\ker \nabla\Phi(\bar{x})^* \cap \partial\theta_{\mathcal{K},B}(0)\big]. \end{aligned}$$

This allows us to find, for all $k \in \mathbb{N}$ sufficiently large, a $b_k \in \mathbb{B}$ such that

$$\frac{\lambda_k - \mu_k}{t_k} - \ell' \Big\| \frac{z_k - \bar{z}}{t_k} - \nabla\Phi(\bar{x})\xi \Big\| b_k \in \partial\theta_{\mathcal{K},B}(\nabla\Phi(\bar{x})\xi) - \big[\ker \nabla\Phi(\bar{x})^* \cap \partial\theta_{\mathcal{K},B}(0)\big]. \tag{4.27}$$

We can see that the left-hand side of inclusion (4.27) converges as $k \to \infty$ to the vector $\widetilde{\eta}$. On the other hand, the right-hand side of this inclusion is the sum of two convex polyhedral sets, and so is closed. This shows that $\widetilde{\eta}$ satisfies to

$$\widetilde{\eta} \in \partial\theta_{\mathcal{K},B}(\nabla\Phi(\bar{x})\xi) - \big[\ker \nabla\Phi(\bar{x})^* \cap \partial\theta_{\mathcal{K},B}(0)\big]. \tag{4.28}$$

Thus we get vectors $\eta \in \partial\theta_{\mathcal{K},B}(\nabla\Phi(\bar{x})\xi)$ and $\eta' \in \ker \nabla\Phi(\bar{x})^* \cap \partial\theta_{\mathcal{K},B}(0)$, which provide the representation $\widetilde{\eta} = \eta - \eta'$. It follows from the relationship (2.15) in Theorem 2.1(iii) that $\eta \in D\partial\theta_{Y,B}(\bar{z}, \bar{\lambda})(\nabla\Phi(\bar{x})\xi)$. Furthermore, employing (4.24) tells us that

$$0 = \nabla_x \Psi(\bar{x}, \bar{\lambda})\xi + \nabla\Phi(\bar{x})^* \widetilde{\eta} = \nabla_x \Psi(\bar{x}, \bar{v})\xi + \nabla\Phi(\bar{x})^* \eta,$$

which contradicts the noncriticality of $\bar{\lambda}$ due to $\xi \neq 0$ and thus completes the proof of the claim.

To finalize verifying implication (i)$\Longrightarrow$(ii) in the theorem, take the neighborhoods U and W from the above claim and shrink them if necessary for the subsequent procedure. Using the claim and arguing similarly to the proof of the conditions in (4.23) give us a constant $\ell' \geq 0$ such that for any $(p_1, p_2) \in U \times W$ and any $(x_{p_1p_2}, \lambda_{p_1p_2}) \in S(p_1, p_2) \cap \mathbb{B}_\varepsilon(\bar{x}, \bar{\lambda})$ we have

$$\operatorname{dist}\big(\lambda_{p_1p_2}; \Lambda(\bar{x})\big) \leq \ell'\big(\,\|x_{p_1p_2} - \bar{x}\| + \|p_1\| + \|p_2\|\,\big). \tag{4.29}$$

Combining it with (4.21) allows us to find $\ell \geq 0$ for which $(p_1, p_2) \in U \times W$ and

$$\|x_{p_1p_2} - \bar{x}\| + \operatorname{dist}\big(\lambda_{p_1p_2}; \Lambda(\bar{x})\big) \leq \ell\big(\,\|p_1\| + \|p_2\|\,\big)$$

whenever $(x_{p_1p_2}, \lambda_{p_1p_2}) \in S(p_1, p_2) \cap \mathbb{B}_\varepsilon(\bar{x}, \bar{\lambda})$. This clearly justifies the semi-isolated calmness property (4.17) and thus finishes the proof of implication (i) $\Longrightarrow$ (ii).

The equivalence between (ii) and (iii) can be verified similarly to the corresponding arguments in the proof of [32, Theorem 4.1], and so we omit them here. Thus it remains to establish the equivalence between assertions (ii) and (iv) of the theorem to complete its proof.

Let us start with checking implication (iv)$\Longrightarrow$(ii). Picking $(p_1, p_2) \in \mathbb{B}_\varepsilon(0, 0)$ and $(x, \lambda) \in S(p_1, p_2) \cap \mathbb{B}_\varepsilon(\bar{x}, \bar{\lambda})$ with ε and ℓ taken from (iv), we get from the definition of S that

$$\Psi(x, \lambda) = p_1 \;\text{ and }\; \lambda \in \partial\theta_{Y,B}(\Phi(x) + p_2). \tag{4.30}$$

It follows from [37, Proposition 12.19] due to the convexity of $\theta_{Y,B}$ that $P\theta_{Y,B} = (I + \partial\theta_{Y,B})^{-1}$, and hence the second inclusion in (4.30) is equivalent to the equality $P\theta_{Y,B}(\lambda + \Phi(x) + p_2) = \Phi(x) + p_2$. Appealing now to (4.18) brings us to the estimates

$$\begin{aligned}\|x - \bar{x}\| + \operatorname{dist}\big(\lambda, \Lambda(\bar{x})\big) &\leq \ell\big(\|\Psi(x, \lambda)\| + \|\Phi(x) - P\theta_{Y,B}(\lambda + \Phi(x)\|\big)\\ &\leq \ell\big(\|p_1\| + \|P\theta_{Y,B}(\lambda + \Phi(x) + p_2) - P\theta_{Y,B}(\lambda + \Phi(x))\| + \|p_2\|\big)\\ &\leq \ell\big(\|p_1\| + \|p_2\| + \|p_2\|\big),\end{aligned}$$

which readily justify the assertion in (ii).

Finally, we verify the converse implication (ii)$\Longrightarrow$(iv). To proceed, pick $(x,\lambda)\in\mathbb{B}_{\varepsilon/2}(\bar{x},\bar{\lambda})$, where ε is taken from (ii). Define the vectors

$$p_2 := P\theta_{Y,B}\big(\lambda+\Phi(x)\big)-\Phi(x)\ \text{ and }\ p_1 := \Psi(x,\lambda-p_2). \tag{4.31}$$

Since Φ and $\nabla\Phi$ are continuous at $\bar{x}$ and since $P\theta_{Y,B}$ is Lipschitz continuous, we assume without loss of generality that $(p_1,p_2)\in\mathbb{B}_{\varepsilon/2}(0,0)$ and $\mathbb{B}_{\varepsilon/2}(0,0)\subset U\times W$, where U and W come from (ii). It follows from (4.31) that $(x,\lambda-p_2)\in S(p_1,p_2)\cap\mathbb{B}_{\varepsilon}(\bar{x},\bar{\lambda})$. Since $\nabla\Phi$ is continuous at $\bar{x}$, we can assume without loss generality that for some $\rho>0$ we have $\|\nabla\Phi(x)\|\le\rho$ for all $x\in\mathbb{B}_{\varepsilon}(\bar{x})$. So we deduce from (4.17) that

$$\begin{aligned}
\|x-\bar{x}\|+\operatorname{dist}\big(\lambda-p_2,\Lambda(\bar{x})\big) &\le \ell\big(\|p_1\|+\|p_2\|\big)\\
&\le \ell\big(\|\Psi(x,\lambda-p_2)\|+\|\Phi(x)-P\theta_{Y,B}(\lambda+\Phi(x))\|\big)\\
&\le \ell\big(\|\Psi(x,\lambda)\|+\rho\|p_2\|+\|\Phi(x)-P\theta_{Y,B}(\lambda+\Phi(x))\|\big)\\
&\le \ell\big(\|\Psi(x,\lambda)\|+(\rho+1)\|\Phi(x)-P\theta_{Y,B}(\lambda+\Phi(x))\|\big).
\end{aligned}$$

Recall that the distance function $\operatorname{dist}\big(\cdot;\Lambda(\bar{x})\big)$ is Lipschitz continuous; so we have

$$\operatorname{dist}\big(\lambda;\Lambda(\bar{x})\big)-\operatorname{dist}\big(\lambda-p_2;\Lambda(\bar{x})\big)\le\|p_2\|=\|\Phi(x)-P\theta_{Y,B}(\lambda+\Phi(x))\|, \tag{4.32}$$

which in combination with the obtained inequalities leads us to

$$\|x-\bar{x}\|+\operatorname{dist}\big(\lambda;\Lambda(\bar{x})\big)\le\ell\|\Psi(x,\lambda)\|+\big(\ell(\rho+1)+1\big)\|\Phi(x)-P\theta_{Y,B}(\lambda+\Phi(x))\|.$$

This verifies (iv) and completes the proof of the theorem. □

To conclude this section, let us mention some connection of the obtained characterizations of noncritical multipliers for variational systems (1.3) with the uniqueness of Lagrange multipliers therein, which is *not* assumed in Theorem 4.7. Indeed, looking more closely at the proof of theorem reveals that the second term in (4.28) is actually *undesired*, since it provides complications for the proof. But, as follows from Theorem 4.5, this terms disappears (reduces to $\{0\}$) if the set of Lagrange multipliers $\Lambda(\bar{x})$ is a singleton. This phenomenon has been recently observed in [33] for the case of constrained optimization problems.

