Springer Monographs in Mathematics

Boris S. Mordukhovich

Variational Analysis and Applications

Springer

Springer Monographs in Mathematics

This series publishes advanced monographs giving well-written presentations of the "state-of-the-art" in fields of mathematical research that have acquired the maturity needed for such a treatment. They are sufficiently self-contained to be accessible to more than just the intimate specialists of the subject, and sufficiently comprehensive to remain valuable references for many years. Besides the current state of knowledge in its field, an SMM volume should ideally describe its relevance to and interaction with neighbouring fields of mathematics, and give pointers to future directions of research.

More information about this series at http://www.springer.com/series/3733

Boris S. Mordukhovich

Variational Analysis and Applications

Springer

Boris S. Mordukhovich
Department of Mathematics
Wayne State University
Detroit, MI, USA

ISSN 1439-7382 ISSN 2196-9922 (electronic)
Springer Monographs in Mathematics
ISBN 978-3-319-92773-2 ISBN 978-3-319-92775-6 (eBook)
https://doi.org/10.1007/978-3-319-92775-6

Library of Congress Control Number: 2018944139

Mathematics Subject Classification: 49J40, 49J52, 49J53, 90C29, 90C30, 90C31, 90C34

This Springer imprint is published by Springer Nature, under the registered company Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

To my beautiful daughters, Lena and Irina,
and also to all my students and collaborators

Preface

All the truths are easy to understand once they are discovered; the point is to discover them.

Galileo Galilei

A moment of truth in mathematics is an instant between infinity when it was considered to be wrong and another infinity when it is considered to be trivial.

Henri Poincaré

Variational analysis, as now understood, is a relatively young area of mathematics. From one side, it can be viewed as an outgrowth of the calculus of variations, constrained optimization, and optimal control, and also of variational principles in mathematical physics and mechanics that go back to the 18th century. On the other hand, modern variational principles and techniques are largely based on perturbations, approximations, and the (unavoidable) usage of generalized differentiation. All of this requires developing new forms of analysis and thus manifests the creation of a new discipline in mathematics that strongly combines and unifies analytic and geometric ideas.

Although some particular aspects of variational analysis have been reflected in the monographic literature earlier (beginning with its starting point–beautiful convex analysis), the first systematic monograph on this subject covering its key ingredients in finite-dimensional spaces was the book by Rockafellar and Wets "Variational Analysis" (Springer, 1998), where this very name was coined. Since then a great many publications have appeared on numerous issues of variational analysis and its applications, including several monographs. Among them is the two-volume monograph by the author "Variational Analysis and Generalized Differentiation, I: Basic Theory, II: Applications" (Springer, 2006) devoted to infinite-dimensional aspects of variational analysis and generalized differentiation with a broad spectrum of applications.

This new book presented to the reader's attention pursues several goals. The *first goal* is to give a systematic and easily understandable exposition of the key concepts and facts of variational analysis with selected applications in finite-dimensional spaces. It is done in Chapters 1–6 that also contain, besides basic material, some recent developments in this vein. We view these chapters as the basis for a self-contained course on variational analysis for beginners, which is accessible to graduate students in mathematics as well as in applied sciences and engineering. To create a *usable text for teaching* variational analysis, with plenty of exercises as well as illustrative figures and examples, is one of the underlying goals of this book.

Here we follow a *dual-space approach*, which doesn't rely on tangential approximations of sets and related constructions for functions and mappings in primal spaces, but instead focuses directly on dual-space approximations, which are *dual to none*. One of the reasons for this is that duality objects generated by any tangential approximation are automatically convex while the latter property provides significant limitations for generalized differentiation and its applications. This issue is revealed and largely discussed in the basic text and commentaries therein. On the other hand, dealing with nonconvex objects requires the usage of new machinery of analysis that is different from the conventional convex separation and the like. The major tool for such analysis is provided by a geometric variational principle known as the *extremal principle* for set systems, which is strongly employed in the book.

This approach leads us to developing an easy path to variational analysis and its applications presented in the book. The finite-dimensional framework allows us to significantly simplify the exposition and proofs of major results. It has been revealed that dual-space objects are actually more beautiful and perfect than their primal-space counterparts and bring us, as a rule, to more natural and complete results. One can observe an analogy with Plato's theory of *Forms* (or Ideas, eidos), which are dual objects to some extent while providing the most accurate representations of reality in the *intelligible realm*.

Yet *another goal* of this book is encouraging the interested readers to learn more on variational analysis and to develop their research skills in this field by performing (at least partly) the exercises presented after the basic material of each chapter. The reader can find hints and references for more difficult exercises and also discussions on challenging open questions in the commentaries. A number of exercises deal with problems in infinite-dimensional spaces (while presenting the corresponding definitions and supporting material), and some of them are referred to in the subsequent chapters of the book.

Chapters 7–10 are devoted concern recent results on applications of variational analysis to important classes of advanced problems in optimization, microeconomics, and related areas. They are presented in full generality of infinite-dimensional spaces and mostly address researchers, graduate students, and practitioners in these (fairly broad) particular fields while may be of interest for larger communities of mathematicians and economists. The results obtained demonstrate the strength of variational analysis and dual-space constructions in solving concrete problems that may not even be of a variational nature.

Let us briefly describe the main content of each of the ten chapters.

Chapter 1 presents the basic constructions of *first-order generalized differentiation* studied and applied in the book. Developing a *geometric* approach to generalized differentiation, we consider first the nonconvex (basic, limiting) *normal cone* to locally closed sets and then define in its terms the *coderivative* of set-valued mappings as well as the *basic* and *singular subdifferentials* of extended-real-valued functions. Various representations and properties of these constructions and their relationships with other objects of generalized differentiability in variational analysis (including tangentially generated ones) are investigated in detail. The given proofs are mostly simplifications and improvements in finite dimensions of those developed in the author's previous book [522] in more general settings. Some new results and proofs are also presented here. Infinite-dimensional extensions and related developments are discussed in the exercise and commentary sections.

The main material of *Chapter 2* concerns *extremal principles* for finitely many and countably many systems of sets, which play a crucial role in the developed dual-space geometric approach to variational analysis and generalized differentiation. Our major extremal principle is expressed in terms of basic normals to finitely many closed sets and can be considered as a *nonconvex variational* counterpart of the classical convex separation. Its proof is given by using the *method of metric approximations* (MMA), which manifests one of the most fundamental ideas of modern variational analysis to implement approximation, perturbation, and limiting procedures. The basic extremal principle and its infinite-dimensional versions (discussed in exercises and commentaries) are strongly employed in all the book chapters. In Chapter 2 this is applied to derive the major *normal cone intersection* and *subdifferential sum rules*. We also present here more recent results concerning extensions of the extremal principle to *countable systems* of sets, which seem to be attractive for their own sake and various applications while being motivated by problems of semi-infinite programming considered in the subsequent chapters. The proofs of the countable versions of the extremal principle are given by using the MMA and reveal some new phenomena even for finitely many closed sets in extremal systems.

Another theme of Chapter 2 concerns *variational principles* for extended-real-valued functions that are different from but somewhat related to extremal principles for sets in both finite and infinite dimensions. The finite-dimensional geometry allows us to derive a general variational principle, which is simple to prove, useful in applications, and contains known versions of such results in finite dimensions. Infinite-dimensional extensions and relationships with lower and upper subdifferential principles for extended-real-valued functions are discussed in the exercise and commentary sections of this chapter.

In *Chapter 3* we combine the study of two major topics of variational analysis, which seem not to be connected at the first glance but actually occur to be deeply interrelated. They concern the main *well-posedness* properties of set-valued mappings (Lipschitzian stability, metric regularity, and linear openness/covering) and their coderivative characterizations—from one side, and a comprehensive *coderivative calculus* from the other. The developed proofs in both directions are based on applying the extremal and variational principles. Furthermore, the usage of coderivative calculus allows us to determine broad classes of parametric variational systems whose solution maps fail to be metrically regular. In the exercise and commentary session of this chapter we discuss, besides infinite-dimensional extensions, a variety of other well-posedness properties useful in variational analysis and its applications and also formulate some challenging open problems in these and related areas.

Chapter 4 is devoted to developing a comprehensive *subdifferential calculus* for both basic and singular limiting subgradients of extended-real-valued functions. A major role is played by evaluating subgradients for a general class of *marginal/optimal value* functions, which is the key for deriving chain, product, quotient, minimum, maximum, and other rules of subdifferential calculus. Another major ingredient of subdifferential calculus highly important in what follows is a variety of *mean value theorems* for nonsmooth functions presented in this chapter together with some impressive applications.

Chapter 5 deals with global and local *monotonicity* of set-valued operators. The importance of such properties has been well recognized in variational analysis, optimization, and numerous applications. There is an enormous amount of publications devoted to these and related topics. Here we present a new view on maximal monotonicity properties by developing their complete *coderivative characterizations* with the usage of machinery of variational analysis and generalized differentiation. The main results are obtained for the notions of *global maximal monotonicity* and *strong local maximal monotonicity*, while we discuss further perspectives, challenging open questions, and formulate several conjectures. Among strong advantages of the obtained characterizations are extensive calculus rules available for coderivatives, which allow us to deal with structural problems and open the gate for further developments. We also discuss in this chapter some related regularity and stability/calmness notions for set-valued mappings, particularly of the subdifferential type.

The first part of *Chapter 6* presents refined necessary optimality conditions for general constrained problems of *nondifferentiable programming* that are expressed in terms of the first-order constructions of generalized differentiation considered in Chapter 1. The obtained optimality conditions are given in both *lower subdifferential* and *upper subdifferential* forms and are derived by direct applications of the extremal and variational principles together with the developed calculus rules. Then we present applications of these results to important classes of *bilevel optimization* problems, which are intrinsically nonsmooth even in the case of smooth initial data. The value function approach allows us to reduce such problems to single-level programs with nonsmooth data and then apply the results obtained above in nondifferentiable programming by using subdifferential rules for marginal/optimal value functions established in Chapter 4. In the exercise and commentary sections of this chapter we discuss other approaches to bilevel programming and draw the reader's attention to unsolved problems in this and related areas.

Chapter 7 is devoted to the systematic application of the underlying constructions and techniques in variational analysis and generalized differentiation to a comprehensive study of *semi-infinite programs* (SIPs) satisfying some *linearity or convexity* assumptions on the problem data. Problems of this type involving infinite linear and convex inequality constraint systems have a long history in optimization theory and applications, especially for systems indexed by compact sets. We show here that the usage of advanced variational techniques, quite recently developed in this area, allows us to offer new viewpoints and derive enhanced results on Lipschitzian stability and optimality conditions for SIPs with arbitrary (in particular, countable) index sets. Furthermore, calculating the basic and singular subgradients of value functions in SIPs with *DC* (difference of convex) objectives leads us to new optimality and stability conditions in DC infinite programs and yields by implementing the value function approach to refined optimality conditions for the class of convex *bilevel SIPs*. Taking into account that SIPs always involve, due to their very essence, infinite dimensionality even in the case of finite-dimensional decision spaces, we present the main material in this and subsequent chapters in general Banach (or Asplund if needed) spaces.

Chapter 8 continues our considerations of SIPs while concentrating on *nonconvex* problems under different assumptions on the functions involved in infinite systems (differentiability, Lipschitz continuity, and lower semicontinuity). Motivated by eventual applications to nonconvex SIPs, various approaches and strategies are tested, which lead us to variational and calculus results of their own importance with large spectra of other applications. We mention here calculations of normals to infinite intersections of nonconvex sets, subdifferentiation of suprema of nonsmooth functions over noncompact index collections, Lipschitzian stability and metric regularity of nonconvex cone-constrained systems, etc. All the results obtained in these directions are quite recent and haven't appeared before in the monographic literature.

Chapter 9 deals with problems of *set* and *set-valued optimization*, which are relatively new in optimization theory and have become particularly attractive to mathematicians, applied scientists, and practitioners during recent years, largely due to practical demands. They are essentially more involved in many aspects in comparison with single-valued vector objectives that are usual in multiobjective optimization. In this chapter we develop a dual-space variational approach to general classes of such problems, which results in establishing existence theorems for Pareto-type optimal solutions and robust necessary optimality conditions for them expressed in terms of coderivatives and novel subdifferential constructions for set-valued mappings with partially ordered values. Our main attention is paid to the so-called *relative Pareto* solutions to multiobjective problems, which unify the conventional efficient and weakly efficient solutions with more flexible notions of set optimality. The basic machinery for the implementation of this approach includes, besides the underlying extremal principle, extended versions of the Ekeland-type and subdifferential variational principles for set-valued mappings. This approach leads us to new results not only for set-valued problems but also for conventional problems of vector optimization in both finite and infinite dimensions.

The final *Chapter 10* concerns applications of the advanced variational and generalized differential techniques presented in this book to *microeconomic modeling*. The main goal is to establish two-sided relationships between models of *welfare economics* and appropriate problems of *set-valued optimization*, and then to study both of them in parallel by using the developed tools and results of variational analysis. This approach occurs to be beneficial in both directions. From one side, it allows us to design a class of set-valued optimization problems and define new types of fully localized solutions to them that correspond to conventional as well as to less understood notions of Pareto-type optimal allocations (set-valuedness of the objective is crucial here!). On the other hand, natural concepts of set optimality for multiobjective problems induce new notions of Pareto-type optimal allocations, which admit adequate economic interpretations. Having the aforementioned equivalence relationships, we apply the developed tools of variational analysis and generalized differentiation, mainly revolved around an appropriate version of the extremal principle, to deriving unified necessary conditions for the corresponding notions of optimal solutions in the designed problems of set-valued optimization, which generate in turn novel versions of the so-called *fundamental second welfare theorem* for marginal price equilibria in nonconvex models of welfare economics with finite- and infinite-dimensional commodity spaces.