4.4 Noncriticality in Extended Nonlinear Programming

Here we concentrate on problems of composite optimization given by (1.2), where $\theta = \theta_{Y,B}$ is taken from (1.1). It means that we are dealing with the class of ENLPs discussed in Section 1. Starting with this section we assume that φ_0 and Φ are not just twice differentiable, but belongs to the class of $\mathcal{C}^2$-smooth mappings around the points in question.

Define the *Lagrangian* of (1.2) by

$$L(x,\lambda) := \varphi_0(x) + \langle \Phi(x), \lambda\rangle - \frac{1}{2}\langle \lambda, B\lambda\rangle \ \text{ for } \ (x,\lambda) \in \mathbb{R}^n \times \mathbb{R}^m \tag{4.33}$$

and observe that the KKT system for (1.2) is written as

$$\nabla_x L(x,\lambda) = 0,\ \lambda \in \partial\theta_{Y,B}(\Phi(x)). \tag{4.34}$$

Thus (4.34) is a particular case of (1.3) with $\Psi := \nabla_x L$. Denoting

$$\Lambda_{\rm com}(\bar{x}) := \big\{\lambda \in \mathbb{R}^m \big|\ \nabla_x L(\bar{x},\lambda) = 0,\ \lambda \in \partial\theta_{Y,B}(\Phi(\bar{x}))\big\}, \tag{4.35}$$

the corresponding set of Lagrange multipliers, we have Definition 4.1 of multiplier criticality as well as all the above results being specified for the KKT system (4.34).

On the other hand, there are some phenomena concerning critical and noncritical Lagrange multipliers that distinguish KKT systems in optimization from general variational systems of type (1.3). We consider them in this and two subsequent sections.

The following theorem provides a certain *second-order sufficient condition* ensuring simultaneously the *strict minimality* of a feasible solution to ENLP (1.2) and the *noncriticality* of the corresponding Lagrange multiplier. In its formulation we use the critical cone $\mathcal{K}$ defined in Theorem 2.1(iii) as well as the notation $\operatorname{rge} A$ for the range of a linear operator A. Note that the existence of Lagrange multipliers corresponding to $\bar{x}$ in (1.2), which is assumed below, is ensured by the first-order qualification condition (5.3) from Lemma 5.1.

Theorem 4.8. (second-order sufficient condition for strict local minimizers and noncritical multipliers in ENLPs). *Let* $(\bar{x}, \bar{\lambda})$ *be a solution to KKT system* (4.34). *Assume further that the second-order sufficient condition*

$$\big\langle \nabla^2_{xx} L(\bar{x}, \bar{\lambda})w, w\big\rangle + 2\theta_{\mathcal{K},B}\big(\nabla\Phi(\bar{x})w\big) > 0 \ \text{ if } \ w \in \mathbb{R}^n \backslash \{0\} \text{ with } \nabla\Phi(\bar{x})w \in \mathcal{K}^* + \operatorname{rge} B \quad (4.36)$$

holds. Then if the following second-order condition

$$\langle \nabla^2_{xx} L(\bar{x}, \bar{\lambda})\xi, \xi\rangle + 2\theta_{\mathcal{K},B}(\nabla\Phi(\bar{x})\xi) > 0 \text{ for all } \xi \neq 0 \text{ and } \xi \in \{w | \nabla\Phi(\bar{x})w \in \mathcal{K}^* + ImB\} \quad (4.37)$$

is satisfied, then $\bar{x}$ *is a strict local minimizer for* (1.2). *Furthermore, the Lagrange multiplier* $\bar{\lambda}$ *satisfying* (4.37) *is noncritical for the KKT system* (4.34) *corresponding to* $\bar{x}$.

Proof. We shall prove the first part of the theorem by contradiction. Suppose that $\bar{x}$ is not a strict local minimizer of the problem (1.2). Then there exists a sequence x_k with $x_k \to \bar{x}$ and $\Phi(x_k) \in \operatorname{dom}\theta$ such that

$$\phi_0(x_k) + \theta(\Phi(x_k)) \leq \phi_0(\bar{x}) + \theta(\Phi(\bar{x})),$$

which is equivalent to

$$\theta(\Phi(x_k)) - \theta(\Phi(\bar{x})) \leq \phi_0(\bar{x}) - \phi_0(x_k). \quad (4.38)$$

Without loss of generality we can assume that $\dfrac{x_k - \bar{x}}{\|x_k - \bar{x}\|} \to \xi$, where $\xi \neq 0$.
Denote $t_k := \|x_k - \bar{x}\|$ and $z_k := \Phi(x_k)$. For any $\lambda \in \partial\theta(\bar{z})$ we have

$$\langle \lambda, z_k - \bar{z}\rangle \leq \theta(z_k) - \theta(\bar{z}). \quad (4.39)$$

Combining the above with (4.38) we get that

$$\langle \lambda, \Phi(x_k) - \Phi(\bar{x})\rangle \leq \theta(\Phi(x_k)) - \theta(\Phi(\bar{x})) \leq -[\phi_0(x_k) - \phi_0(\bar{x})]$$

for any $\lambda \in \partial\theta(\bar{z})$. Dividing both sides by t_k and passing to limit give us

$$\langle \lambda, \nabla\Phi(\bar{x})\xi\rangle \leq -\nabla\phi_0(\bar{x})\xi \text{ for all } \lambda \in \partial\theta(\bar{z}). \quad (4.40)$$

Notice that it follows from $\bar{\lambda} \in \Lambda_{\text{com}}(\bar{x})$ that $\nabla_x L(\bar{x}, \bar{\lambda}) = 0$, and therefore $\nabla\Phi(\bar{x})^*\bar{\lambda} =$

$-\nabla\phi_0(\bar{x})$. Thus

$$\langle\bar{\lambda}, \nabla\Phi(\bar{x})\xi\rangle = \langle\nabla\Phi(\bar{x})^*\bar{\lambda}, \bar{\lambda}\rangle = -\nabla\phi_0(\bar{x})\xi, \tag{4.41}$$

and so by Proposition 2.1(iv) we have

$$d\theta(\bar{z})(\nabla\Phi(\bar{x})\xi) = -\nabla\phi_0(\bar{x})\xi = \langle\bar{\lambda}, \nabla\Phi(\bar{x})\xi\rangle, \tag{4.42}$$

which by [37, Proposition 13.9] shows that $d^2\theta(\bar{z},\bar{\lambda})(\nabla\Phi(\bar{x})\xi)$ is finite. Since $d^2\theta(\bar{z},\bar{\lambda}) = 2\theta_{\mathcal{K},B}$, it follows that $\theta_{\mathcal{K},B}(\nabla\Phi(\bar{x})\xi) < \infty$, or $\nabla\Phi(\bar{x})\xi \in \mathrm{dom}\theta_{\mathcal{K},B} = \mathcal{K}^* + ImB$. We know that $\mathrm{dom}\partial\theta_{\mathcal{K},B} = \mathrm{dom}\theta_{\mathcal{K},B}$, therefore $\nabla\Phi(\bar{x})\xi \in \mathrm{dom}\partial\theta_{\mathcal{K},B}$. Let $\eta \in \partial\theta_{\mathcal{K},B}(\nabla\Phi(\bar{x})\xi) = (N_{\mathcal{K}} + B)^{-1}(\nabla\Phi(\bar{x})\xi)$. This is equivalent to the following

$$\begin{cases} \eta \in \mathcal{K} = K_Y(\bar{\lambda}, \bar{z} - B\bar{\lambda}) = T_Y(\bar{\lambda}) \cap \{\bar{z} - B\bar{\lambda}\}^{\perp}, \\ \nabla\Phi(\bar{x})\xi - B\eta \in \mathcal{K}^*, \\ \langle\nabla\Phi(\bar{x})\xi - B\eta, \eta\rangle = 0. \end{cases} \tag{4.43}$$

Since $\eta \in \partial\theta_{\mathcal{K},B}(\nabla\Phi(\bar{x})\xi) = \arg\max_{y\in\mathcal{K}}\{\langle\nabla\Phi(\bar{x})\xi, y\rangle - \dfrac{1}{2}\langle y, By\rangle\}$, we have

$$\begin{aligned} \theta_{\mathcal{K},B}(\nabla\Phi(\bar{x})\xi) &= \langle\nabla\Phi(\bar{x})\xi, \eta\rangle - \tfrac{1}{2}\langle\eta, B\eta\rangle \\ &= \langle\nabla\Phi(\bar{x})^*\eta, \xi\rangle - \tfrac{1}{2}\langle\eta, B\eta\rangle. \end{aligned} \tag{4.44}$$

Now set $\lambda_k := \bar{\lambda} + t_k\eta$. Since Y is a closed polyhedral and $\eta \in T_Y(\bar{\lambda})$, for sufficiently large k vectors $\lambda_k \in Y$. By the definition of function θ we obtain

$$\langle\Phi(x_k), \lambda_k\rangle - \frac{1}{2}\langle\lambda_k, B\lambda_k\rangle \le \theta(\Phi(x_k)).$$

Furthermore, since $\bar{\lambda} \in \partial\theta(\bar{z}) = \partial\theta(\Phi(\bar{x}))$ we have

$$\langle\Phi(\bar{x}), \bar{\lambda}\rangle - \frac{1}{2}\langle\bar{\lambda}, B\bar{\lambda}\rangle = \theta(\Phi(\bar{x})).$$