It should be emphasized that giving a large number of *exercises* in this book plays a special and highly important role in its design. Besides the standard intention of exercises to help readers in better understanding the basic material, they encourage them to significantly develop research skills and the ability to work independently in the broad areas covered by the book. On the other hand, precise definitions and result formulations in many exercises make this part of the book a handy *reference source* to enormous material available now in the (first-order) state-of-the-art variational analysis and its applications in both finite-dimensional and infinite-dimensional spaces. We also formulate in the exercise sections some *open problems* and *conjectures* and then discuss them in the corresponding commentaries. Such a book design allows us to present here a massive amount of fundamental and newer developments together with further perspectives.

Each chapter of the book ends with an extensive *commentary* section. The main purpose of the commentaries is to emphasize the essence of major results, track the genesis of ideas, provide historical comments, and illuminate challenging open questions and directions of future research from the author's viewpoints. The book includes a large (definitely incomplete) list of references related to the topics and results mentioned in the text that may help the reader in the further study of variational analysis and its applications. For the reader's convenience, we list the titles of all the statements, remarks, and exercises together with the glossary of notation and acronyms as well as an abundant subject index that illustrates the broadness of topics covered by the book and the alternative terminology widely spread in variational analysis and generalized differentiation. In the latter we mostly follow the monographs by Rockafellar and Wets [678] and by the author [522, 523]. The detailed subject index allows the reader to quickly find topics of particular interest and directs him/her, through the commentaries and reference list, to additional sources.

We envision that the book will be useful for large groups of graduate students, researchers, and practitioners in various areas of mathematical sciences, operations research, and applications, particularly to those in economics, mechanics, engineering, and behavior sciences. Our trust is that the book will help the reader to share the author's admiration of the beauty and harmony of variational analysis. We also hope that it will encourage the reader to study more in this exciting area, to employ variational ideas and results in different fields of mathematics and applications, and to get involved in further active research.

Parts of this book have been used by the author in teaching many classes at Wayne State University and other institutions worldwide. The author much appreciates useful feedback that has come from his former and current graduate students over the years. Special thanks go to Truong Bao, Hong Do, Alexander Kruger, Nguyen Mau Nam, Tran Nghia, Dat Pham, Ebrahim Sarabi, and Bingwu Wang. All the figures are made by Nguyen Van Hang, in addition to her great help in reviewing the manuscript. The author is gratefully indebted to three anonymous referees of the initial book submission in June 2017 and two referees of the revision for their helpful remarks and suggestions, which have been fully incorporated into the final version.

A lot of material presented here is based on the author's joint papers with his brilliant collaborators that are listed in the references. Many thanks are owed to each of them. Grants from the National Science Foundation and the Air Force Office of Scientific Research were essential to carry out and complete this project. The author is also very grateful to Elizabeth Loew, Executive Editor in Mathematics of Springer New York, for her attention and excellent pieces of advice in all the stages of preparation and publication of this book.

Above all, the deepest gratitude and love go to my wife Margaret for her enormous support, understanding, and help over the years.

Ann Arbor, MI, USA Boris S. Mordukhovich
January 2018

Contents

Chapter 1
Constructions of Generalized Differentiation

This chapter is devoted to the exposition of basic tools of first-order generalized differentiation in variational analysis. We follow here the routes of the *dual-space geometric approach* to generalized differentiation in the vein of [507, 522], which revolves around approximation techniques and set extremality. Starting with the *nonconvex robust* construction of the *normal cone* to sets, we continue with the *coderivative* of single-valued and set-valued mappings and the *subdifferential* of (extended-)real-valued functions . For simplicity of the exposition and to emphasize the essence of major variational ideas, our main presentation in Chapters 1–6 is given in finite-dimensional spaces, while we discuss infinite-dimensional extensions in exercises and commentaries to each chapter with the hints and references therein.

Thus, unless otherwise stated, all the spaces under consideration in Chapters 1–6 are *finite-dimensional* and *Euclidean* with the inner product $\langle\cdot,\cdot\rangle$ and the norm $\|\cdot\|$; we often use the standard notation $X = \mathbb{R}^n$ for them. By $\mathbb{B}_X$, or simply by $\mathbb{B}$ if no confusion arises, we denote the *closed unit ball* centered at the origin of the space in question, while $B_r(x)$ stands for the *closed ball centered at* x with radius $r > 0$. In the same way, the closed unit ball in the *dual* space X^*–when it appears–is often denoted by $\mathbb{B}_{X^*}$ or simply by $\mathbb{B}^*$.

Given a nonempty set $\Omega \subset X$, the symbols

$$\mathrm{cl}\,\Omega,\ \ \mathrm{co}\,\Omega,\ \ \mathrm{clco}\,\Omega,\ \ \mathrm{bd}\,\Omega,\ \ \text{and}\ \ \mathrm{int}\,\Omega$$

stand for the standard notions of the *closure*, *convex hull*, *closed convex hull*, *boundary*, and *interior* of the set Ω, respectively.

Recall that a set C is a *cone* in X if $0 \in C$ and $\lambda x \in C$ for all $x \in C$ and $\lambda \geq 0$. The *conic hull* of $\Omega \subset X$ is defined by

$$\mathrm{cone}\,\Omega := \big\{\alpha x \in X \big|\ \alpha \geq 0,\ x \in \Omega\big\}$$

B. S. Mordukhovich, *Variational Analysis and Applications*, Springer Monographs in Mathematics, https://doi.org/10.1007/978-3-319-92775-6_1

unless otherwise stated. In some situations, which will be specifically emphasized (mostly in Chapters 7, 8), the symbol "cone Ω" signifies the *convex* conic hull of the set in question. The *linear combination* of two sets $\Omega_1, \Omega_2 \subset X$ is

$$\alpha_1\Omega_1 + \alpha_2\Omega_1 := \big\{\alpha_1 x_1 + \alpha_2 x_2 \big|\ x_1 \in \Omega_1,\ x_2 \in \Omega_2\big\},$$

where the symbol := means "equal by definition" and where $\alpha_1, \alpha_2 \in \mathbb{R}$ are scalars from $(-\infty, \infty)$. Dealing with the empty set $\emptyset$, we use the conventions that $\Omega + \emptyset := \emptyset$, that $\alpha\emptyset := \emptyset$ if $\alpha \in \mathbb{R} \setminus \{0\}$ and $\alpha\emptyset := \{0\}$ if $\alpha = 0$, and that $\inf\emptyset := \infty$, $\sup\emptyset := -\infty$, and $\|\emptyset\| := \infty$.

Along with single-valued mappings usually denoted by $f\colon \mathbb{R}^n \to \mathbb{R}^m$, we often consider set-valued mappings (or multifunctions) $F\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ with values $F(x) \subset \mathbb{R}^m$ in the collection of all the subsets of $\mathbb{R}^m$ (and similarly, of course, in infinite dimensions). The limiting construction

$$\mathop{\mathrm{Lim\,sup}}_{x\to\bar{x}} F(x) := \Big\{y \in \mathbb{R}^m \Big|\ \exists\, x_k \to \bar{x},\ y_k \to y \ \text{ with } \ y_k \in F(x_k) \ \text{ for all } \ k \in \mathbb{N} := \{1, 2, \ldots\}\Big\} \tag{1.1}$$

is known as the *Painlevé-Kuratowski outer/upper limit* of F at $\bar{x}$. All the mappings considered below are *proper*, i.e., $F(x) \neq \emptyset$ for some $x \in X$.

1.1 Normals and Tangents to Closed Sets

In our *geometric approach* to generalized differentiation, we start with constructing *normals* to nonempty sets $\Omega \subset \mathbb{R}^n$, which is crucial for the whole theory. Given $\bar{x} \in \Omega$, suppose in what follows (unless otherwise stated) that Ω is *locally closed* around $\bar{x} \in \Omega$, i.e., there is $r > 0$ such that the set $\Omega \cap B_r(\bar{x})$ is closed. This doesn't actually restrict the generality since otherwise we can pass to the *closure* of Ω. Anyway, the closedness of sets is truly essential for furnishing most of the *variational arguments* involving limiting procedures.

Although the local closedness assumption for sets, together with the corresponding closed-graph assumption for (set-valued) mappings and lower semicontinuity one for (extended-real-valued) functions, is standing in this book, from time to time, we'll remind the reader about it to emphasize the issue.

1.1.1 Generalized Normals

Given a set $\Omega \subset \mathbb{R}^n$, associate with it the *distance function*

$$\mathrm{dist}(x; \Omega) = d_\Omega(x) := \inf_{z\in\Omega} \|x - z\|, \quad x \in \mathbb{R}^n, \tag{1.2}$$

and define the *Euclidean projector* of $x \in \mathbb{R}^n$ to Ω by

$$\Pi(x;\Omega) = \Pi_\Omega(x) := \big\{ w \in \Omega \;\big|\; \|x - w\| = \text{dist}(x;\Omega) \big\}. \tag{1.3}$$

Under the imposed local closedness of Ω around $\bar{x} \in \Omega$, we have $\Pi(x;\Omega) \neq \emptyset$ for all $x \in \mathbb{R}^n$ sufficiently close to this point.

Definition 1.1 (Basic Normals to Sets). *Let $\Omega \subset \mathbb{R}^n$ with $\bar{x} \in \Omega$. The (basic)* NORMAL CONE *to Ω at $\bar{x}$ is defined by*

$$N(\bar{x};\Omega) = N_\Omega(\bar{x}) := \operatorname*{Lim\,sup}_{x \to \bar{x}} \big[\text{cone}\big(x - \Pi(x;\Omega)\big)\big] \tag{1.4}$$

via the outer limit (1.1). *Each $v \in N(\bar{x};\Omega)$ is called a* BASIC *or* LIMITING NORMAL *to Ω at $\bar{x}$ and is represented as follows: there are sequences $x_k \to \bar{x}$, $w_k \in \Pi(x_k;\Omega)$, and $\alpha_k \geq 0$ such that $\alpha_k(x_k - w_k) \to v$ as $k \to \infty$.*

It is obvious that (1.4) is a closed cone in $\mathbb{R}^n$. A remarkable property of this cone is the possibility to use it for a complete characterization of *boundary points* for locally closed sets, which can be treated as a nonconvex counterpart of the supporting hyperplane theorem for convex sets; cf. Proposition 1.7.

Proposition 1.2 (Normal Cone Characterization of Boundary Points). *For $\bar{x} \in \Omega$ to be a boundary point of Ω, it is necessary and sufficient that $N(\bar{x};\Omega) \neq \{0\}$, i.e., the normal cone* (1.4) *is nontrivial at $\bar{x}$.*

Proof. It is obvious from (1.4) that $N(\bar{x};\Omega) = \{0\}$ if $\bar{x} \in \text{int}\,\Omega$. When $\bar{x} \in \text{bd}\,\Omega$, there is a sequence $\{x_k\} \subset \mathbb{R}^n \setminus \Omega$ such that $x_k \to \bar{x}$ as $k \to \infty$. Pick now a projection $w_k \in \Pi(x_k;\Omega)$ for all k sufficiently large, denote $\alpha_k := \|x_k - w_k\|^{-1}$, and consider the vectors $v_k := \alpha_k(x_k - w_k)$ with $\|v_k\| = 1$. Taking a subsequence of $\{v_k\}$ that converges to some $v \in \mathbb{R}^n$ with $\|v\| = 1$, we get $v \in N(\bar{x};\Omega)$ by the normal cone construction in Definition 1.1. △

Another important property of the normal cone (1.4), which can be easily deduced from the definition, is its *robustness*, i.e., stability with respect to small perturbations of the initial point. In what follows we use the notation

$$x \xrightarrow{\Omega} \bar{x} \iff x \to \bar{x} \;\text{ with }\; x \in \Omega.$$

Proposition 1.3 (Robustness of Basic Normals). *We always have*

$$N(\bar{x};\Omega) = \operatorname*{Lim\,sup}_{x \xrightarrow{\Omega} \bar{x}} N(x;\Omega), \quad \bar{x} \in \Omega.$$

The following simple but useful product property of the normal cone is also a direct consequence of the definition.

Proposition 1.4 (Basic Normals to Products of Sets). *Let* $\Omega_1 \subset \mathbb{R}^n$, $\Omega_2 \subset \mathbb{R}^m$ *with* $(\bar{x}_1, \bar{x}_2) \in \Omega_1 \times \Omega_2$. *Then we have the product formula*

$$N\big((\bar{x}_1, \bar{x}_2); \Omega_1 \times \Omega_2\big) = N(\bar{x}_1; \Omega_1) \times N(\bar{x}_2; \Omega_2).$$

Recall that a set Ω is *convex* if $z + \alpha(x - z) \in \Omega$ for any $x, z \in \Omega$ and $\alpha \in [0, 1]$, i.e., together with any points $x, z \in \Omega$, it contains the entire line segment connecting these points. The following example illustrates that the normal cone (1.4) may be *nonconvex* in very simple settings.