Combining the two above inequalities with (4.38) gives us

$$\langle\Phi(x_k), \lambda_k\rangle - \frac{1}{2}\langle\lambda_k, B\lambda_k\rangle - \big(\langle\Phi(\bar{x}), \bar{\lambda}\rangle - \frac{1}{2}\langle\bar{\lambda}, B\bar{\lambda}\rangle\big) \le \phi_0(\bar{x}) - \phi_0(x_k),$$

or equivalently

$$L(x_k, \lambda_k) - L(\bar{x}, \bar{\lambda}) \le 0 \text{ for all large } k. \tag{4.45}$$

The Taylor series of $L(x,\lambda)$ at $(\bar{x},\bar{\lambda})$ gives us

$$\begin{aligned} L(x_k,\lambda_k)-L(\bar{x},\bar{\lambda}) = \; & \nabla_x L(\bar{x},\bar{\lambda})(x_k-\bar{x})+\nabla_\lambda L(\bar{x},\bar{\lambda})(\lambda_k-\bar{\lambda})+ \\ & \tfrac{1}{2}\langle\nabla^2_{xx}L(\bar{x},\bar{\lambda})(x_k-\bar{x}),x_k-\bar{x}\rangle+\tfrac{1}{2}\langle\nabla^2_{\lambda\lambda}L(\bar{x},\bar{\lambda})(\lambda_k-\bar{\lambda}),\lambda_k-\bar{\lambda}\rangle+ \\ & \langle\nabla^2_{x\lambda}L(\bar{x},\bar{\lambda})(\lambda_k-\bar{\lambda}),x_k-\bar{x}\rangle+o\big(\|x_k-\bar{x}\|^2+\|\lambda_k-\bar{\lambda}\|^2\big). \end{aligned} \tag{4.46}$$

Proceeding with the calculations we get the following:

- $\nabla_x L(\bar{x},\bar{\lambda})=0$
- $\nabla_\lambda L(\bar{x},\bar{\lambda})(\lambda_k-\bar{\lambda})=\langle\Phi(\bar{x})-B\bar{\lambda},\lambda_k-\bar{\lambda}\rangle=\langle\bar{z}-B\bar{\lambda},\lambda_k-\bar{\lambda}\rangle$
- $\langle\nabla^2_{\lambda\lambda}L(\bar{x},\bar{\lambda})(\lambda_k-\bar{\lambda}),\lambda_k-\bar{\lambda}\rangle=-\langle B(\lambda_k-\bar{\lambda}),\lambda_k-\bar{\lambda}\rangle$
- $\langle\nabla^2_{x\lambda}L(\bar{x},\bar{\lambda})(\lambda_k-\bar{\lambda}),x_k-\bar{x}\rangle=\langle\nabla\Phi(\bar{x})^*(\lambda_k-\bar{\lambda}),x_k-\bar{x}\rangle$.

Therefore

$$\begin{aligned} L(x_k,\lambda_k)-L(\bar{x},\bar{\lambda}) &= \langle\bar{z}-B\bar{\lambda},\lambda_k-\bar{\lambda}\rangle+\tfrac{1}{2}\langle\nabla^2_{xx}L(\bar{x},\bar{\lambda})(x_k-\bar{x}),x_k-\bar{x}\rangle \\ &\quad -\tfrac{1}{2}\langle B(\lambda_k-\bar{\lambda}),\lambda_k-\bar{\lambda}\rangle+\langle\nabla\Phi(\bar{x})^*(\lambda_k-\bar{\lambda}),x_k-\bar{x}\rangle+o(t_k^2+\|\eta\|\,t_k^2) \\ &= \langle\bar{z}-B\bar{\lambda},t_k\eta\rangle+\tfrac{1}{2}\langle\nabla^2_{xx}L(\bar{x},\bar{\lambda})(x_k-\bar{x}),x_k-\bar{x}\rangle- \\ &\quad \tfrac{1}{2}\langle B(\lambda_k-\bar{\lambda}),\lambda_k-\bar{\lambda}\rangle+\langle\nabla\Phi(\bar{x})^*(\lambda_k-\bar{\lambda}),x_k-\bar{x}\rangle+o(t_k^2). \end{aligned}$$

Since $\eta\in\mathcal{K}=T_Y(\bar{\lambda})\cap\{\bar{z}-B\bar{\lambda}\}^\perp$, we get that $\langle\bar{z}-B\bar{\lambda},\eta\rangle=0$, and hence

$$L(x_k,\lambda_k)-L(\bar{x},\bar{\lambda})=\frac{1}{2}\langle\nabla^2_{xx}L(\bar{x},\bar{\lambda})(x_k-\bar{x}),x_k-\bar{x}\rangle-\frac{1}{2}t_k^2\langle B\eta,\eta\rangle+t_k\langle\nabla\Phi(\bar{x})^*\eta,x_k-\bar{x}\rangle+o(t_k^2).$$

Dividing both sides by t_k^2 and passing to limit and also combining with (4.44) we get

$$\begin{aligned} \lim_{k\to\infty}\frac{L(x_k,\lambda_k)-L(\bar{x},\bar{\lambda})}{t_k^2} &= \tfrac{1}{2}\langle\nabla^2_{xx}L(\bar{x},\bar{\lambda})\xi,\xi\rangle-\tfrac{1}{2}\langle B\eta,\eta\rangle+\langle\nabla\Phi(\bar{x})^*\eta,\xi\rangle \\ &= \tfrac{1}{2}\langle\nabla^2_{xx}L(\bar{x},\bar{\lambda})\xi,\xi\rangle+\theta_{\mathcal{K},B}(\nabla\Phi(\bar{x})\xi) \end{aligned}$$

By the assumption of the theorem, the right hand side of the above equality is always positive. However, by (4.45) the left hand side is nonpositive, which is a contradiction. This shows that $\bar{x}$ must be a strict local minimizer of the problem (1.2).

It is easy to see that if a vector $v \in \Lambda_{\rm com}(\bar{x})$ satisfies the condition (4.37), the system (4.4) associated with v can only have a solution (ξ, η) with $\xi = 0$, which implies by Theorem 4.2 that v is noncritical. This completes the proof of the theorem. $\square$

The next example, which revisits Example 4.3 in the ENLP framework, illustrates the possibility to use the second-order sufficient condition (4.37) to justify the strict optimality of a feasible solution to (1.2) and the noncriticality of the corresponding Lagrange multiplier.

Example 4.9. (multiplier noncriticality via the second-order sufficient condition). Consider the ENLP from (1.2), where $m = n$, $\varphi_0(x) := x_1^2 + \ldots + x_n^2$ and $\Phi(x) := x$, and where Y and B are taken from Example 4.3. Then we have

$$\begin{aligned} \theta_{Y,B}\big(\Phi(x)\big) &= \sup_{y \in \mathbb{R}^n_+} \Big\{ \langle y, \Phi(x) \rangle - \frac{1}{2}\langle y, y \rangle \Big\} \\ &= \sup_{(y_1,\ldots,y_n) \in \mathbb{R}^n_+} \Big\{ \sum_{i=1}^n \big(x_i y_i - \frac{1}{2} y_i^2 \big) \Big\} \\ &= \frac{1}{2} \sum_{i=1}^n \big(\max\{x_i, 0\} \big)^2. \end{aligned} \tag{4.47}$$

Let us check that condition (4.37) holds when $\bar{x} = 0$ and $\bar{\lambda} = 0$, which confirms by Theorem 4.8 that $\bar{x}$ is a strict minimizer for this ENLP and $\bar{\lambda}$ is the corresponding noncritical multiplier. Indeed, it follows from Example 4.3 that $\bar{\lambda} \in \partial\theta(\bar{z})$, where $\bar{z} := \Phi(\bar{x}) = 0$. By the structure of $L(x, \lambda)$ we have the expressions

$$\nabla_x L(x,\lambda) = (2x_1 + \lambda_1, \ldots, 2x_n + \lambda_n) \ \text{ and } \ \nabla^2_{xx} L(x,\lambda) = 2I.$$

Then $\nabla_x L(\bar{x}, \bar{\lambda}) = 0$ and hence $\bar{\lambda} \in \Lambda_{\rm com}(\bar{x})$. Since ${\rm rge}\, B = \mathbb{R}^n$, it follows that $\{w |\ \nabla\Phi(\bar{x})w \in \mathcal{K}^* + {\rm rge}\, B\} = \mathbb{R}^n$, and therefore the sufficient condition in Theorem 4.8 reads as

$$\langle \nabla^2_{xx} L(\bar{x}, \bar{\lambda})\xi, \xi \rangle + 2\theta_{\mathcal{K},B}\big(\nabla\Phi(\bar{x})\xi\big) > 0 \ \text{ for all } \ \xi \neq 0,$$

which is equivalently presented by

$$2\langle \xi, \xi \rangle + 2\theta_{\mathcal{K},B}\big(\nabla\Phi(\bar{x})\xi\big) > 0 \ \text{ for all } \ \xi \neq 0. \tag{4.48}$$

Furthermore, Example 4.3 tells us that $\mathcal{K} = \mathbb{R}^n_+ \cap \{\bar{z}\}^\perp$ and so $\mathcal{K} = \mathbb{R}^n_+ = Y$. Combining this with (4.47), the sufficient condition (4.37) now becomes

$$2\langle\xi,\xi\rangle+2\theta_{Y,B}\big(\nabla\Phi(\bar{x})\xi\big)>0\ \text{ for all }\ \xi\neq 0. \tag{4.49}$$

Since $\theta_{Y,B}$ from (4.47) is always nonnegative, condition (4.49) holds, and thus it confirms the strict minimality of $\bar{x}$ and the noncriticality of $\bar{\lambda}$.