Example 1.5 (Nonconvexity of the Basic Normal Cone). Consider the closed set $\Omega := \{(x, y) \in \mathbb{R}^2 |\ y \ge -|x|\}$. It is easy to see that

$$N\big((0, 0); \Omega\big) = \big\{(v, v) \in \mathbb{R}^2 \big|\ v \le 0\big\} \cup \big\{(v, -v) \in \mathbb{R}^2 \big|\ v \ge 0\big\},$$

which is a nonconvex subset of $\mathbb{R}^2$; see Fig. 1.1.

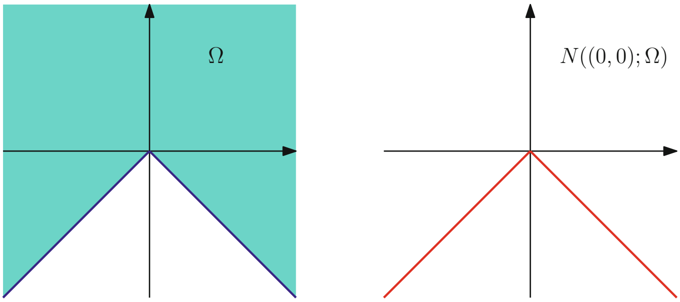


Fig. 1.1 Nonconvexity of the basic normal cone

The next theorem shows that the normal cone (1.4) to Ω at $\bar{x}$ can be equivalently described via the outer limit (1.1) of some convex sets of generalized normals to Ω at points near $\bar{x}$.

Given $x \in \Omega$, define the collection of *regular normals* to Ω at x by

$$\widehat{N}(x; \Omega) = \widehat{N}_\Omega(x) := \Big\{v \in \mathbb{R}^n \ \Big|\ \limsup_{z \xrightarrow{\Omega} x} \frac{\langle v, z - x\rangle}{\|z - x\|} \le 0\Big\} \tag{1.5}$$

and for every $\varepsilon > 0$ consider its ε-*enlargement*

$$\widehat{N}_\varepsilon(x; \Omega) := \Big\{v \in \mathbb{R}^n \ \Big|\ \limsup_{z \xrightarrow{\Omega} x} \frac{\langle v, z - x\rangle}{\|z - x\|} \le \varepsilon\Big\}, \tag{1.6}$$

which reduces to $\widehat{N}(\bar{x}; \Omega) = \widehat{N}_0(\bar{x}; \Omega)$ when $\varepsilon = 0$.

Observe that the convex cone (1.5) may be *trivial*, i.e., $\widehat{N}(\bar{x}; \Omega) = \{0\}$, for boundary points of closed sets as in Example 1.5 with $\bar{x} = (0, 0)$. This phenomenon violates a natural expectation from any normal cone to a closed set at boundary points. On the other hand, the following Theorem 1.6 tells us that elements of $\widehat{N}(x;\Omega)$ at *nearby* points can be used for constructing "real" normals to sets. It motivates us to label the collection of regular normals (1.5) as the *prenormal cone* to Ω at $\bar{x}$; it is also used in the literature as the "regular normal cone." Note that the second representation in (1.7) shows that the limiting process therein is *stable* with respect to ε-enlargements of the prenormal cone. Such a stability is essential to justify a number of significant results of variational analysis and generalized differentiation; see below.

Theorem 1.6 (Equivalent Descriptions of Basic Normals). *Given any $\bar{x} \in \Omega \subset \mathbb{R}^n$, we have the following representations of the basic normal cone:*

$$N(\bar{x};\Omega) = \mathop{\mathrm{Lim\,sup}}_{x \xrightarrow{\Omega} \bar{x}} \widehat{N}(x;\Omega) = \mathop{\mathrm{Lim\,sup}}_{\substack{x \xrightarrow{\Omega} \bar{x}\\ \varepsilon \downarrow 0}} \widehat{N}_\varepsilon(x;\Omega). \tag{1.7}$$

Proof. We split the proof into several steps, which are of their own interest.

Step 1: *If $x \in \mathbb{R}^n$ and $w \in \Pi(x;\Omega)$, then $x - w \in \widehat{N}(w;\Omega)$ and thus*

$$N(\bar{x};\Omega) \subset \mathop{\mathrm{Lim\,sup}}_{x \xrightarrow{\Omega} \bar{x}} \widehat{N}(x;\Omega).$$

Indeed, pick $z \in \Omega$ and get by the choice of w that $\|w - x\|^2 \le \|z - x\|^2 = \|(w - x) + (z - w)\|^2$; hence $0 \le \|z - w\|^2 + 2\langle w - x, z - w\rangle$. This yields

$$\limsup_{z \xrightarrow{\Omega} x} \frac{\langle x - w, z - w\rangle}{\|z - x\|} \le \frac{1}{2} \limsup_{z \xrightarrow{\Omega} x} \|z - w\| = 0,$$

which shows that $x - w \in \widehat{N}(w;\Omega)$. To justify now the displayed inclusion, for any $v \in N(\bar{x};\Omega)$, we have $\alpha_k(x_k - w_k) \to v$ with some $x_k \to \bar{x}$, $w_k \in \Pi(x_k;\Omega)$, and $\alpha_k \ge 0$. It follows from the above that $x_k - w_k \in \widehat{N}(w_k;\Omega)$ and thus $\alpha_k(x_k - w_k) \in \widehat{N}(w_k;\Omega)$ with $w_k \xrightarrow{\Omega} \bar{x}$ due to $\|w_k - x_k\| \le \|x_k - \bar{x}\|$ for all $k \in \mathbb{N}$. This gives us the claimed inclusion.

Step 2: *For any elements $w_\alpha \in \Pi(x + \alpha v;\Omega)$ with $0 \ne v \in \widehat{N}_\varepsilon(x;\Omega)$, $x \in \Omega$, $\varepsilon \ge 0$, and $\alpha > 0$, we have the relationship*

$$\limsup_{\alpha \downarrow 0} \frac{\|w_\alpha - x\|}{\alpha} \le 2\varepsilon.$$

Indeed, it follows from the choice of w_α that $\|(x + \alpha v) - w_\alpha\|^2 \le \|(x + \alpha v) - x\|^2 = \|\alpha v\|^2$, which implies the equivalent conditions

$$\Big[\|w_\alpha - x\|^2 + 2\alpha\langle v, x - w_\alpha\rangle \le 0\Big] \Longleftrightarrow \Big[\frac{\|w_\alpha - x\|}{\alpha} \le 2\frac{\langle v, w_\alpha - x\rangle}{\|w_\alpha - x\|}\Big].$$

It follows further from the classical Cauchy-Schwarz inequality that

$$\|w_\alpha - x\|^2 \le 2\alpha\langle v, w_\alpha - x\rangle \le 2\alpha\|v\| \cdot \|w_\alpha - x\|$$

and so $\|w_\alpha - x\| \le 2\alpha\|v\| \to 0$ as $\alpha \downarrow 0$. Thus the choice of v yields

$$\limsup_{\alpha\downarrow 0} \frac{\langle v, w_\alpha - x\rangle}{\|x - w_\alpha\|} \le \limsup_{z \xrightarrow{\Omega} x} \frac{\langle v, z - x\rangle}{\|z - x\|} \le \varepsilon,$$

which justifies by (1.6) the claimed estimate.

Step 3: *We have the inclusion*

$$\operatorname*{Lim\,sup}_{\substack{x \xrightarrow{\Omega} \bar{x} \\ \varepsilon\downarrow 0}} \widehat{N}_\varepsilon(x; \Omega) \subset \operatorname*{Lim\,sup}_{x \xrightarrow{\Omega} \bar{x}} \widehat{N}(x; \Omega).$$

To show this, take any v from the left-hand side set above and by (1.1) find $\varepsilon_k \downarrow 0$, $x_k \xrightarrow{\Omega} \bar{x}$, and $v_k \in \widehat{N}_{\varepsilon_k}(x_k; \Omega)$ such that $v_k \to v$ as $k \to \infty$. By Step 2 there are $w_k \in \Omega$ and $\alpha_k \downarrow 0$ satisfying

$$w_k \in \Pi(x_k + \alpha_k v_k; \Omega) \ \text{ and } \ \|w_k - x_k\| \le 2\varepsilon_k\alpha_k, \quad k \in \mathbb{N},$$

which implies that $w_k \to \bar{x}$ when $k \to \infty$. As shown in Step 1, $(x_k + \alpha_k v_k) - w_k \in \widehat{N}(w_k; \Omega)$ and so, since $\widehat{N}(w_k; \Omega)$ is a cone,

$$v_k + \frac{1}{\alpha_k}(x_k - w_k) = \frac{1}{\alpha_k}\Big((x_k + \alpha_k v_k) - w_k\Big) \in \widehat{N}(w_k; \Omega).$$

The latter implies that $v_k + \frac{1}{\alpha_k}(x_k - w_k) \to v$ as $k \to \infty$, which therefore justifies the statement claimed in this step.

Step 4: *We have the inclusion*

$$\widehat{N}(x; \Omega) \subset N(x; \Omega) \ \text{ for all } \ x \in \Omega.$$

To verify it, take any $v \in \widehat{N}(x; \Omega)$ and for large $k \in \mathbb{N}$ define $z_k := x + \frac{1}{k}v$ and pick $w_k \in \Pi(x + \frac{1}{k}v; \Omega)$. Then we get $v = k(z_k - x) = v_k + k(w_k - x)$, where $v_k := k(z_k - w_k) \in \operatorname{cone}\big(z_k - \Pi(z_k; \Omega)\big)$ and $z_k \to x$. It follows from Step 2 that $k(w_k - x) \to 0$ and so $v_k \to v$, which justifies the claimed statement.

Step 5: *We have the inclusion*

$$\operatorname*{Lim\,sup}_{x \xrightarrow{\Omega} \bar{x}} \widehat{N}(x; \Omega) \subset N(\bar{x}; \Omega).$$

Indeed, taking v from the set on the left-hand side above gives us $v_k \to v$ and $x_k \xrightarrow{\Omega} \bar{x}$ with $v_k \in \widehat{N}(x_k; \Omega)$. Denote $G(z) := \operatorname{cone}[z - \Pi(z; \Omega)]$ and get from (1.4) and Step 4 that $\widehat{N}(x; \Omega) \subset \operatorname{Lim\,sup}_{z \to x} G(z)$. Hence for each $k \in \mathbb{N}$ we find $z_k \in$

$\mathbb{R}^n$ and $y_k \in G(z_k)$ with $\|z_k - x_k\| \le 1/k$ and $\|y_k - v_k\| \le 1/k$. Since $z_k \to \bar{x}$ and $y_k \to v$, this ensures that $v \in \operatorname{Lim\,sup}_{x\to\bar{x}} G(x) = N(\bar{x};\Omega)$, which justifies the claimed inclusion and thus completes the proof. △

The next proposition shows that for convex sets Ω both constructions (1.4) and (1.5) reduce to the normal cone of convex analysis.

Proposition 1.7 (Normals to Convex Sets). *Let Ω be convex, and let $\bar{x}$ be any point of Ω. Then we have the representations*

$$\widehat{N}_\varepsilon(\bar{x};\Omega) = \Big\{ v \in \mathbb{R}^n \Big|\ \langle v, x - \bar{x}\rangle \le \varepsilon\|x - \bar{x}\| \text{ for all } x \in \Omega \Big\},\ \varepsilon \ge 0, \quad (1.8)$$

$$N(\bar{x};\Omega) = \widehat{N}(\bar{x};\Omega) = \Big\{ v \in \mathbb{R}^n \Big|\ \langle v, x - \bar{x}\rangle \le 0 \text{ for all } x \in \Omega \Big\}. \quad (1.9)$$

Proof. The inclusion "$\supset$" in (1.8) obviously holds for an arbitrary set Ω. To verify the opposite inclusion in (1.8) when Ω is convex, fix any $\varepsilon \ge 0$, take $v \in \widehat{N}_\varepsilon(\bar{x};\Omega)$, and then fix $x \in \Omega$. By the convexity of Ω, we have that $x_\alpha := \bar{x} + \alpha(x - \bar{x}) \in \Omega$ for all $0 \le \alpha \le 1$ with $x_\alpha \to \bar{x}$ as $\alpha \downarrow 0$. Taking any $\gamma > 0$ and using definition (1.6) give us

$$\langle v, x_\alpha - \bar{x}\rangle \le (\varepsilon + \gamma)\|x_\alpha - \bar{x}\| \ \text{ for all small } \alpha > 0.$$

Substituting the expression for x_α into this inequality justifies (1.8). The representation (1.9) for $N(\bar{x};\Omega)$ follows from (1.8) taken at any $x \in \Omega$ by passing to the limit due to Theorem 1.6. △

1.1.2 Tangential Preduality

It follows from (1.9) that Proposition 1.2 reduces for convex sets Ω to the fact that for any $\bar{x} \in \operatorname{bd}\Omega$ there is $0 \ne v \in \mathbb{R}^n$ with $\langle v, x\rangle \le \langle v, \bar{x}\rangle$ whenever $x \in \Omega$. This is the classical *supporting hyperplane theorem*, which is equivalent to the *separation theorem* for convex sets and plays a fundamental role in convex analysis and its various extensions. One of the implementations of this fundamental result is the *duality/polarity* correspondence

$$N(\bar{x};\Omega) = T^*(\bar{x};\Omega) := \{v \in \mathbb{R}^n \Big|\ \langle v, w\rangle \le 0 \ \text{ for all } \ w \in T(\bar{x};\Omega)\Big\} \quad (1.10)$$

between the normal cone to *convex* sets given in (1.9) and the tangent cone $T(\bar{x};\Omega) := \operatorname{cl}\{w \in \mathbb{R}^n |\ \exists \alpha > 0 \text{ with } \bar{x} + \alpha w \in \Omega\}$ of convex analysis.