CHAPTER 5 NONCRITICALITY AND ITS APPLICATIONS IN STABILITY

5.1 Critical Multipliers and Full Stability of Minimizers in ENLPs

This section also deals with constrained minimization problems of the ENLP type and delivers as important message for both theoretical and numerical aspects of optimization. As discussed in Section 1, critical multipliers are particularly responsible for slow convergence of major primal-dual algorithms of optimization and are desired to be excluded for a given local minimizer. It is natural to suppose that seeking not arbitrary while just “nice" and stable in some sense local minimizers allows us to rule out the appearance of critical multipliers associated with such local optimal solutions. It is conjectured in [21] that fully stable local minimizers in the sense of [15] are appropriate candidate for excluding critical multipliers. This conjecture is affirmatively verified in [31] for problems (1.2) with $\theta=\theta_{Y,B}$ where $B=0$. Now we are able to extend this result to the general case of (1.1) with an arbitrary symmetric positive-semidefinite matrix B.

To proceed, we first specify the definition of fully stable local minimizers from [15] for problems (1.2) with term (1.1). Consider their *canonically perturbed* version described by

$$\text{minimize }\ \varphi_0(x)+\theta\big(\Phi(x)+p_2\big)-\langle p_1,x\rangle\ \text{ subject }\ x\in\mathbb{R}^n \tag{5.1}$$

with parameter pairs $(p_1,p_2)\in\mathbb{R}^n\times\mathbb{R}^m$. Fix $\gamma>0$ and $(\bar{x},\bar{p}_1,\bar{p}_2)$ with $\Phi(\bar{x})+\bar{p}_2\in\operatorname{dom}\theta$ and then define the parameter-depended optimal value function for (5.1) by

$$m_\gamma(p_1,p_2):=\inf_{\|x-\bar{x}\|\le\gamma}\big\{\varphi_0(x)+\theta\big(\Phi(x)+p_2\big)-\langle p_1,x\rangle\big\}$$

together with the parameterized set of optimal solutions to (5.1) given by

$$M_\gamma(p_1,p_2) := \mathop{\arg\min}_{\|x-\bar{x}\|\le\gamma} \big\{\varphi_0(x) + \theta\big(\Phi(x)+p_2\big) - \langle p_1, x\rangle\big\} \tag{5.2}$$

with the convention that $\arg\min := \emptyset$ when the expression under minimization in (5.2) is ∞. We say that $\bar{x}$ is a *fully stable* local optimal solution to problem (1.2) if there exist a number $\gamma > 0$ and neighborhoods U of $\bar{p}_1$ and W of $\bar{p}_2$ such that the mapping $(p_1,p_2) \mapsto M_\gamma(p_1,p_2)$ is single-valued and Lipschitz continuous with $M_\gamma(\bar{p}_1,\bar{p}_2) = \{\bar{x}\}$ and that the function $(p_1,p_2) \mapsto m_\gamma(p_1,p_2)$ is likewise Lipschitz continuous on $U \times W$.

Note that [15, Proposition 3.5] deduces the local Lipschitz continuity of m_γ from the *basic constraint qualification* (5.3) formulated in the following lemma, which is obtained in [37, Exercise 13.26]. The second-order necessary condition presented below can be viewed as a “no-gap" version of the second-order sufficient one used in Theorem 4.8 with the notation therein.

Lemma 5.1. (second-order necessary optimality condition for composite optimization problems). Let $\bar{x}$ be a local optimal solution to problem (1.2) with $\theta = \theta_{Y,B}$ taken from (1.1), and let the basic constraint qualification

$$N_{\operatorname{dom}\theta_{Y,B}}(\Phi(\bar{x})) \cap \ker \nabla\Phi(\bar{x})^* = \{0\} \tag{5.3}$$

be satisfied, and so $\Lambda_{\rm com}(\bar{x}) \neq \emptyset$. Then we have second-order necessary optimality condition

$$\max_{\lambda\in\Lambda_{\rm com}(\bar{x})} \big\langle \nabla^2_{xx}L(\bar{x},\lambda)w, w\big\rangle + 2\theta_{\mathcal{K},B}\big(\nabla\Phi(\bar{x})w\big) \ge 0 \tag{5.4}$$

valid for all $w \in \mathbb{R}^n$ with $\nabla\Phi(\bar{x})w \in \mathcal{K}^* + \operatorname{rge} B$.

Now we are ready to establish the aforementioned result in the general ENLP setting.

Theorem 5.2. (excluding critical multipliers by full stability of local minimizers). *Let $\bar{x}$ be a fully stable local optimal solution to problem* (1.2)*, and let θ be taken from* (1.1)*. Then the Lagrange multiplier set $\Lambda_{\rm com}(\bar{x})$ in* (4.35) *is nonempty and does not include critical multipliers.*

Proof. First we show that the full stability of $\bar{x}$ ensures the validity of the qualification condition (5.3). Indeed, pick any $\eta \in N_{\mathrm{dom}\,\theta_{Y,B}}(\Phi(\bar{x})) \cap \ker \nabla\Phi(\bar{x})^*$. Select $p_1 = \bar{p}_1 := 0$ and $p_2 := t\eta$ as $t \downarrow 0$. It follows from the full stability of $\bar{x}$ that there exist a Lipschitz constant $\ell \geq 0$ and the unique solution $x_{p_1p_2}$ to problem (5.1) such that

$$\|x_{p_1p_2} - \bar{x}\| \leq \ell t \|\eta\| . \tag{5.5}$$

Since $\Phi(x_{p_1p_2}) + p_2 \in \mathrm{dom}\,\theta_{Y,B}$ and $\eta \in N_{\mathrm{dom}\,\theta_{Y,B}}(\Phi(\bar{x}))$, we get $\langle \eta, \Phi(x_{p_1p_2}) + p_2 - \Phi(\bar{x})\rangle \leq 0$. This gives us the relationships

$$\begin{aligned} 0 \ &\geq \big\langle \eta, \nabla\Phi(\bar{x})(x_{p_1p_2} - \bar{x}) + o(\|x_{p_1p_2} - \bar{x}\|) + p_2 \big\rangle \\ &= \big\langle \nabla\Phi(\bar{x})^*\eta, x_{p_1p_2} - \bar{x} \big\rangle + \big\langle \eta, o(\|x_{p_1p_2} - \bar{x}\|) + p_2 \big\rangle \\ &= \big\langle \eta, o(\|x_{p_1p_2} - \bar{x}\|) \big\rangle + t\|\eta^2\|. \end{aligned}$$

Using estimate (5.5) and letting $t \downarrow 0$ lead to $\eta = 0$. Thus the basic constraint qualification (5.3) is satisfied, which ensures that $\Lambda_{\mathrm{com}}(\bar{x}) \neq \emptyset$.

Next we pick any $\bar{\lambda} \in \Lambda_{\mathrm{com}}(\bar{x})$ and show that it is noncritical for the unperturbed KKT system (4.34) corresponding to $\bar{x}$. Consider the KKT system for the perturbed problem (5.1) that can be written as

$$\begin{pmatrix} p_1 \\ \\ p_2 \end{pmatrix} \in \begin{pmatrix} \nabla_x L(x,\lambda) \\ \\ -\Phi(x) \end{pmatrix} + \begin{pmatrix} 0 \\ \\ (\partial\theta_{Y,B})^{-1}(\lambda) \end{pmatrix}. \tag{5.6}$$

Let $S_{KKT} \colon \mathbb{R}^n \times \mathbb{R}^m \rightrightarrows \mathbb{R}^n \times \mathbb{R}^m$ be the solution map to (5.6) given by

$$S_{KKT}(p_1,p_2) := \big\{ (x,\lambda) \in \mathbb{R}^n \times \mathbb{R}^m \big|\ p_1 = \nabla_x L(x,\lambda),\ \lambda \in \partial\theta_{Y,B}\big(p_2 + \Phi(x)\big) \big\}. \tag{5.7}$$

Employing Theorem 4.7, we only need to prove that there exist numbers $\varepsilon > 0$ and $\ell \geq 0$ as well as neighborhoods U of $0 \in \mathbb{R}^n$ and W of $0 \in \mathbb{R}^m$ such that for any $(p_1,p_2) \in U \times V$ and any $(x_{p_1p_2}, \lambda_{p_1p_2}) \in S_{KKT}(p_1,p_2) \cap (\mathbb{B}_\varepsilon(\bar{x}) \times \mathbb{B}_\varepsilon(\bar{\lambda}))$, estimate (4.17) holds with replacing $\Lambda(\bar{x})$ by the set of Lagrange multipliers $\Lambda_{\mathrm{com}}(\bar{x})$ taken from (4.35).

To this end we deduce from the full stability of $\bar{x}$ in (5.1) with $(\bar{p}_1,\bar{p}_2) = (0,0)$ due to the result of [31, Proposition 6.1] that there exist neighborhoods $\widetilde{U} \times \widetilde{W}$ of $(0,0)$ and $\widetilde{V}$ of

$\bar{x}$ for which the set-valued mapping

$$(p_1,p_2)\mapsto Q(p_1,p_2):=\big\{x\in\mathbb{R}^n\big|\ p_1\in\nabla\varphi_0(x)+\nabla\Phi(x)^*\partial\theta_{Y,B}(\Phi(x)+p_2)\big\}$$

admits a Lipschitzian single-valued graphical localization on $\widetilde{U}\times\widetilde{W}\times\widetilde{V}$. This means that there exists a Lipschitzian single-valued mapping $g\colon\widetilde{U}\times\widetilde{W}\mapsto\widetilde{V}$ such that $(\operatorname{gph}Q)\cap(\widetilde{U}\times\widetilde{W}\times\widetilde{V})=\operatorname{gph}g$. Denote $U:=\widetilde{U}$, $W:=\widetilde{W}$ and take $\varepsilon>0$ so small that $\mathbb{B}_\varepsilon(\bar{x})\subset\widetilde{V}$. The Lipschitzian single-valued graphical localization property of Q allows us to find a constant $\ell\geq 0$ such that for any $(p_1,p_2)\in U\times W$ and any $(x_{p_1p_2},\lambda_{p_1p_2})\in S_{KKT}(p_1,p_2)\cap\big(\mathbb{B}_\varepsilon(\bar{x})\times\mathbb{B}_\varepsilon(\bar{\lambda})\big)$ we have the inclusion $x_{p_1p_2}\in Q(p_1,p_2)$, and hence

$$\|x_{p_1p_2}-\bar{x}\|=\|x_{p_1p_2}-x_{\bar{p}_1\bar{p}_2}\|\leq\ell\big(\,\|p_1\|+\|p_2\|\,\big).$$

Using now the error bound estimate (4.29) from the proof of Theorem 4.7 with replacing $\Lambda(\bar{x})$ by $\Lambda_{\rm com}(\bar{x})$ and adjusting ε if necessary give us the semi-isolated calmness property (4.17), which is equivalent to the noncriticality of $\bar{\lambda}$ that was chosen arbitrary from the Lagrange multiplier set $\Lambda_{\rm com}(\bar{x})$. This therefore completes the proof of theorem. $\square$

The result of Theorem 5.2 calls for the deriving verifiable conditions for full stability of local minimizers to (1.2) expressed entirely via the problem data and the given minimizer. Such conditions allow us to efficiently exclude slow convergence of primal-dual algorithms to seek fully stable minimizers based on the initial data. Some characterizations of full stability of local minimizers for ENLPs of type (1.2) are obtained in [31, Theorem 7.3] under rather strong assumptions. Relaxing these assumptions is a challenging goal of our future research.

5.2 Noncriticality and Lipschitzian Stability of Solutions to ENLPs

In this section we use the machinery developed above to investigate other notions of Lipschitzian stability, which occur to be related to noncriticality of multipliers for ENLPs. The following theorem provides characterizations of both isolated calmness and robust isolated calmness properties of the KKT solution map (5.7) associated with ENLP (1.2) in terms of the second-order sufficient condition (4.37) as well as noncriticality and uniqueness of Lagrange multipliers.

Theorem 5.3. (characterizations of robust isolated calmness of solution maps). *Let $\bar{x}$ be a feasible solution to ENLP* (1.2) *with θ taken from* (1.1)*, and let $\bar{\lambda} \in \Lambda_{\rm com}(\bar{x})$ be a corresponding Lagrange multiplier from* (4.35)*. The following assertions are equivalent:*

- **(i)** *The solution map S_{KKT} from* (5.7) *is robustly isolatedly calm at the point $\big((0,0),(\bar{x},\bar{\lambda})\big) \in \mathbb{R}^{n+m} \times \mathbb{R}^{n+m}$, and $\bar{x}$ is a local optimal solution to* (1.2).
- **(ii)** *The second-order sufficient condition* (4.37) *holds, and $\Lambda_{\rm com}(\bar{x}) = \{\bar{\lambda}\}$.*
- **(iii)** *$\Lambda_{\rm com}(\bar{x}) = \{\bar{\lambda}\}$, $\bar{x}$ is a local optimal solution to* (1.2)*, and $\bar{\lambda}$ is a noncritical multiplier for* (1.3) *with $\Psi = \nabla_x L$ that is associated with the optimal solution $\bar{x}$.*
- **(iv)** *S_{KKT} is isolatedly calm at $\big((0,0),(\bar{x},\bar{\lambda})\big)$, and $\bar{x}$ is a local optimal solution to* (1.2).

Proof. The outline of the proof is as follows. We sequentially verify implications (ii)$\Longrightarrow$(iii), (iii)$\Longrightarrow$(iv), (iv)$\Longrightarrow$(iii), (iii)$\Longrightarrow$(ii), and (i) $\Longleftrightarrow$ (iv).

To prove (ii)$\Longrightarrow$(iii), assume the validity of (4.37) and that $\Lambda_{\rm com}(\bar{x}) = \{\bar{\lambda}\}$. Then Theorem 4.8 tells us that $\bar{x}$ is a strict local minimizer of (1.2) and that $\bar{\lambda}$ is a noncritical multiplier of (1.3) with $\Psi = \nabla_x L$ corresponding to $\bar{x}$, and thus (iii) is satisfied.

Suppose next that all the conditions in (iii) hold. Since $\bar{\lambda}$ is noncritical, we derive the semi-isolated calmness of S_{KKT} at $\big((0,0),(\bar{x},\bar{\lambda})\big)$. This together with $\Lambda_{\rm com}(\bar{x}) = \{\bar{\lambda}\}$ results in the existence of a number $\ell \geq 0$ as well as neighborhoods U of $(0,0)$ and V of $(\bar{x},\bar{\lambda})$ such that

$$S_{KKT}(p_1,p_2) \cap V \subset \big\{(\bar{x},\bar{\lambda})\big\} + \ell\,\|(p_1,p_2)\|\,\mathbb{B} \ \text{ for all } \ (p_1,p_2) \in U. \tag{5.8}$$

Thus S_{KKT} enjoys the isolated calmness property at $\big((0,0,(\bar{x},\bar{\lambda}))\big)$, and we arrive at (iv).

To verify the opposite implication (iv)$\Longrightarrow$(iii), let us show that the isolated calmness of S_{KKT} at $\big((0,0),(\bar{x},\bar{\lambda})\big)$ in (iv) yields $\Lambda_{\rm com}(\bar{x}) = \{\bar{\lambda}\}$. Indeed, suppose on the contrary that $\Lambda_{\rm com}(\bar{x})$ is not a singleton. Then there exists $\widehat{\lambda} \in \Lambda_{\rm com}(\bar{x})$ with $\widehat{\lambda} \neq \bar{\lambda}$. Since the set $\Lambda_{\rm com}(\bar{x})$ is convex, every point of the line segment connecting $\bar{\lambda}$ and $\widehat{\lambda}$ belongs to $\Lambda_{\rm com}(\bar{x})$. The isolated calmness of S_{KKT} at $\big((0,0),(\bar{x},\bar{\lambda})\big)$ amounts to (5.8), and hence we can find

$\lambda' \neq \bar{\lambda}$ with $\lambda' \in \Lambda_{\rm com}(\bar{x})$ and such that λ' is sufficiently close to $\bar{\lambda}$, i.e., $(\bar{x}, \lambda') \in V$. Then it follows from (5.8) that

$$\|\lambda' - \bar{\lambda}\| \leq \ell \cdot 0 = 0,$$

which yields $\lambda' = \bar{\lambda}$, a contradiction ensuring that $\Lambda_{\rm com}(\bar{x})$ is a singleton. Theorem 4.7 tells us that $\bar{\lambda}$ is a noncritical multiplier of (1.3) corresponding to $\bar{x}$, and thus (iii) holds.

Next we verify implication (iii)$\Longrightarrow$(ii). Let us first deduce from $\Lambda_{\rm com}(\bar{x}) = \{\bar{\lambda}\}$ in (iii) that the qualification condition (5.3) in (ii) is satisfied. Supposing the contrary, find a normal $v \in N_{{\rm dom}\, \theta_{Y,B}}(\Phi(\bar{x}))$ with $v \neq 0$ such that $\nabla\Phi(\bar{x})^* v = 0$. Letting $\lambda' := \bar{\lambda} + v$, we get $\lambda' \neq \bar{\lambda}$ and $\nabla_x L(\bar{x}, \lambda') = 0$ for the Lagrangian function (4.33). By the choice of v and the normal cone definition (2.3) we get from the above that

$$\langle \lambda', z - \Phi(\bar{x}) \rangle \leq \theta_{Y,B}(z) - \theta_{Y,B}(\Phi(\bar{x})) \ \text{ for all } \ z \in {\rm dom}\, \theta_{Y,B},$$

which shows that $\lambda' \in \partial\theta_{Y,B}(\Phi(\bar{x}))$ and hence $\lambda' \in \Lambda_{\rm com}(\bar{x})$ due to $\nabla_x L(\bar{x}, \lambda') = 0$. Since $\lambda' \neq \bar{v}$, it gives us a contradiction with the assumption of $\Lambda_{\rm com}(\bar{x}) = \{\bar{\lambda}\}$ in (iii) and thus justifies the validity of the qualification condition (5.3). Employing now Lemma 5.1 tells us that the second-order *necessary* optimality condition (5.4) is satisfied.