Note to this end that the duality scheme of type (1.10) has been conventionally used in nonsmooth analysis to define normal cones to nonconvex sets via *some* tangential approximations. It is easy to see that any normal cone obtained in this scheme is *automatically convex*, even when the generating tangential approximation is not. This shows that our basic normal cone (1.4) *cannot be tangentially generated* due to its intrinsic nonconvexity. However, it is not the case for the prenormal cone (1.5), which is convex and in fact can be obtained by the duality scheme from the following tangential approximation.

Definition 1.8 (Contingent Cone). *Given $\Omega \subset \mathbb{R}^n$ and $\bar{x} \in \Omega$, the* CONTINGENT CONE *to Ω at $\bar{x}$ is defined by*

$$T(\bar{x};\Omega) := \operatorname*{Lim\,sup}_{t\downarrow 0} \frac{\Omega - \bar{x}}{t} \tag{1.11}$$

via the outer limit (1.1). Each $w \in T(\bar{x};\Omega)$ is called a TANGENT *to Ω at $\bar{x}$ and is represented as follows: there are sequences $\{x_k\} \subset \Omega$ and $\{\alpha_k\} \subset \mathbb{R}_+$ such that $x_k \to \bar{x}$ and $\alpha_k(x_k - \bar{x}) \to w$ as $k \to \infty$.*

When Ω is convex, the contingent cone (1.11) agrees with the classical tangent cone of convex analysis, while in general it may be *nonconvex* as for the set $\Omega := \{(x_1, x_2) \in \mathbb{R}^2 |\ x_2 = |x_1|\}$ at $\bar{x} = (0,0)$, where $T(\bar{x};\Omega) = \Omega$. Let us now show that its (convex) dual cone is exactly the prenormal cone (1.5).

Proposition 1.9 (Duality Between the Prenormal and Contingent Cones). *For any $\Omega \subset \mathbb{R}^n$ and $\bar{x} \in \Omega$, we have the duality correspondence*

$$\widehat{N}(\bar{x};\Omega) = T^*(\bar{x};\Omega)$$

between the prenormal cone (1.5) and the contingent cone (1.11).

Proof. Fix any vectors $v \in \widehat{N}(\bar{x};\Omega)$ and $w \in T(\bar{x};\Omega)$. By (1.11) there are sequences $t_k \downarrow 0$ and $w_k \to w$ with $\bar{x} + t_k w_k \in \Omega$ for all $k \in \mathbb{N}$. Substituting this combination into (1.5) and picking any $\gamma > 0$, we get

$$\langle v, w_k\rangle \le \gamma \|w_k\| \ \text{ for all large } \ k \in \mathbb{N}.$$

Passing here to the limit as $k \to \infty$ shows that $\langle v, w\rangle \le 0$, and thus we get $\widehat{N}(\bar{x};\Omega) \subset T^*(\bar{x};\Omega)$ by the dual cone definition in (1.10).

To verify the converse inclusion, fix $v \notin \widehat{N}(\bar{x};\Omega)$ and find by (1.5) a positive number γ and a sequence $x_k \xrightarrow{\Omega} \bar{x}$ such that

$$\langle v, x_k - \bar{x}\rangle > \gamma \|x_k - \bar{x}\| \ \text{ for all large } \ k \in \mathbb{N};$$

so $x_k \ne \bar{x}$. Let $\alpha_k := \|x_k - \bar{x}\|^{-1}$ and suppose without loss of generality that

$$\frac{x_k - \bar{x}}{\|x_k - \bar{x}\|} \to w \ \text{ as } \ k \to \infty \ \text{ for some } \ w \in \mathbb{R}^n.$$

By construction (1.11) we have $w \in T(\bar{x};\Omega)$ while $\langle v, w\rangle \ge \gamma > 0$ by passing to the limit above. Thus we get $v \notin T^*(\bar{x};\Omega)$, which justifies the inclusion $T^*(\bar{x};\Omega) \subset \widehat{N}(\bar{x};\Omega)$ and completes the proof of the proposition. △

Combining Theorem 1.6 and Proposition 1.9 tells us that, although the normal cone (1.4) cannot be tangentially generated at the point in question, it admits an approximation by tangentially generated normals to the set at points nearby. This phenomenon can be naturally labeled as *tangential preduality* for basic normals. However, it is essentially *finite-dimensional*; see [522] and Section 1.5 below for more details.

1.1.3 Smooth Variational Description

We conclude this section with a *variational* property of regular normals giving their *smooth* description, which is convenient for applications. By Theorem 1.6 this provides a *smooth limiting description* of the normal cone (1.4). Everywhere we understand *differentiability* of $\varphi\colon \mathbb{R}^n \to \mathbb{R}$ at $\bar{x}$ with the derivative/gradient $\nabla\varphi(\bar{x}) \in \mathbb{R}^n$ in the standard (Fréchet) sense

$$\lim_{x\to\bar{x}} \frac{\varphi(x) - \varphi(\bar{x}) - \langle \nabla\varphi(\bar{x}), x - \bar{x}\rangle}{\|x - \bar{x}\|} = 0 \tag{1.12}$$

while the *smoothness* (of class $\mathcal{C}^1$) property of φ *around* $\bar{x}$ is its differentiability on a neighborhood U of $\bar{x}$ with the continuous gradient $\nabla\varphi\colon U \to \mathbb{R}^n$.

Theorem 1.10 (Smooth Variational Description of Regular Normals). *Let $\Omega \subset \mathbb{R}^n$ with $\bar{x} \in \Omega$. Then regular normals to $\bar{x}$ can be described in the following two equivalent ways:*

(i) *We have $v \in \widehat{N}(\bar{x};\Omega)$ if and only if there is a neighborhood U of $\bar{x}$ and a function $\psi\colon U \to \mathbb{R}$ such that ψ is differentiable at $\bar{x}$ with $\nabla\psi(\bar{x}) = v$ and ψ achieves its local maximum relative to Ω at $\bar{x}$.*

(ii) *We have $v \in \widehat{N}(\bar{x};\Omega)$ if and only if there is a smooth and concave function ψ on $\mathbb{R}^n$ such that $\nabla\psi(\bar{x}) = v$ and ψ achieves its global maximum relative to Ω uniquely at $\bar{x}$.*

Proof. It is not hard to verify (i) based on definition (1.5). Indeed, for any $\psi\colon U \to \mathbb{R}$ with the properties from (i), we have

$$\psi(x) = \psi(\bar{x}) + \langle v, x - \bar{x}\rangle + o(\|x - \bar{x}\|) \le \psi(\bar{x}) \text{ for all } x \in U.$$

Hence $\langle v, x - \bar{x}\rangle + o(\|x - \bar{x}\|) \le 0$ and $v \in \widehat{N}(\bar{x};\Omega)$ by (1.5). Conversely, for any $v \in \widehat{N}(\bar{x};\Omega)$, consider the function

$$\psi(x) := \begin{cases} \min\big\{0, \langle v, x - \bar{x}\rangle\big\} & \text{if } x \in \Omega,\\ \langle v, x - \bar{x}\rangle & \text{otherwise,}\end{cases}$$

which surely satisfies the properties listed in (i).

To justify (ii), we need to verify the "only if" part the proof of which is essentially more involved. We split it into several steps.

Step 1: *Let $\rho\colon [0,\infty) \to [0,\infty)$ be a function having the right-hand derivative $\rho'_+(0)$ and satisfying the conditions*

$$\rho(0) = \rho'_+(0) = 0 \ \textit{ and } \ \rho(t) \le \alpha + \beta t \ \textit{ for all } \ t \ge 0$$

with some positive constants α and β. Then there exist $\gamma > 0$ and a nondecreasing, convex, and $\mathcal{C}^1$-smooth function $\sigma : [0, 2\gamma) \to [0,\infty)$ such that

$$\sigma(0) = \sigma'_+(0) = 0 \ \textit{ and } \ \sigma(t) > \rho(t) \ \textit{ for } \ t \in (0, 2\gamma).$$

To construct σ, choose a sequence of $a_k > 0$ with $a_{k+1} < \frac{1}{2}a_k$ and

$$\rho(t) + t^2 < 2^{-(k+3)}t \ \text{ if } \ t \in [0, a_k] \ \text{ for all } \ k \in \mathbb{N}.$$

Let $\gamma := \frac{1}{2}a_1$ and define $r\colon [0, 2\gamma] \to [0,\infty)$ by $r(0) := 0$, $r(a_k) := 2^{-k}$, and so that r is linear on $[a_{k+1}, a_k]$ for all $k \in \mathbb{N}$. Then define $\sigma : [0, 2\gamma) \to [0,\infty)$ by

$$\sigma(t) := \int_0^t r(\xi)d\xi \ \text{ for } \ t \in [0, 2\gamma)$$

and show that it possesses all the required properties. Its smoothness, monotonicity, convexity, and the equalities $\sigma(0) = \sigma'_+(0) = 0$ follow directly from the definition and standard facts of real analysis. To check the remaining properties of σ, fix $t \in (0, 2\gamma)$ and observe that $t \in [a_{k+1}, a_k)$ for some $k \in \mathbb{N}$. Then, by the above constructions of σ and r, we get

$$\begin{aligned}\sigma(t) &\ge \int_{a_{k+1}}^t r(\xi)d\xi + \int_{\frac{1}{2}a_{k+1}}^{a_{k+1}} r(\xi)d\xi \ge \int_{a_{k+1}}^t 2^{-(k+1)}d\xi + \int_{\frac{1}{2}a_{k+1}}^{a_{k+1}} 2^{-(k+2)}d\xi \\ &= \frac{t - a_{k+1}}{2^{k+1}} + \frac{a_{k+1}}{2^{k+3}} \ge \frac{t}{2^{k+3}} > \rho(t),\end{aligned}$$

which justifies all the properties of $\sigma(t)$ listed above.

Step 2: *Let $\rho\colon [0,\infty) \to [0,\infty)$ be given as in Step 1. Then there is a nondecreasing, convex, and $\mathcal{C}^1$-smooth function $\tau : [0,\infty) \to [0,\infty)$ such that*

$$\tau(0) = \tau'_+(0) = 0 \ \textit{ and } \ \tau(t) > \rho(t) \ \textit{ for all } \ t > 0.$$

Given the numbers $\alpha, \beta > 0$ and the function $\sigma(t)$ built above, choose $\lambda > 1$ with $\lambda\sigma(\gamma) > \alpha + \beta\gamma$, and consider the following two cases in constructing the function $\tau(t)$ with the claimed properties:

(a) Let $\lambda\sigma'(\gamma) \le \beta$. Take $\mu \ge \lambda$ with $\mu\sigma'(\gamma) = \beta$ and define

$$\tau(t) := \begin{cases} \mu\sigma(t) & \text{if } \ 0 \le t \le \gamma, \\ \mu\sigma(\gamma) + \beta(t-\gamma) & \text{if } \ t > \gamma. \end{cases}$$

It is easy to see that this function is nondecreasing, convex, and continuous everywhere on $[0,\infty)$ including $t=\gamma$. Moreover, $\tau'_-(\gamma)=\mu\sigma'(\gamma)$ and $\tau'_+(\gamma)=\beta=\mu\sigma'(\gamma)$ due to the choice of μ, which implies the continuous differentiability of τ on $[0,\infty)$. It follows from the definition of τ and the assumptions on ρ that $\tau(0)=\tau'_+(0)=0$, that $\tau(t)\ge\sigma(t)>\rho(t)$ for $0<t\le\gamma$, and that $\tau(t)=\mu\sigma(\gamma)+\beta(t-\gamma)>\alpha+\beta t\ge\rho(t)$ for $t>\gamma$. This ensures the required properties of $\tau(\cdot)$ in the case under consideration.

(b) Let $\lambda\sigma'(\gamma)>\beta$. In this case we define a nondecreasing and convex function $\tau\colon[0,\infty)\to[0,\infty)$ by

$$\tau(t):=\begin{cases}\lambda\sigma(t) & \text{if}\quad 0\le t\le\gamma,\\ \lambda\sigma(\gamma)-\lambda\gamma\sigma'(\gamma)+\lambda\sigma'(\gamma)t & \text{if}\quad t>\gamma.\end{cases}$$

Again, a straightforward verification yields that $\tau(t)$ is a $\mathcal{C}^1$-smooth function $[0,\infty)$ satisfying all the requirements on $[0,\gamma]$. By the choice of λ, we get

$$\tau(t)\ge\alpha+\beta\gamma+\lambda\sigma'(\gamma)(t-\gamma)>\alpha+\beta\gamma+\beta(t-\gamma)=\alpha+\beta t\ge\rho(t)$$

for $t>\gamma$, which verifies the statement claimed in Step 2.

Step 3: *Let $v\in\widehat{N}(\bar{x};\Omega)$. Then there is a function $\psi\colon\mathbb{R}^n\to\mathbb{R}$ having all the properties listed in assertion* (ii).