To finish the verification of (iii)$\Longrightarrow$(ii), we need to prove that the second-order *sufficient* optimality condition (4.37) holds under the assumptions in (iii). Supposing the contrary gives us a nonzero element $\xi_0 \in \{w |\ \nabla\Phi(\bar{x})w \in \mathcal{K}^* + {\rm rge}\, B\}$ such that

$$\big\langle \nabla^2_{xx} L(\bar{x}, \bar{\lambda})\xi_0, \xi_0 \big\rangle + 2\theta_{\mathcal{K},B}\big(\nabla\Phi(\bar{x})\xi_0\big) \leq 0.$$

Since $\Lambda_{\rm com}(\bar{x}) = \{\bar{\lambda}\}$, it is easy to see that the second-order necessary condition (5.4) can be equivalently written as

$$\big\langle \nabla^2_{xx} L(\bar{x}, \bar{\lambda})w, w \big\rangle + 2\theta_{\mathcal{K},B}\big(\nabla\Phi(\bar{x})w\big) \geq 0 \quad \text{for all } w \in \mathbb{R}^n \ \text{ with } \ \nabla\Phi(\bar{x})w \in {\rm dom}\, \theta_{\mathcal{K},B}.$$

Furthermore, employing the equalities

$$\nabla\Phi(\bar{x})\xi_0 \in \mathcal{K}^* + {\rm rge}\, B = \big(\mathcal{K} \cap \ker B\big)^* = {\rm dom}\, \theta_{\mathcal{K},B}$$

allows us to deduce from the equivalent form of the second-order necessary condition that

$$\big\langle \nabla^2_{xx} L(\bar{x},\bar{\lambda})\xi_0,\xi_0\big\rangle + 2\theta_{\mathcal{K},B}\big(\nabla\Phi(\bar{x})\xi_0\big)=0.$$

This in turn implies that the vector ξ_0 is an *optimal solution* to the problem

$$\min_{\xi\in\mathbb{R}^n}\ \frac{1}{2}\big\langle \nabla^2_{xx} L(\bar{x},\bar{\lambda})\xi,\xi\big\rangle + \theta_{\mathcal{K},B}\big(\nabla\Phi(\bar{x})\xi\big).$$

Applying the subdifferential Fermat rule to the latter problem and then using the elementary sum rule for convex subgradients together with the chain rule from [37, Exercise 10.22(b)] yield

$$\begin{aligned} 0 &\in \nabla^2_{xx} L(\bar{x},\bar{\lambda})\xi_0 + \nabla\Phi(\bar{x})^*\partial\theta_{\mathcal{K},B}(\nabla\Phi(\bar{x})\xi_0) \\ &= \nabla^2_{xx} L(\bar{x},\bar{\lambda})\xi_0 + \nabla\Phi(\bar{x})^* D\partial\theta_{Y,B}(\Phi(\bar{x}),\bar{\lambda})(\nabla\Phi(\bar{x})\xi_0), \end{aligned}$$

where the last equality comes from (2.15). Since $\xi_0\neq 0$, it shows by Definition 4.1 that $\bar{\lambda}$ is a critical multiplier. This contradicts the assumption in (iii) that $\bar{\lambda}$ is a noncritical multiplier and therefore verifies the validity of (4.37) and the entire implication (iii)$\Longrightarrow$(ii).

Our next step is to prove implication (i)$\Longrightarrow$(iv), which clearly holds. To complete the proof of the theorem, it remains to verify implication (iv)$\Longrightarrow$(i). To achieve this implication, we only need to show that there are neighborhoods U of $(0,0)$ and V of $(\bar{x},\bar{\lambda})$ such that $S_{KKT}(p_1,p_2)\cap V\neq\emptyset$ for all $(p_1,p_2)\in U$. To this end, define the set-valued mapping $Q\colon \mathbb{R}^m\rightrightarrows\mathbb{R}^n$ by

$$Q(p):=\big\{x\in\mathbb{R}^n\big|\ \Phi(x)+p\in\operatorname{dom}\theta_{Y,B}\big\},\quad p\in\mathbb{R}^n.$$

Having already proved (iv) and (iii) are equivalent, we have the qualification condition (5.3) because of the assumptions in (iii). As proved above, (iii) and (ii) are equivalent. Thus the second-order sufficient condition (4.37) is satisfied and implies by Theorem 4.8 that $\bar{x}$ is a strict local minimizer for (1.2). This gives a neighborhood O of $\bar{x}$ for which we have

$$\varphi_0(\bar{x})+\theta_{Y,B}\big(\Phi(\bar{x})\big)<\varphi_0(x)+\theta_{Y,B}\big(\Phi(x)\big)\quad\text{for all}\quad x\in O. \tag{5.9}$$

Applying [20, Theorem 4.37(ii)] to the mapping Q with the initial point $(0,\bar{x})$ gives us numbers $r>0$ and $\ell\geq 0$ such that

$$Q(p)\cap\mathbb{B}_r(\bar{x})\subset Q(p')+\ell\,\|p-p'\|\,\mathbb{B}\ \text{ for all }\ p,p'\in\mathbb{B}_r(0), \tag{5.10}$$

where r can be chosen such that $\mathbb{B}_r(\bar{x})\subset O$. Consider now the optimization problem

$$\text{minimize }\ \varphi_0(x)+\theta_{Y,B}\big(\Phi(x)+p_2\big)-\langle p_1,x\rangle\ \text{ subject to }\ x\in\mathbb{B}_r(\bar{x})\cap Q(p_2). \tag{5.11}$$

It is clear that this problem admits an optimal solution $x_{p_1p_2}$ for any pair $(p_1,p_2)\in\mathbb{R}^n\times\mathbb{B}_r(0)$ since the cost function therein is lower semicontinuous while the constraint set is obviously compact. Let us now show that there is a number $\varepsilon>0$ with $\mathbb{B}_\varepsilon(0,0)$ such that

$$x_{p_1p_2}\in\operatorname{int}\mathbb{B}_r(\bar{x})\ \text{ for any }\ (p_1,p_2)\in\mathbb{B}_\varepsilon(0,0). \tag{5.12}$$

Suppose the contrary and then find sequences $(p_{1k},p_{2k})\to(0,0)$ and $x_{p_{1k}p_{2k}}$ for which $\|x_{p_{1k}p_{2k}}-\bar{x}\|=r$. We get without loss of generality that $x_{p_{1k}p_{2k}}\to x_0$ as $k\to\infty$ and so $\|x_0-\bar{x}\|=r$. This yields $x_0\neq\bar{x}$. Since $x_{p_{1k}p_{2k}}$ is an optimal solution to (5.11), it follows that

$$\varphi_0(x_{p_{1k}p_{2k}})+\theta_{Y,B}\big(\Phi(x_{p_{1k}p_{2k}})+p_{2k}\big)-\langle p_{1k},x_{p_{1k}p_{2k}}\rangle\leq\varphi_0(x)+\theta_{Y,B}\big(\Phi(x)+p_{2k}\big)-\langle p_{1k},x\rangle \tag{5.13}$$

for all $x\in\mathbb{B}_r(\bar{x})\cap Q(p_{2k})$. Pick any $x\in\mathbb{B}_{\frac{r}{2}}(\bar{x})\cap Q(0)$ and $k\in\mathbb{N}$ so large that $p_{2k}\in\alpha\mathbb{B}$ with $\alpha<\min\{\frac{r}{2\ell},r\}$. It follows from (5.10) that there exist $x'\in Q(p_{2k})$ and $b\in\mathbb{B}$ satisfying

$$\|x'-\bar{x}\|\leq\|x-\bar{x}\|+\ell\,\|p_{2k}\|\leq\frac{r}{2}+\ell\frac{r}{2\ell}=r,\ \ \text{where }\ x:=x'+\ell\,\|p_{2k}\|\,b.$$

Thus $x'\in\mathbb{B}_r(\bar{x})\cap Q(p_{2k})$, and it follows from (5.13) that

$$\begin{aligned}\varphi_0(x_{p_{1k}p_{2k}})\quad&+\theta_{Y,B}\big(\Phi(x_{p_{1k}p_{2k}})+p_{2k}\big)-\langle p_{1k},x_{p_{1k}p_{2k}}\rangle\leq\varphi_0\big(x-\ell\,\|p_{2k}\|\,b\big)\\&+\theta_{Y,B}\big(\Phi(x-\ell\,\|p_{2k}\|\,b)+p_{2k}\big)-\langle p_{1k},x-\ell\,\|p_{2k}\|\,b\rangle.\end{aligned}$$

Passing to the limit at the latter inequality as $k\to\infty$ gives us the estimate

$$\varphi_0(x_0)+\theta_{Y,B}\big(\Phi(x_0)\big)\leq\varphi_0(x)+\theta_{Y,B}\big(\Phi(x)\big),$$

which holds for all $x \in \mathbb{B}_{\frac{r}{2}}(\bar{x}) \cap Q(0)$. In particular, we have

$$\varphi_0(x_0) + \theta_{Y,B}\big(\Phi(x_0)\big) \le \varphi_0(\bar{x}) + \theta_{Y,B}\big(\Phi(\bar{x})\big), \tag{5.14}$$

which contradicts (5.9) since $x_0 \neq \bar{x}$ and $x_0 \in \mathbb{B}_r(\bar{x}) \subset O$, and thus we arrive at (5.12).