To proceed, consider the positive-valued function

$$\rho(t):=\sup\Big\{\langle v,x-\bar{x}\rangle\ \Big|\ x\in\Omega,\ \|x-\bar{x}\|\le t\Big\}\ \text{ for }\ t\ge 0,\qquad(1.13)$$

which clearly satisfies all the assumptions formulated in Step 1 due to the definition of regular normals. By Step 2 we get the corresponding function $\tau\colon[0,\infty)\to[0,\infty)$ and construct $\psi\colon\mathbb{R}^n\to\mathbb{R}$ by

$$\psi(x):=-\tau(\|x-\bar{x}\|)-\|x-\bar{x}\|^2+\langle v,x-\bar{x}\rangle,\quad x\in\mathbb{R}^n.$$

Note that this function is concave on $\mathbb{R}^n$ with $\psi(\bar{x})=0$ since $\tau(\cdot)$ is convex and nondecreasing on $[0,\infty)$ with $\tau(0)=0$. We also have

$$\psi(x)+\|x-\bar{x}\|^2\le-\rho(\|x-\bar{x}\|)+\langle v,x-\bar{x}\rangle\le 0=\psi(\bar{x})\ \text{ for all }\ x\in\Omega,$$

which implies that $\psi(x)$ achieves its global maximum over Ω uniquely at $\bar{x}$. Observe that $\psi(x)$ is differentiable at any $x\ne\bar{x}$ due the smoothness of the function $\tau(\cdot)$ and the Euclidean norm $\|\cdot\|$ at nonzero points of $\mathbb{R}^n$. To justify (ii), it remains to observe that $\psi(x)$ is differentiable at $x=\bar{x}$ with $\nabla\psi(\bar{x})=v$, which follows from the smoothness of $\tau(t)$ with $\tau'_+(0)=0$ by the classical chain rule. This completes the proof of the theorem. △

1.2 Coderivatives of Mappings

In this section we consider generalized differentiation of set-valued mappings/multifunctions $F\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ with the values $F(x) \subset \mathbb{R}^m$ which may be, in particular, empty or singletons. If the latter is the case for all x, we usually use the standard notation $f\colon \mathbb{R}^n \to \mathbb{R}^m$ for single-valued mappings.

1.2.1 Set-Valued Mappings

Given $F\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, we say that it is *closed-valued*, *convex-valued*, ..., if all the values $F(x)$ are closed, convex, ..., respectively. With each mapping F, we associate its main geometric description—the *graph*

$$\operatorname{gph} F := \big\{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m \big|\ y \in F(x)\}$$

and denote its *domain*, *kernel*, and *range* by

$$\operatorname{dom} F := \big\{x \in \mathbb{R}^n \big|\ F(x) \neq \emptyset\big\}, \quad \ker F := \{x \in \mathbb{R}^n \big|\ 0 \in F(x)\},$$

$$\operatorname{rge} F := \big\{y \in \mathbb{R}^m \big|\ \exists\, x \in \mathbb{R}^n \ \text{ with } \ y \in F(x)\big\},$$

respectively. The (direct) *image* of a set $\Omega \subset \mathbb{R}^n$ under F is

$$F(\Omega) := \big\{y \in \mathbb{R}^m \big|\ \exists\, x \in \Omega \ \text{ with } \ y \in F(x)\big\},$$

while the *inverse image/preimage* of $\Theta \subset \mathbb{R}^m$ under this mapping is

$$F^{-1}(\Theta) := \big\{x \in \mathbb{R}^n \big|\ F(x) \cap \Theta \neq \emptyset\big\},$$

which reduces to $f^{-1}(\Theta) = \{x \in \mathbb{R}^n|\ f(x) \in \Theta\}$ in the single-valued case. The *inverse mapping* $F^{-1}\colon \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ to F is defined by

$$F^{-1}(y) := \big\{x \in \mathbb{R}^n \big|\ y \in F(x)\big\}.$$

It is clear that $\operatorname{dom} F^{-1} = \operatorname{rge} F$, that $\operatorname{rge} F^{-1} = \operatorname{dom} F$, and that

$$\operatorname{gph} F^{-1} = \big\{(y, x) \in \mathbb{R}^m \times \mathbb{R}^n \big|\ (x, y) \in \operatorname{gph} F\big\}.$$

We say that $F\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is *locally bounded* around $\bar{x}$ if there is a neighborhood U of $\bar{x}$ such that the image set $F(U)$ is bounded in $\mathbb{R}^m$.

Recall also that a mapping $F\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is *positively homogeneous* if $0 \in F(0)$ and $F(\alpha x) \supset \alpha F(x)$ for all $\alpha > 0$, $x \in \mathbb{R}^n$; i.e., its graph is a cone in $\mathbb{R}^n \times \mathbb{R}^m$. The *norm* of a positively homogeneous mapping is given by

$$\|F\| := \sup\big\{\|y\| \ \big|\ y \in F(x) \ \text{ and } \ \|x\| \le 1\big\}. \tag{1.14}$$

1.2.2 Coderivative Definition and Elementary Properties

Now we are ready to define our main generalized differential concept for mappings called the *coderivative*. We proceed geometrically and associate the coderivative with the *normal cone* (1.4) to the *graph* of the given set-valued or single-valued mapping. The term "**co**derivative" reflects the dual-space nature of this construction for mappings generated by the normal cone to sets. As follows from the discussion in Section 1.1, the basic coderivative defined below is a nonconvex-valued mapping that is *not dual* to any derivative-like objects generated by tangential approximations of sets.

In accordance with Section 1.1, we consider without loss of generality set-valued mappings whose *graphs* are *locally closed* around the reference points.

Definition 1.11 (Basic Coderivative of Set-Valued Mappings). *Consider* $F\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ *with* $\mathrm{dom}\, F \neq \emptyset$*, and let* $(\bar{x}, \bar{y}) \in \mathrm{gph}\, F$*. The* (*basic*) CODERIVATIVE *of* F *at* $(\bar{x}, \bar{y})$ *is a multifunction* $D^*F(\bar{x}, \bar{y})\colon \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ *with the values*

$$D^*F(\bar{x}, \bar{y})(v) := \big\{u \in \mathbb{R}^n \,\big|\, (u, -v) \in N\big((\bar{x}, \bar{y}); \mathrm{gph}\, F\big)\big\}, \quad v \in \mathbb{R}^m, \tag{1.15}$$

generated by the normal cone (1.4) *to the graph of* F *at* $(\bar{x}, \bar{y})$.

Defining then the *precoderivative* (known also as the *regular coderivative*) of F at $(\bar{x}, \bar{y})$ via the prenormal cone (1.5) by

$$\widehat{D}^*F(\bar{x}, \bar{y})(v) := \big\{u \in \mathbb{R}^n \,\big|\, (u, -v) \in \widehat{N}\big((\bar{x}, \bar{y}); \mathrm{gph}\, F\big)\big\}, \quad v \in \mathbb{R}^m, \tag{1.16}$$

and, employing Theorem 1.6, we get the limiting representation

$$D^*F(\bar{x}, \bar{y})(\bar{v}) = \mathop{\mathrm{Lim\,sup}}_{\substack{(x,y)\xrightarrow{\mathrm{gph}\, F}(\bar{x},\bar{y})\\ v \to \bar{v}}} \widehat{D}^*F(x, y)(v) \tag{1.17}$$

and the similar one in terms of the ε-enlargements $\widehat{D}^*_\varepsilon F(x, y)$ of (1.16) defined via $\widehat{N}_\varepsilon((x, y); \mathrm{gph}\, F)$ as $\varepsilon \downarrow 0$. In what follows we omit $\bar{y}$ in notation (1.15) and (1.16) if the mapping is single-valued at $\bar{x}$.

It should be mentioned that employing in (1.15) the basic normal cone construction from (1.4) to the graphical set $\mathrm{gph}\, F \subset \mathbb{R}^n \times \mathbb{R}^m$ requires the usage of the *Euclidean norm* $\|(x, y)\| = \sqrt{\|x\|^2 + \|y\|^2}$ on the product space, which is very beneficial in many situations due to remarkable variational and smoothness (off the origin) properties of the Euclidean norm. However, in those proofs below which are based on the equivalent representations of basic normals from Theorem 1.6 implemented in (1.17), it is more convenient to employ the *sum norm* on the product $X \times Y$ given by

$$\|(x, y)\| := \|x\| + \|y\| \text{ for all } x \in X,\ y \in Y. \tag{1.18}$$

It is not hard to check that the representations in (1.7) and (1.17) are *invariant* with respect to any equivalent norm used on the space in question. Recall to this end that all the norms on a finite-dimensional space are equivalent.

Observe that both basic and regular coderivatives are *positively homogeneous* with respect of their argument v. We show next that for single-valued mappings $F = f\colon \mathbb{R}^n \to \mathbb{R}^m$ smooth around the reference point $\bar{x}$, they both are single-valued and linear in v, thus being reduced to the *adjoint/transpose* Jacobian matrix $\nabla f(\bar{x})^*\colon \mathbb{R}^m \to \mathbb{R}^n$ applied to v; we keep the notation $\nabla f(\bar{x})$ for the Jacobian matrix. As always, by *smoothness* (i.e., of class $\mathcal{C}^1$) of f *around* $\bar{x}$, we mean its continuous differentiability on a neighborhood of $\bar{x}$. Note that the vast majority, if not all, of the results given in this book for smooth mappings hold true for those, which are merely *strictly differentiable at* $\bar{x}$ with the strict derivative operator $\nabla f(\bar{x})\colon \mathbb{R}^n \to \mathbb{R}^m$ in the sense that

$$\lim_{x,z\to\bar{x}} \frac{f(x) - f(z) - \nabla f(\bar{x})(x-z)}{\|x-z\|} = 0. \tag{1.19}$$

However, proofs in the strict differentiable case are usually more involved, and we restrict ourselves to $\mathcal{C}^1$-smooth mappings for simplicity; cf. [522, 523].

Proposition 1.12 (Coderivatives of Smooth Mappings). *Let the mapping $f\colon \mathbb{R}^n \to \mathbb{R}^m$ be of class $\mathcal{C}^1$ around $\bar{x}$. Then we have the representations*

$$D^* f(\bar{x})(v) = \widehat{D}^* f(\bar{x})(v) = \big\{\nabla f(\bar{x})^* v\big\} \ \textit{for all} \ \ v \in \mathbb{R}^m.$$

Proof. Note first that the inclusion $u \in \widehat{D}^* f(x)(v)$ means by definition that

$$\langle u, z-x\rangle - \langle v, f(z) - f(x)\rangle \le \gamma\big(\|z-x\| + \|f(z) - f(x)\|\big)$$

for an arbitrary number $\gamma > 0$ when z is sufficiently close to x. On the other hand, by the differentiability of f at x, we have that

$$\langle u - \nabla f(x)^* v, z-x\rangle \le \gamma\|z-x\|.$$

Combining these facts with the definition of the adjoint operator shows that $\widehat{D}^* f(x)(v) = \big\{\nabla f(x)^* v\}$ for all x close to $\bar{x}$. Passing here to the limit as $x \to \bar{x}$ and using the continuity of ∇f together with the coderivative representation (1.17) justify the formula for $D^* f(\bar{x})(v)$. △

Another simple and expected coderivative representation holds for *convex-graph* multifunctions, i.e., those for which the set gph F is convex.

Proposition 1.13 (Coderivatives of Convex-Graph Mappings). *Let the graph of $F\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ be convex. Then*

$$\begin{aligned} D^* F(\bar{x}, \bar{y})(v) &= \widehat{D}^* F(\bar{x}, \bar{y})(v) \\ &= \Big\{u \in \mathbb{R}^n \,\Big|\, \langle u, \bar{x}\rangle - \langle v, \bar{y}\rangle = \max_{(x,y)\in \mathrm{gph}\, F} \big[\langle u, x\rangle - \langle v, y\rangle\big]\Big\} \end{aligned}$$

for all $(\bar{x}, \bar{y}) \in$ gph F and $v \in \mathbb{R}^m$.

Proof. It follows from the normal cone representations in Proposition 1.9. △

In general the coderivative may take *nonconvex* and also *empty* values. Let us illustrate this by direct calculations based on the definition.

Example 1.14 (Coderivative Calculations).
(i) Consider first the function $f(x) := |x|$ on $\mathbb{R}$ and calculate its coderivative at $\bar{x} = 0$. Using the normal cone definition (1.4) gives us (see Fig. 1.2)

$$N\big((0,0);\operatorname{gph} f\big) = \big\{(x,y) \in \mathbb{R}^2 \big|\ y = |x| \ \ \&\ \ y \le -|x|\big\}.$$

Thus the coderivative (1.15) of this function is calculated by

$$D^*f(0)(v) = \begin{cases} [-v, v] & \text{if } \ v \ge 0, \\ \{-v, v\} & \text{if } \ v < 0 \end{cases}$$

and has, in particular, nonconvex values when $v < 0$. Note that the precoderivative (1.16) in this case is given by

$$\widehat{D}^*f(0)(v) = \begin{cases} [-v, v] & \text{if } \ v \ge 0, \\ \emptyset & \text{if } \ v < 0. \end{cases}$$

(ii) For another function $f(x) := |x|^\alpha$ with $\alpha \in (0,1)$, we have

$$N\big((0,0);\operatorname{gph} f\big) = \big\{(x,y) \in \mathbb{R}^2 \big|\ y \le 0\big\}$$

(see Fig. 1.3), and hence the coderivative (1.15) takes empty values

$$D^*f(0)(v) = \widehat{D}^*f(0)(v) = \begin{cases} \mathbb{R} & \text{if } \ v \ge 0, \\ \emptyset & \text{if } \ v < 0. \end{cases}$$

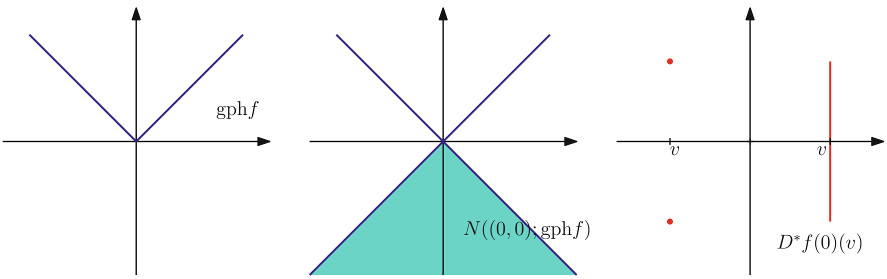


Fig. 1.2 Coderivative of $f(x) = |x|$

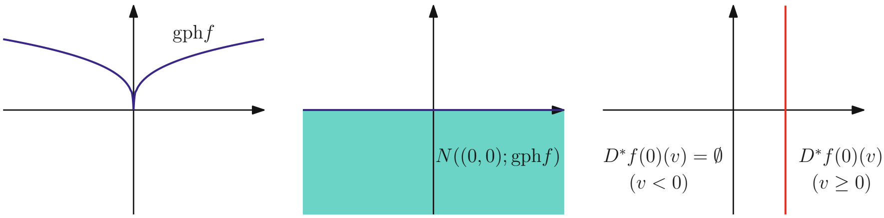


Fig. 1.3 Coderivative of $f(x) = |x|^\alpha$, $0 < \alpha < 1$

1.2.3 Extremal Property of Convex-Valued Multifunctions

Now we present an important result revealing an *extremal property* of *convex-valued* multifunctions formulated via the basic coderivative. This property is useful for various applications; see, e.g., Section 1.5. The proof is simple enough due to the usage of some previous considerations.

A set-valued mapping $F\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is said to be *inner/lower semicontinuous* at the *domain* point $\bar{x} \in \operatorname{dom} F$ if we have

$$F(\bar{x}) = \operatorname*{Lim\,inf}_{x\to\bar{x}} F(x) := \Big\{ y \Big| \; \forall x_k \xrightarrow{\operatorname{dom} F} \bar{x} \; \exists y_k \to y, \;\; y_k \in F(x_k) \Big\} \tag{1.20}$$

in terms of the *Painlevé-Kuratowski inner/lower limit* F at $\bar{x}$.

Theorem 1.15 (Extremal Property of Convex-Valued Mappings via Their Basic Coderivative). *Let* $F\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ *be inner semicontinuous at* $\bar{x} \in \operatorname{dom} F$ *and convex-valued around this point, and let* $v \in \operatorname{dom} D^*F(\bar{x},\bar{y})$ *for some* $\bar{y} \in F(\bar{x})$. *Then we have the extremal property*

$$\langle v, \bar{y}\rangle = \min_{y\in F(\bar{x})} \langle v, y\rangle. \tag{1.21}$$

Proof. By $v \in \operatorname{dom} D^*F(\bar{x},\bar{y})$ and the coderivative definition (1.15), there is $u \in \mathbb{R}^n$ with $(u,-v) \in N((\bar{x},\bar{y}); \operatorname{gph} F)$. By Theorem 1.6 we find sequences $(x_k,y_k) \to (\bar{x},\bar{y})$ with $y_k \in F(x_k)$ and $(u_k,v_k) \to (u,v)$ such that

$$\limsup_{(x,y)\xrightarrow{\operatorname{gph} F}(x_k,y_k)} \frac{\langle u_k, x-x_k\rangle - \langle v_k, y-y_k\rangle}{\|(x,y)-(x_k,y_k)\|} \le 0, \quad k \in \mathbb{N}.$$

Putting there $x = x_k$ shows that $-v_k \in \widehat{N}(y_k; F(x_k))$. Since all the sets $F(x_k)$ are convex, we get from Proposition 1.9 that $\langle v_k, y-y_k\rangle \ge 0$ for any $y \in F(x_k)$. Suppose now that there is $\widetilde{y} \in F(\bar{x})$ such that $\langle v,\widetilde{y}\rangle < \langle v,\bar{y}\rangle$. Then the inner semicontinuity of F at $\bar{x}$ gives us a sequence $\widetilde{y}_k \to \widetilde{y}$ with $\widetilde{y}_k \in F(x_k)$, and so

$$\langle v_k, \widetilde{y}_k - y_k\rangle < 0 \;\text{ for all large }\; k.$$

The obtained contradiction completes the proof of the theorem. △

The following example with two parts shows that both assumptions of Theorem 1.15 are essential for the validity of the extremal property (1.21).

Example 1.16 (Assumptions of Theorem 1.15 Are Essential for the Validity of the Extremal Property).

(i) First we show that the convex-valuedness assumption is essential for the fulfillment of the extremal property (1.21) of inner semicontinuous mappings. Consider the set-valued mapping $F\colon \mathbb{R} \rightrightarrows \mathbb{R}$ defined by

$$F(x) := \big\{ -|x|, |x| \big\} \text{ for } x \in \mathbb{R} \tag{1.22}$$

(see Fig. 1.4), which is clearly nonconvex-valued at any $x \neq 0$ while being inner semicontinuous at $\bar{x} = 0$ due to the equalities

$$\operatorname*{Lim\,inf}_{x \to 0} F(x) = \{0\} = F(0).$$

It is easy to see that the normal cone to the graph of (1.22) at $(0, 0)$ is

$$N\big((0,0); \operatorname{gph} F\big) = \big\{(x,y) \in \mathbb{R}^2 \big|\ y = x\big\} \cup \big\{(x,y) \in \mathbb{R}^2 \big|\ y = -x\big\},$$

and so the coderivative $D^*F(0,0)(v)$ of (1.22) is calculated by

$$D^*F(0,0)(v) = \begin{cases} \{-v, v\} & \text{for } v > 0, \\ 0 & \text{for } v = 0, \\ \{v, -v\} & \text{for } v < 0. \end{cases}$$

It follows from here that for $v = 1 \in \operatorname{dom} D^*F(0,0)$, we have $\langle v, 0\rangle = 0$ while

$$\min_{y \in F(0)} \langle v, y\rangle = \min_{y \in \mathbb{R}} \langle v, y\rangle \neq 0,$$

and thus the extremal property (1.21) fails for F from (1.22).

(ii) Next we demonstrate that property (1.21) may be violated for convex-valued multifunctions, which are not inner semicontinuous at the reference points. Define the convex-valued mapping $F\colon \mathbb{R} \rightrightarrows \mathbb{R}$ by (see Fig. 1.5)

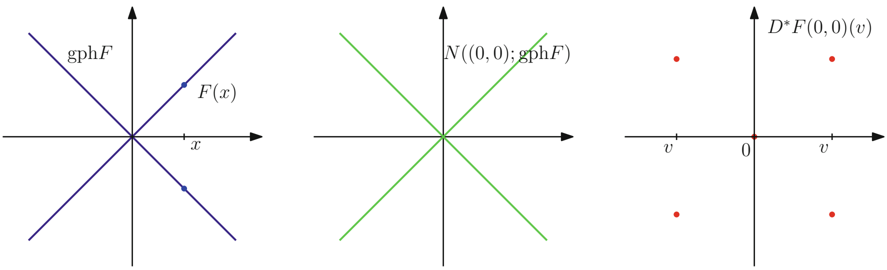


Fig. 1.4 Coderivative of $F(x) := \big\{ -|x|, |x| \big\}$

$$F(x) := \begin{cases} 1 & \text{for } x > 0, \\ [-1, 1] & \text{for } x = 0, \\ -1 & \text{for } x < 0, \end{cases} \tag{1.23}$$

which is not inner semicontinuous at $\bar{x} = 0$ due to

$$\mathop{\text{Lim inf}}_{x \to 0} F(x) = \emptyset \neq F(0) = [-1, 1].$$

Then for the point $(\bar{x}, \bar{y}) = (0, 1) \in \text{gph}\, F$, we have

$$N\big((0, 1); \text{gph}\, F\big) = \big\{(u, v) \in \mathbb{R}^2 \big|\ u \leq 0,\ v \geq 0\big\} \cup \big\{(u, v) \in \mathbb{R}^2 \big|\ uv = 0\big\},$$

which readily implies that dom $D^*F(0, 1) = \mathbb{R}$. Hence

$$\min_{y \in F(0)} vy = -v < v \cdot 1 \ \text{ for any } \ v > 0$$

This shows that the extremal property (1.21) fails for F from (1.23).

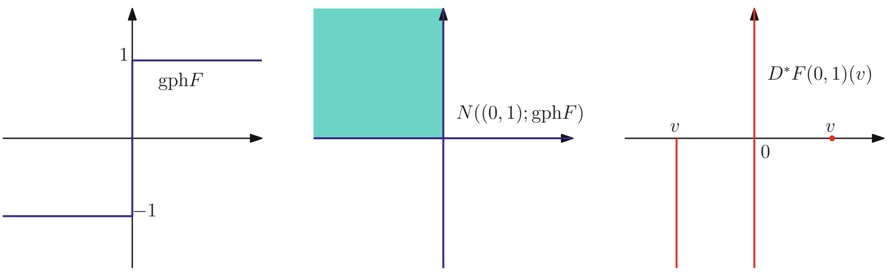


Fig. 1.5 Violation of the extremal property without inner semicontinuity

1.3 First-Order Subgradients of Nonsmooth Functions

This section presents the major first-order subdifferential constructions for extended-real-valued functions mainly used in what follows and then describes some of their fundamental properties and interrelations.

1.3.1 Extended-Real-Valued Functions

In this book we make a terminological distinction between *mappings* and *functions*. By (single-valued or set-valued) mappings, we understand correspondences with values in multidimensional (finite-dimensional or infinite-dimensional) spaces, without any ordering on them. The term "functions" is used for mappings that take real values with the natural order on $\mathbb{R}$. In fact, it is more convenient for various reasons to consider *extended-real-valued* functions, which may take values in the *extended real line* $\overline{\mathbb{R}} := (-\infty, \infty] = \mathbb{R} \cup \{\infty\}$. One of the reasons is to include *sets* into the functional framework by associating a set $\Omega \subset \mathbb{R}^n$ with its *indicator function*

$$\delta(x;\Omega) := \begin{cases} 0 & \text{if } x \in \Omega, \\ \infty & \text{otherwise.} \end{cases}$$

We always suppose that a function $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ is *proper*, i.e.,

$$\operatorname{dom}\varphi := \big\{x \in \mathbb{R}^n \big|\ \varphi(x) < \infty\big\} \neq \emptyset$$

for its *domain*. Note that for definiteness the main attention is paid to "lower" properties of functions largely motivated by applications to *minimization* problems; that's why we exclude the value $-\infty$ from consideration. The "upper" properties and the corresponding upper constructions for φ can be obtained symmetrically by passing to $-\varphi$. We'll do it when it becomes necessary.

From the viewpoint of lower properties, the most appropriate general concept for functions under consideration in variational analysis and optimization is lower semicontinuity, in contrast to continuity in classical analysis. Recall that $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ is *lower semicontinuous* (l.s.c.) at $\bar{x} \in \operatorname{dom}\varphi$ if

$$\varphi(\bar{x}) \le \liminf_{x \to \bar{x}} \varphi(x).$$

Unless otherwise stated, in what follows we consider extended-real-valued functions φ that are *l.s.c. around* the reference point $\bar{x}$, i.e., have this property at any point in some neighborhood of $\bar{x}$. This corresponds to the local closedness of the epigraphical set, or the *epigraph*,

$$\operatorname{epi}\varphi := \big\{(x,\alpha) \in \mathbb{R}^n \times \mathbb{R}\big|\ \alpha \ge \varphi(x)\big\}$$

around the point $(\bar{x}, \varphi(\bar{x})) \in \operatorname{gph}\varphi$. Throughout the book we use the notation

$$x \xrightarrow{\varphi} \bar{x} \Longleftrightarrow x \to \bar{x} \ \text{ with } \ \varphi(x) \to \varphi(\bar{x}),$$

where the condition $\varphi(x) \to \varphi(\bar{x})$ is redundant if φ is continuous at $\bar{x}$. Note that for the indicator function $\varphi(x) = \delta(x;\Omega)$, the notation $x \xrightarrow{\varphi} \bar{x}$ agrees with $x \xrightarrow{\Omega} \bar{x}$ for sets in Section 1.1 and that the lower semicontinuity of φ around $\bar{x} \in \operatorname{dom}\varphi$ reduces to the local closedness of Ω around $\bar{x} \in \Omega$.

1.3.2 Subgradients from Normals to Epigraphs

Similarly to the coderivative case for mappings, we define next the basic and singular limiting subdifferentials (collections of the corresponding subgradients) of extended-real-valued functions geometrically via the basic normals taken from

Definition 1.1. But instead of applying normals to *graphs*, we deal now with *epigraphs* of functions exploiting the natural order structure on $\mathbb{R}$. First we observe the following structure of the normal cone (1.4) to epigraphs.