At the last step of the proof, denote by $\Lambda_{\rm com}(x_{p_1p_2})$ be the set of Lagrange multipliers associated with the optimal solution $x_{p_1p_2}$ to problem (5.11). It follows from the validity of the qualification condition (5.3) and its robustness with respect to perturbations of the initial point that this qualification condition is also satisfied for the perturbed problem (5.11). This implies in turn that $\Lambda_{\rm com}(x_{p_1p_2}) \neq \emptyset$ for all (p_1, p_2) sufficiently close to $(0,0) \in \mathbb{R}^n \times \mathbb{R}^m$. Assume without loss of generality that $\Lambda_{\rm com}(x_{p_1p_2}) \neq \emptyset$ for all $(p_1, p_2) \in \mathbb{B}_\varepsilon(0,0)$, where ε is taken from (5.12). Using a similar argument as (4.23) and (4.29) via the Hoffman lemma gives us a constant $\ell' \ge 0$ such that for any $(p_1, p_2) \in \mathbb{B}_\varepsilon(0,0)$ and any $\lambda_{p_1p_2} \in \Lambda_{\rm com}(x_{p_1p_2})$ we have

$$\|\lambda_{p_1p_2} - \bar{\lambda}\| = {\rm dist}\big(\lambda_{p_1p_2}; \Lambda_{\rm com}(\bar{x})\big) \le \ell'\big(\, \|x_{p_1p_2} - \bar{x}\| + \|p_1\| + \|p_2\|\,\big).$$

This clearly proves the existence of a neighborhood V of $(\bar{x}, \bar{\lambda})$ such that $S_{KKT}(p_1,p_2) \cap V \neq \emptyset$ for all $(p_1, p_2) \in \mathbb{B}_\varepsilon(0,0)$ and so finishes the proof of implication (iv)$\Longrightarrow$(i). $\square$

The final piece of this thesis concerns yet another well-recognized Lipschitzian type property, which seems to be the most natural extension of *robust* Lipschitzian behavior to set-valued mapping. For this reason we label it as the Lipschitz-like property [20] while it is also known as the pseudo-Lipschitz or Aubin one. It is said that a set-valued mapping/multifunction $F\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is *Lipschitz-like* around $(\bar{x}, \bar{y}) \in {\rm gph}\, F$ if there exists a constant $\ell \ge 0$ together with neighborhoods U of $\bar{x}$ and V of $\bar{y}$ such that we have the inclusion

$$F(x') \cap V \subset F(x) + \ell\|x - x'\|\mathbb{B} \;\text{ for all }\; x, x' \in U. \tag{5.15}$$

To formulate a convenient characterization of property (5.15), we recall first the notion of the *normal cone* to a set $\Omega \subset \mathbb{R}^n$ at a point $\bar{x} \in \Omega$ defined by

$$N_\Omega(\bar{x}) := \Big\{ v \in \mathbb{R}^n \Big| \text{ there exist } x_k \overset{\Omega}{\to} \bar{x},\ v_k \to v \text{ with } \limsup_{x \to x_k} \frac{\langle v_k, x - x_k \rangle}{\|x_k - x\|} \le 0 \Big\}.$$

The *coderivative* of a set-valued mapping $F\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ at $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ is given by

$$D^*F(\bar{x}, \bar{y})(v) := \big\{ u \in \mathbb{R}^n \big| (u, -v) \in N_{\operatorname{gph} F}(\bar{x}, \bar{y}) \big\}, \quad v \in \mathbb{R}^m.$$

The following characterization of the Lipschitz-like property for any closed-graph mapping $F\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ around $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ is known as the *Mordukhovich criterion* from [37, Theorem 9.40], where the proof is different from the original one; see [18, Theorem 5.7] as well as its infinite-dimensional extension given in [20, Theorem 4.10]:

$$D^*F(\bar{x}, \bar{y})(0) = \{0\}. \tag{5.16}$$

Note the results obtained therein provide also a precise computation of the *exact bound*/infimum of Lipschitzian moduli $\{\ell\}$ in (5.15) via the coderivative norm at $(\bar{x}, \bar{y})$.

Full *coderivative calculus* developed for coderivatives, which is based on variational/extremal principles of variational analysis and can be found in [22, 20], allows us apply the general characterization (5.16) to specific multifunctions given in some structural forms. The next theorem employs (5.16) and coderivative calculus to characterize the Lipschitz-like property of the solution map (5.7) to the canonically perturbed KKT system (5.6).

Theorem 5.4. (Lipschitz-like property of solution maps). *Let $(\bar{x}, \bar{\lambda}) \in S_{KKT}(0,0)$ for the solution map S_{KKT} defined in* (5.7) *with θ taken from* (1.1). *Then S_{KKT} is Lipschitz-like around $\big((0,0), (\bar{x}, \bar{\lambda})\big)$ if and only if we have the implication*

$$\begin{cases} \nabla^2_{xx} L(\bar{x}, \bar{\lambda})\xi + \nabla\Phi(\bar{x})^*\eta = 0 \\ \eta \in \big(D^*\partial\theta_{Y,B}\big)(\Phi(\bar{x}), \bar{\lambda})\big(\nabla\Phi(\bar{x})\xi\big) \end{cases} \Longrightarrow (\xi, \eta) = (0,0). \tag{5.17}$$

Proof. Consider the mapping G from (4.12) with $\Psi = \nabla_x L$. We easily deduce from the coderivative definition and the form of S that

$$(\xi, \eta) \in D^*S_{KKT}\big((0,0), (\bar{x}, \bar{\lambda})\big)(w_1, w_2) \Longleftrightarrow -(w_1, w_2) \in D^*G\big((\bar{x}, \bar{\lambda}), (0,0)\big)(-\xi, -\eta) \tag{5.18}$$

for all $(\xi,\eta)\in\mathbb{R}^n\times\mathbb{R}^m$ and $(w_1,w_2)\in\mathbb{R}^n\times\mathbb{R}^m$. Using the structure of G and employing the coderivative sum rule in the equality form from [22, Theorem 3.9] yield

$$\begin{aligned} D^*G\big((\bar{x},\bar{\lambda}),(0,0)\big)(\xi,\eta) &= \begin{bmatrix} \nabla^2_{xx}L(\bar{x},\bar{\lambda}) & -\nabla\Phi(\bar{x})^* \\ \nabla\Phi(x) & 0 \end{bmatrix}\begin{bmatrix} \xi \\ \eta \end{bmatrix}+\begin{bmatrix} 0 \\ D^*(\partial\theta_{Y,B})^{-1}(\bar{\lambda},\Phi(\bar{x}))(\eta) \end{bmatrix} \\ &= \begin{bmatrix} \nabla^2_{xx}L(\bar{x},\bar{\lambda})\xi-\nabla\Phi(\bar{x})^*\eta \\ \nabla\Phi(x)\xi+D^*(\partial\theta_{Y,B})^{-1}(\bar{\lambda},\Phi(\bar{x}))(\eta) \end{bmatrix}. \end{aligned} \tag{5.19}$$

It follows from (5.18) and the coderivative criterion (5.15) that S_{KKT} is Lipschitz-like around $\big((0,0),(\bar{x},\bar{\lambda})\big)$ if and only if we have the implication

$$(0,0)\in D^*G\big((\bar{x},\bar{\lambda}),(0,0)\big)(\xi,\eta)\Longrightarrow(\xi,\eta)=(0,0),$$

which leads us together the coderivative representation for G in (5.19) to characterization (5.17) of the Lipschitz-like property of the solution map S_{KKT}. □

Combining finally the obtained characterization of the Lipschitz-like property in Theorem 5.4 with some known facts of variational analysis allows us to reveal a relationship between the latter property of the solution map S_{KKT} and its isolated calmness at the same point.

Theorem 5.5. (Lipschitz-like property of solution maps implies their isolated calmness). *Let S_{KKT} be the solution map* (5.7) *of the canonically perturbed KKT system* (5.6) *with the piecewise linear-quadratic term* (1.1)*, and let $(\bar{x},\bar{\lambda})\in S_{KKT}(0,0)$. If S_{KKT} is Lipschitz-like around $\big((0,0),(\bar{x},\bar{\lambda})\big)$, then it enjoys the isolated calmness property at this point.*

Proof. Assuming that S_{KKT} has the Lipschitz-like property around $\big((0,0),(\bar{x},\bar{\lambda})\big)$, we get implication (5.17) by Theorem 5.4. On the other hand, we proceed similarly to the proof of Theorem 5.4 and get counterparts of the equalities in (5.18) and (5.19) with replacing the coderivative by the graphical derivative therein. The latter one is due to the easily checkable sum rule for graphical derivatives of summations with one smooth term as in (4.12).