Proposition 1.17 (Basic Normals to Epigraphs). *Let* $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ *and* $(\bar{x}, \bar{\alpha}) \in \operatorname{epi}\varphi$. *Then* $\lambda \geq 0$ *for every* $(v, -\lambda) \in N((\bar{x}, \bar{\alpha}); \operatorname{epi}\varphi)$, *and so there are uniquely defined subsets* $D, D^\infty \subset \mathbb{R}^n$ *providing the representation*

$$N\big((\bar{x}, \varphi(\bar{x})); \operatorname{epi}\varphi\big) = \big\{\lambda(v, -1)\big|\ v \in D,\ \lambda > 0\big\} \cup \big\{(v, 0)\big|\ v \in D^\infty\big\}.$$

Proof. Taking any $(v, -\lambda) \in N((\bar{x}, \bar{\alpha}); \operatorname{epi}\varphi)$ and using Theorem 1.6, find sequences $(x_k, \alpha_k) \xrightarrow{\operatorname{epi}\varphi} (\bar{x}, \bar{\alpha})$, $v_k \to v$, and $\lambda_k \to \lambda$ such that

$$\limsup_{(x,\alpha)\xrightarrow{\operatorname{epi}\varphi}(x_k,\alpha_k)} \frac{\langle v_k, x - x_k\rangle - \lambda_k(\alpha - \alpha_k)}{\|(x, \alpha) - (x_k, \alpha_k)\|} \leq 0 \quad \text{for all } k \in \mathbb{N}.$$

Letting here $x = x_k$, $\alpha = \alpha_k + 1$ and then passing to the limit as $k \to \infty$, we get $\lambda \geq 0$. This easily implies the claimed representation, where the closedness of the sets D and D^∞ follows from that of the normal cone (1.4). △

The set D in Proposition 1.17 describes "sloping" normals, while D^∞ consists of "horizontal" normals to the epigraph. We define via these sets the basic and singular subdifferentials of the function φ at $\bar{x}$ as follows.

Definition 1.18 (Basic and Singular Subdifferentials of Functions). *Let* $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ *be finite at* $\bar{x} \in \operatorname{dom}\varphi$. *Then the collection of* BASIC SUBGRADIENTS, *or the* (*basic*) SUBDIFFERENTIAL, *of* φ *at* $\bar{x}$ *is defined by*

$$\partial\varphi(\bar{x}) := \big\{v \in \mathbb{R}^n \big|\ (v, -1) \in N\big((\bar{x}, \varphi(\bar{x})); \operatorname{epi}\varphi\big)\big\}. \tag{1.24}$$

The collection of SINGULAR SUBGRADIENTS, *or the* SINGULAR SUBDIFFERENTIAL, *of* φ *at this point is defined by*

$$\partial^\infty\varphi(\bar{x}) := \big\{v \in \mathbb{R}^n \big|\ (v, 0) \in N\big((\bar{x}, \varphi(\bar{x})); \operatorname{epi}\varphi\big)\big\}. \tag{1.25}$$

We'll see below that the subgradient sets (1.24) and (1.25) are much different from each other and play significantly distinct roles in variational analysis and optimization while they enjoy similar and rather comprehensive calculus rules. The basic subdifferential $\partial\varphi(\bar{x})$ reduces to the usual gradient $\{\nabla\varphi(\bar{x})\}$ for smooth functions and to the subdifferential of convex analysis when φ is convex. The singular subdifferential $\partial^\infty\varphi(\bar{x})$ reduces to $\{0\}$ for locally Lipschitzian functions; so it has never appeared in classical analysis and has not been designated in the subdifferential framework of convex analysis as well.

We begin with the extended-real-valued setting of indicator functions when constructions (1.24) and (1.25) agree and reduce to the normal cone (1.4) for the set in question. This easily follows from the definitions and Proposition 1.4.

Proposition 1.19 (Subgradients of Indicator Functions). *For any set $\Omega \subset \mathbb{R}^n$ and point $\bar{x} \in \Omega$, we have the representations*

$$\partial\delta(\bar{x}; \Omega) = \partial^{\infty}\delta(\bar{x}; \Omega) = N(\bar{x}; \Omega).$$

Let us present another property, which is shared by both subdifferential constructions from Definition 1.18 and easily follows from Proposition 1.3.

Proposition 1.20 (Robustness of the Basic and Singular Subdifferentials). *For any $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ and $\bar{x} \in \operatorname{dom}\varphi$, we have*

$$\partial\varphi(\bar{x}) = \operatorname*{Lim\,sup}_{x \xrightarrow{\varphi} \bar{x}} \partial\varphi(x) \quad \text{and} \quad \partial^{\infty}\varphi(\bar{x}) = \operatorname*{Lim\,sup}_{x \xrightarrow{\varphi} \bar{x}} \partial^{\infty}\varphi(x).$$

Next we calculate both basic and singular subdifferentials from Definition 1.18 and illustrate some of their properties for simple functions on $\mathbb{R}$.

Example 1.21 (Subgradients of Simple Functions on $\mathbb{R}$).

(i) Consider first the convex function $\varphi(x) := |x|$. Then we easily see from definition (1.4) or representation (1.9) that

$$N\big((0,0); \operatorname{epi}\varphi\big) = \big\{(x,y) \in \mathbb{R}^2 \big|\ y \le -|x|\big\} \quad \text{and thus} \quad \partial\varphi(0) = [-1, 1]$$

in accordance with convex analysis; see Fig. 1.6. However, changing the sign of the function gives us a completely different picture. Indeed, for $\varphi(x) := -|x|$, the normal cone $N((0,0); \operatorname{epi}\varphi)$ is calculated in Example 1.5, and thus $\partial\varphi(0) = \{-1, 1\}$, i.e., the subdifferential (1.24) is *nonconvex*; see Fig. 1.7. Note that in both cases of $\varphi(x) = |x|$ and $\varphi(x) = -|x|$, we have $\partial^{\infty}\varphi(0) = \{0\}$.

(ii) Next consider the continuous while not Lipschitz continuous function $\varphi(x) := x^{1/3}$ for which we easily get from the definitions that

$$N\big((0,0); \operatorname{epi}\varphi\big) = \big\{(x,0) \in \mathbb{R}^2 \big|\ x \ge 0\big\} \quad \text{with} \quad \partial\varphi(0) = \emptyset,\ \partial^{\infty}\varphi(0) = [0, \infty),$$

which illustrates that the subdifferential (1.24) may be *empty*; see Fig. 1.8.

(iii) If we replace the function in (ii) by $\varphi(x) := x^{1/3}$ for $x < 0$ and $\varphi(x) := 0$ for $x \ge 0$, then

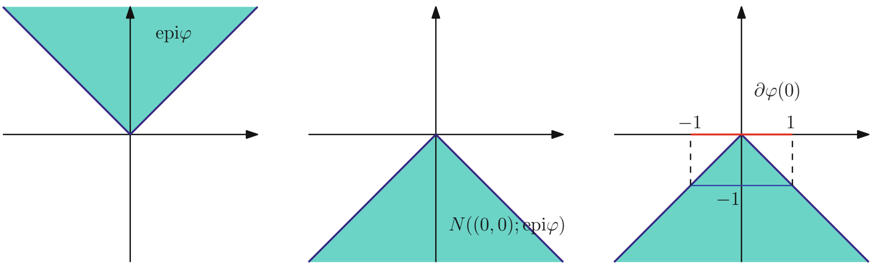


Fig. 1.6 Subdifferential of $\varphi(x) = |x|$

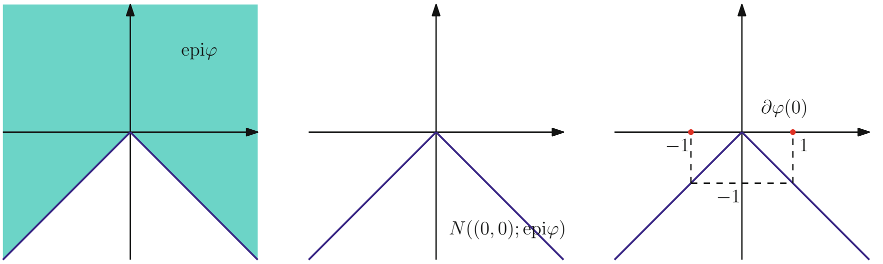


Fig. 1.7 Subdifferential of $\varphi(x) = -|x|$

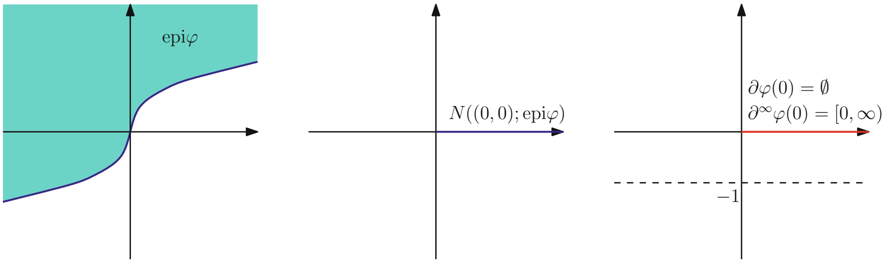


Fig. 1.8 Subdifferential and singular subdifferential of $\varphi(x) = x^{1/3}$

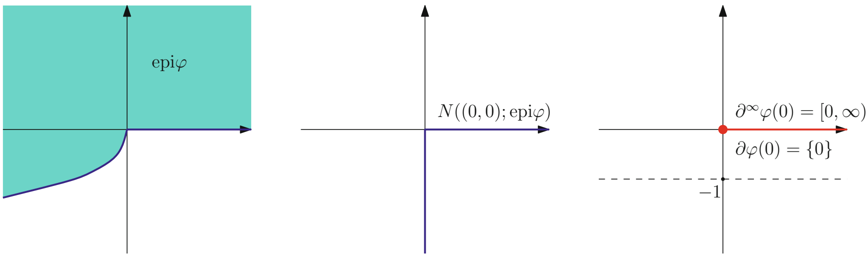


Fig. 1.9 Subdifferential and singular subdifferential of $\varphi(x) = x^{1/3}$ if $x < 0$ and $\varphi(x) = 0$ if $x \geq 0$

$$N\big((0,0); \text{epi}\,\varphi\big) = \big\{(x,0) \in \mathbb{R}^2 \big|\ x \geq 0\big\} \cup \big\{(0,y) \in \mathbb{R}^2 \big|\ y \leq 0\big\}$$

with $\partial\varphi(0) = \{0\}$ and $\partial^\infty\varphi(0) = [0, \infty)$; see Fig. 1.9. This shows, in particular, that the basic subdifferential (1.24) of a continuous function may be a *singleton*, while the function is nonsmooth around the point in question.

(iv) The last example in this vein illustrates yet another, rather opposite feature of the subdifferential (1.24): it may *not be a singleton* for a continuous function that is *differentiable* at the reference point (1.12) but not *strictly* differentiable at it and hence not of class $\mathcal{C}^1$ around this point. Indeed, define $\varphi(x) := x^2 \sin(1/x)$ for $x \neq 0$ and $\varphi(0) := 0$. This function is obviously differentiable at zero with $\varphi'(0) = 0$, while $\partial\varphi(0) = [-1, 1]$.

It is easy to deduce from Proposition 1.2 that we *may* have $\partial\varphi(\bar{x}) = \emptyset$ *only* when $\partial^\infty\varphi(\bar{x}) \neq \{0\}$. Indeed, since $(\bar{x}, \varphi(\bar{x}))$ is a boundary point of the epigraph $\operatorname{epi}\varphi$ which is locally closed around it, there is a nonzero vector $(v, -\lambda) \in N((\bar{x}, \varphi(\bar{x})); \operatorname{epi}\varphi)$. The emptiness of $\partial\varphi(\bar{x})$ implies that $\lambda = 0$, and thus $0 \neq v \in \partial^\infty\varphi(\bar{x})$. Note that the *triviality condition* $\partial^\infty\varphi(\bar{x}) = \{0\}$ is *not necessary* for the nonemptiness of the basic subdifferential $\partial\varphi(\bar{x})$. The latter is always the case for the indicator function in Proposition 1.19 and may also occur when φ is continuous around $\bar{x}$ as demonstrated in Example 1.21(iii).

On the other hand, in the examples given above, the triviality condition $\partial^\infty\varphi(\bar{x}) = \{0\}$ relates to the local Lipschitz continuity of φ around $\bar{x}$. The next theorem shows that it is indeed a *characterization* and describes behavior of the basic subdifferential of locally Lipschitzian functions. Recall that a mapping $f\colon \mathbb{R}^n \to \mathbb{R}^m$ defined near $\bar{x}$ is *locally Lipschitzian* around this point with some *modulus* $\ell \geq 0$ if there is a neighborhood U of $\bar{x}$ such that

$$\|f(x) - f(z)\| \leq \ell\|x - z\| \ \text{ for all } \ x, z \in U. \tag{1.26}$$

Theorem 1.22 (Subdifferentials of Locally Lipschitzian Functions). *Let $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ with $\bar{x} \in \operatorname{dom}\varphi$. Then it is locally Lipschitzian around $\bar{x}$ with some modulus $\ell \geq 0$ if and only if $\partial^\infty\varphi(\bar{x}) = \{0\}$. In this case $\partial\varphi(\bar{x}) \neq \emptyset$ and, for a fixed Lipschitz modulus ℓ, we have*

$$\|v\| \leq \ell \ \textit{ whenever } \ v \in \partial\varphi(\bar{x}). \tag{1.27}$$

Proof. Suppose that φ is Lipschitz continuous on some convex neighborhood U of $\bar{x}$ with modulus ℓ, and show that for any $\lambda \geq 0$, we have the implication

$$(v, -\lambda) \in N\big((\bar{x}, \varphi(\bar{x})); \operatorname{epi}\varphi\big) \Longrightarrow \|v\| \leq \ell\lambda. \tag{1.28}$$

By the normal cone definition (1.4) with the usage of the Euclidean norm on $\mathbb{R}^n \times \mathbb{R}$, it suffices to verify that

$$(w, \mu) \in \Pi\big((x, \alpha); \operatorname{epi}\varphi\big) \Longrightarrow \|w - x\| \leq \ell|\mu - \alpha|$$

for the Euclidean projector $\Pi(\cdot; \operatorname{epi}\varphi)$. Assuming the contrary gives us

$$x \neq w \ \text{ and } \ \gamma := \frac{\|x - w\| - \ell(\mu - \alpha)}{(\ell^2 + 1)\|x - w\|} > 0.$$

Denoting $z := w + \gamma(x - w)$ and $\nu := \mu + \gamma\ell\|x - w\|$, we have that $z \in U$ by the convexity of U. The Lipschitz continuity of φ ensures that $(z, \nu) \in \operatorname{epi}\varphi$. It is not hard to check for the Euclidean norm $\|\cdot\|$ that

$$\|(x, \alpha) - (z, \nu)\| \leq \frac{\ell\|x - w\| + (\mu - \alpha)}{\sqrt{\ell^2 + 1}} < \|(x, \alpha) - (w, \mu)\|,$$

which contradicts the choice of $(w, \mu) \in \Pi((x,\alpha); \operatorname{epi}\varphi)$ and so justifies (1.28). This yields that $\partial^\infty\varphi(\bar{x}) = \{0\}$ for $\lambda = 0$ in (1.28) and that $\|v\| \le \ell$ for $\lambda > 0$.