Having this, we apply the Levy-Rockafellar criterion of isolated calmness (4.11) to the solution map (5.7) and thus conclude that the isolated calmness of S_{KKT} at $\big((0,0),(\bar{x},\bar{\lambda})\big)$ is equivalent to

$$\begin{cases} \nabla^2_{xx}L(\bar{x},\bar{\lambda})\xi + \nabla\Phi(\bar{x})^*\eta = 0 \\ \eta \in \big(D\partial\theta_{Y,B}\big)(\Phi(\bar{x}),\bar{\lambda})\big(\nabla\Phi(\bar{x})\xi\big) \end{cases} \Longrightarrow (\xi,\eta) = (0,0). \tag{5.20}$$

Comparing (5.17) and (5.20), we see that the only difference is in terms involving $(D^*\partial\theta_{Y,B})(\Phi(\bar{x}),\bar{\lambda})$ and $(D\partial\theta_{Y,B})(\Phi(\bar{x}),\bar{\lambda})$. To this end we use the derivative-coderivative relationship from [37,Theorem 13.57], which tells us that the inclusion

$$(D\partial\theta_{Y,B})(\Phi(\bar{x}),\bar{\lambda})(u) \subset (D^*\partial\theta_{Y,B})(\Phi(\bar{x}),\bar{\lambda})(u) \ \text{ for all } \ u \in \mathbb{R}^m$$

holds under the assumptions that are automatically satisfied for the piecewise linear-quadratic function $\theta_{Y,B}$ from (1.1). This therefore completes the proof of the theorem. $\square$

REFERENCES

[1] J. F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems*, Springer, New York, 2000.

[2] C. Ding, D. Sun and L. Zhang, Characterization of the robust isolated calmness for a class of conic programming problems, *SIAM J. Optim. SIAM J. Optim.* **27** (2017), 67–90.

[3] Hong Do, Boris S. Mordukhovich and M. Ebrahim Sarabi, *Criticality of Lagrange Multipliers in Extended Nonlinear Optimization*, to appear in Optimization; arXiv:1901.01469.

[4] Hong Do, Boris S. Mordukhovich and M. Ebrahim Sarabi, *Complete characterizations of full stability in extended nonlinear programming*; final step of preparation.

[5] A. L. Dontchev and R. T. Rockafellar, *Characterizations of strong regularity for variational inequalities over polyhedral sets*, SIAM J. Optim., 6 (1996), pp. 1087-1105.

[6] A. L. Dontchev and R. T. Rockafellar, *Implicit Functions and Solution Mappings: A View from Variational Analysis*, 2nd edition, Springer, New York, 2014.

[7] F. Facchinei and J.-S. Pang, *Finite-Dimesional Variational Inequalities and Complementarity Problems*, Springer, New York, 2003.

[8] R. Henrion, A. Y. Kruger, and J. V. Outrata, *Some remarks on stability of generalized equations*, J. Optim. Theory Appl. 159 (2013), 681âĂŞ697.

[9] A. F. Izmailov, On the analytical and numerical stability of critical Lagrange multipliers, *Comput. Math. Math. Phys.* **45** (2005), 930–946.

[10] A. F. Izmailov and M. V. Solodov, *Newton-Type Methods for Optimization and Variational Problems*, Springer, New York, 2014.

[11] A. F. Izmailov and M. V. Solodov, Critical Lagrange multipliers: what we currently know about them, how they spoil our lives, and what we can do about it, *TOP* **23** (2015), 1–26.

[12] D. Kinderlehrer and G. Stampacchia, *An Introduction to Variational Inequalities and Their Applications*, SIAM, Philadelphia, PA, 2000.

[13] M. Kojima, *Strongly stable stationary solutions in nonlinear programs*, in Analysis and Computation of Fixed Points, S. M. Robinson (ed.), pp. 93âĂŞ138, Academic Press, New York, 1980.

[14] J. Kyparisis, *Sensitivity analysis for nonlinear programs and variational inequalities with nonunique multipliers*, Math. Oper. Res. 15 (1990), 286âĂŞ298.

[15] A. B. Levy, R. A. Poliquin and R. T. Rockafellar, Stability of locally optimal solutions, *SIAM J. Optim.* **10** (2000), 580–604.

[16] A. S. Lewis and S. Zhang, *Partial smoothness, tilt stability, and generalized Hessians*, SIAM J. Optim., 23 (2013), pp.74-94.

[17] S. Lu and S. M. Robinson, *Variational inequalities over perturbed polyhedral convex sets*, Math. Oper. Res. 33 (2008), 689âĂŞ711.

[18] B. S. Mordukhovich, Complete characterizations of openness, metric regularity, and Lipschitzian properties of multifunctions, *Trans. Amer. Math. Soc.* **340** (1993), 1-35.

[19] B. S. Mordukhovich, *Sensitivity analysis in nonsmooth optimization*, in Theoretical Aspects of Industrial Design, D. A. Field and V. Komkov, eds., SIAM, Philadelphia, 1992, pp. 32-46.

[20] B. S. Mordukhovich, *Variational Analysis and Generalized Differentiation, I: Basic Theory*, Springer, Berlin, 2006.

[21] B. S. Mordukhovich, Comments on: Critical Lagrange multipliers: what we currently know about them, how they spoil our lives, and what we can do about it, *TOP* **23** (2015), 35–42.

[22] B. S. Mordukhovich, *Variational Analysis and Applications*, Springer, Cham, Switzerland, 2018.

[23] B. S. Mordukhovich, N. M. Nam and N. T. Y. Nhi, *Partial second-order subdifferentials in variational analysis and optimization*, Numer. Func. Anal. Optim. 35 (2014), 1113âĂŞ1151.

[24] B. S. Mordukhovich and J. V. Outrata, *Tilt stability in nonlinear programming under Mangasarian-Fromovitz constraint qualification*, Kybernetika, 49 (2013), pp. 446-464.

[25] B. S. Mordukhovich and T. T. A. Nghia, *Second-order variational analysis and characterizations of tilt-stable optimal solutions in infinite-dimensional spaces*, Nonlinear Anal. 86 (2013), 159âĂŞ180.

[26] B. S. Mordukhovich and T. T. A. Nghia, *Full Lipschitzian and HÂÍolderian stability in optimization with applications to mathematical programming and optimal control*, SIAM J. Optim. 24 (2014), 1344âĂŞ1381.

[27] B. S. Mordukhovich and T. T. A. Nghia, *Second-order characterizations of tilt stability with applications to nonlinear programming*, Math. Program. 149 (2015), 83âĂŞ104.

[28] B. S. Mordukhovich and T. T. A. Nghia, *Local monotonicity and full stability for parametric variational systems*, SIAM J. Optim. 26 (2016), 1032âĂŞ1059.

[29] B. S. Mordukhovich, T. T. A. Nghia and R. T. Rockafellar, Full stability in finite-dimensional optimization, *Math. Oper. Res.* **40** (2015), 226–252.

[30] B. S. Mordukhovich and R. T. Rockafellar, Second-order subdifferential calculus with applications to tilt stability in optimization, *SIAM J. Optim.* **22** (2012), 953–986.

[31] B. S. Mordukhovich, R. T. Rockafellar and M. E. Sarabi, Characterizations of full stability in constrained optimization, *SIAM J. Optim.* **23** (2013), 1810–1849.

[32] B. S. Mordukhovich and M. E. Sarabi, Critical multipliers in variational systems via second-order generalized differentiation, *Math. Program.* **169** (2018), 605–648.

[33] B. Mordukhovich and M. E. Sarabi, Criticality of Lagrange multipliers in variational systems, *SIAM J. Optim.*, to appear, https://arxiv.org/pdf/1808.03802.pdf.

[34] R. A. Poliquin and R. T. Rockafellar, *Tilt stability of a local minimum*, SIAM J. Optim., 8 (1998), pp. 287-299.

[35] R. A. Poliquin and R. T. Rockafellar, *A calculus of prox-regularity*, J. Convex Anal., 17 (2010), pp. 203-210.

[36] R. T. Rockafellar, Extended nonlinear programming, in *Nonlinear Optimization and Related Topics*, G. Di Pillo and F. Giannessi (eds.), *Applied Optimization*, Vol. 36, pp. 381–399, Kluwer Academic Publishers, Dordrecht, 2000.

[37] R. T. Rockafellar and R. J-B. Wets, *Variational Analysis*, Springer, Berlin, 2006.

ABSTRACT

SECOND-ORDER GENERALIZED DIFFERENTIATION OF PIECEWISE LINEAR-QUADRATIC FUNCTIONS AND ITS APPLICATIONS

by

HONG DO

May 2019

Advisor: Boris Mordukhovich

Major: Mathematics

Degree: Doctor of Philosophy

The area of second-order variational analysis has been rapidly developing during the recent years with many important applications in optimization. This dissertation is devoted to the study and applications of the second-order generalized differentiation of a remarkable class of convex extended-real-valued functions that is highly important in many aspects of nonlinear and variational analysis, specifically those related to optimization and stability.

The first goal of this dissertation is to compute the second-order subdifferential of the functions described above, which will be applied in the study of the stability of composite optimization problems associated with piecewise linear-quadratic functions, known as extended nonlinear programming (ENLPs). In the second part of the dissertation, the multiplier criticality is studied systematically for variational systems of composite optimization problems with applications to KKT systems in ENLPs.

AUTOBIOGRAPHICAL STATEMENT

Hong Do was born in Hanoi, Vietnam. She did her undergraduate work at the Faculty of Mathematics and Mechanics of the St. Petersburg State University, Russian Federation. She received her Specialist Degree in Mathematics in 2008. After that she returned her country Vietnam and began working as an instructor at the Faculty of Mathematics, Mechanics and Informatics, VNU University of Science, until summer 2014. During 2011-2013 she also taught at the HUS High School for Gifted Students. In August 2014, she entered the PhD program in Mathematics at the Wayne State University.