To complete the proof, it remains to show that the condition $\partial^\infty\varphi(\bar{x}) = \{0\}$ implies that φ is locally Lipschitzian around $\bar{x}$. This follows from the *coderivative criterion* for the Lipschitz-like property of general multifunctions derived in Theorem 3.3; see also Theorem 4.15 for another proof. △

1.3.3 Subgradients from Coderivatives

It is clear from Definition 1.18 that both basic and singular subdifferentials of $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ at $\bar{x} \in \operatorname{dom}\varphi$ can be expressed via the *coderivative*

$$\partial\varphi(\bar{x}) = D^*E_\varphi\big(\bar{x},\varphi(\bar{x})\big)(1), \quad \partial^\infty\varphi(\bar{x}) = D^*E_\varphi\big(\bar{x},\varphi(\bar{x})\big)(0)$$

of the *epigraphical multifunction* $E_\varphi\colon \mathbb{R}^n \rightrightarrows \mathbb{R}$ associate with φ by

$$E_\varphi(x) := \big\{\alpha \in \mathbb{R}\big|\ \alpha \ge \varphi(x)\big\}. \tag{1.29}$$

The next theorem important in what follows shows that, for the class of l.s.c. functions under consideration, we can replace E_φ in the coderivative representation of $\partial\varphi(\bar{x})$ by the *function φ itself*, having also a useful relationship between $\partial^\infty\varphi(\bar{x})$ and $D^*\varphi(\bar{x})(0)$ when φ is continuous around $\bar{x}$.

Theorem 1.23 (Subdifferentials from Coderivatives of l.s.c. and Continuous Functions). *Let $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ be finite around $\bar{x}$. Then we have*

$$\partial\varphi(\bar{x}) = D^*\varphi(\bar{x})(1). \tag{1.30}$$

If in addition φ is continuous around $\bar{x}$, then

$$\partial^\infty\varphi(\bar{x}) \subset D^*\varphi(\bar{x})(0). \tag{1.31}$$

Proof. We split the proof into several steps remembering that φ is *l.s.c. around $\bar{x}$*, which is our standing assumption.

Step 1: *For any sequence $(x_k,\alpha_k) \xrightarrow{\operatorname{epi}\varphi} (\bar{x},\varphi(\bar{x}))$ as $k\to\infty$, there is a subsequence $\{x_{k_j}\}$ of $\{x_k\}$ such that*

$$\big(x_{k_j},\varphi(x_{k_j})\big) \longrightarrow \big(\bar{x},\varphi(\bar{x})\big)\ \text{ as }\ j\to\infty.$$

To proceed, assume first that the set $S := \{x_k|\ \varphi(\bar{x}) \le \varphi(x_k),\ k\in\mathbb{N}\}$ consists of infinitely many elements. By passing to the limit in

$$\varphi(\bar{x}) \le \varphi(x_k) \le \alpha_k \ \text{ for all }\ x_k \in S$$

and taking into account that $\alpha_k \to \varphi(\bar{x})$ as $k \to \infty$, we get $\lim\limits_{x_k \xrightarrow{S} \bar{x}} \varphi(x_k) = \varphi(\bar{x})$ that verifies the claim in this case. In the remaining case where the set S is finite, we suppose without loss of generality that $\varphi(x_k) \le \varphi(\bar{x})$ for all $k \in \mathbb{N}$ and thus get $\limsup\limits_{k\to\infty} \varphi(x_k) \le \varphi(\bar{x})$, which implies in turn that

$$\lim_{k\to\infty} \varphi(x_k) = \varphi(\bar{x})$$

since φ is l.s.c. at $\bar{x}$. It justifies the claim in this case as well.

Step 2: *We have the inclusion* $D^*\varphi(\bar{x})(1) \subset \partial\varphi(\bar{x})$.

This means that the following implication holds:

$$(v, -1) \in N\big((\bar{x}, \varphi(\bar{x})); \operatorname{gph}\varphi\big) \Longrightarrow (v, -1) \in N\big((\bar{x}, \varphi(\bar{x})); \operatorname{epi}\varphi\big).$$

To verify it, pick any $(v, -1) \in N((\bar{x}, \varphi(\bar{x})); \operatorname{gph}\varphi)$ and find by Theorem 1.6 sequences $(v_k, \lambda_k) \to (\bar{x}, -1)$ and $x_k \to \bar{x}$ such that the inclusions $(v_k, \lambda_k) \in \widehat{N}((x_k, \varphi(x_k)); \operatorname{gph}\varphi)$ hold for all $k \in \mathbb{N}$. Suppose without loss of generality that $\lambda_k = -1$ for all $k \in \mathbb{N}$ and show now that $(v_k, -1) \in \widehat{N}((x_k, \varphi(x_k)); \operatorname{epi}\varphi)$ along some subsequence of $\{x_k\}$. In fact, we select this subsequence from Step 1 with no relabeling.

Arguing by contradiction, assume that the claimed inclusion is violated for some fixed k and then find a number $\gamma \in (0, 1)$ and a sequence of pairs $(z_j, \alpha_j) \xrightarrow{\operatorname{epi}\varphi} (x_k, \varphi(x_k))$ as $j \to \infty$ such that

$$\langle v_k, z_j - x_k\rangle + (\varphi(x_k) - \alpha_j) > \gamma\|(z_j, \alpha_j) - (x_k, \varphi(x_k))\| \text{ for all } j \in \mathbb{N}.$$

Since $\alpha_j \ge \varphi(z_j)$ and $\varphi(z_j) \to \varphi(x_k)$ as $j \to \infty$, we have

$$\|(z_j - x_k, \varphi(z_j) - \varphi(x_k))\| \le \|(z_j - x_k, \alpha_j - \varphi(x_k))\| + \alpha_j - \varphi(z_j),$$

which implies in turn the estimate

$$\langle v_k, z_j - x_k\rangle + \varphi(x_k) - \varphi(z_j) > \gamma\|(z_j, \varphi(z_j)) - (x_k, \varphi(x_k))\|$$

for all $j \in \mathbb{N}$. This means that $(v_k, -1) \notin \widehat{N}((x_k, \varphi(x_k)); \operatorname{gph}\varphi)$, which is a contradiction by taking into account the choice of the (sub)sequence $\{x_k\}$ from Step 1. Thus we have the inclusion $D^*\varphi(\bar{x})(1) \subset \partial\varphi(\bar{x})$ in (1.30).

Step 3: *For any set* $\Omega \subset \mathbb{R}^n$ *locally closed around* $\bar{x}$*, we have*

$$N(\bar{x}; \Omega) \subset N(\bar{x}; \operatorname{bd}\Omega) \text{ at every } \bar{x} \in \operatorname{bd}\Omega.$$

To verify this, take $0 \ne v \in N(\bar{x}; \Omega)$ and by Theorem 1.6 find sequences $x_k \xrightarrow{\Omega} \bar{x}$ and $v_k \to v$ with $v_k \in \widehat{N}(x_k; \Omega)$ for all $k \in \mathbb{N}$. Since $\|v_k\| > 0$ when k is sufficiently large, this implies by (1.5) that $x_k \in \operatorname{bd}\Omega$ for such k. The claim now follows from the observation that

$$\widehat{N}(\bar{x};\Omega_1) \subset \widehat{N}(\bar{x};\Omega_2) \ \text{ whenever } \ \Omega_2 \subset \Omega_1 \ \text{ and } \ \bar{x} \in \Omega_2,$$

which can be easily checked by definition (1.5).

Step 4: *We have the inclusion* $\partial\varphi(\bar{x}) \subset D^*\varphi(\bar{x})(1)$.

Since the set $\operatorname{epi}\varphi$ is closed around $(\bar{x},\varphi(\bar{x}))$, it follows from Step 3 that

$$N\big((\bar{x},\varphi(\bar{x}));\operatorname{epi}\varphi\big) \subset N\big((\bar{x},\varphi(\bar{x}));\operatorname{bd}(\operatorname{epi}\varphi)\big),$$

and so it remains to verify the implication

$$\big[(v,-1) \in N\big((\bar{x},\varphi(\bar{x}));\operatorname{bd}(\operatorname{epi}\varphi)\big)\big] \Longrightarrow \big[(v,-1) \in N\big((\bar{x},\varphi(\bar{x}));\operatorname{gph}\varphi\big)\big].$$

To proceed, pick $(v,-1) \in N((\bar{x},\varphi(\bar{x}));\operatorname{bd}(\operatorname{epi}\varphi))$ and find $(v_k,\lambda_k) \to (v,-1)$ and $(x_k,\alpha_k) \stackrel{\operatorname{bd}(\operatorname{epi}\varphi)}{\longrightarrow} (\bar{x},\varphi(\bar{x}))$ as $k \to \infty$ such that $(v_k,-\lambda_k) \in \widehat{N}((x_k,\varphi(x_k));\operatorname{bd}(\operatorname{epi}\varphi))$ whenever $k \in \mathbb{N}$. Let $\lambda_k \equiv -1$ without loss of generality and for all $(x,\alpha) \in \big[B_{1/k}(x_k) \times (\alpha_k - \frac{1}{k}, \alpha_k + \frac{1}{k})\big] \cap \operatorname{bd}(\operatorname{epi}\varphi)$ get

$$\langle v_k, x - x_k\rangle - (\alpha - \alpha_k) \le \frac{1}{k}\Big(\|x - x_k\| + |\alpha - \alpha_k|\Big) \tag{1.32}$$

when k is large. Similarly to Step 2, select by the lower semicontinuity of φ a subsequence of $\{x_k\}$ (no relabeling) such that $(x_k,\varphi(x_k)) \to (\bar{x},\varphi(\bar{x}))$ as $k \to \infty$. Then (1.32) implies that along this subsequence we have

$$\langle v_k, x-x_k\rangle - \big(\alpha-\varphi(x_k)\big) - \big(\varphi(x_k)-\alpha_k\big) \le \frac{1}{k}\Big(\|x-x_k\|+|\alpha-\varphi(x_k)|+|\varphi(x_k)-\alpha_k|\Big)$$

for all $(x,\alpha) \in \big[B_{1/k}(x_k) \times (\alpha_k - r_k, \alpha_k + r_k)\big] \cap \operatorname{bd}(\operatorname{epi}\varphi)$, where such a sequence $r_k \downarrow 0$ exists due to $\alpha_k - \varphi(x_k) \to 0$. By $(x_k,\alpha_k) \in \operatorname{bd}(\operatorname{epi}\varphi) \subset \operatorname{epi}\varphi$ due to the l.s.c. of φ, it yields $\varphi(x_k) \le \alpha_k$ and therefore

$$\langle v_k, x - x_k\rangle - \big(\alpha - \varphi(x_k)\big) \le \frac{1}{k}\Big(\|x - x_k\| + |\alpha - \varphi(x_k)|\Big)$$

for all $(x,\alpha) \in \operatorname{gph}\varphi$ close to $x_k, \varphi(x_k)$. Thus we arrive at

$$\limsup_{(x,\alpha)\stackrel{\operatorname{gph}\varphi}{\longrightarrow}(x_k,\varphi(x_k))} \frac{\langle v_k, x - x_k\rangle - \big(\alpha - \varphi(x_k)\big)}{\|x - x_k\| + |\alpha - \varphi(x_k)|} \le \frac{1}{k},$$

which means that $(v_k,-1) \in \widehat{N}_{\frac{1}{k}}((x_k,\varphi(x_k));\operatorname{gph}\varphi)$ for each k. It tells us by passing to the limit as $k \to \infty$ that $(v,-1) \in N((\bar{x},\varphi(\bar{x}));\operatorname{gph}\varphi)$, i.e., $v \in D^*\varphi(\bar{x})(1)$. This justifies the claim and hence representation (1.30).

Step 5: *If* φ *is continuous around* $\bar{x}$*, then we have inclusion* (1.31).

Indeed, it follows from the continuity of φ around $\bar{x}$ that $\operatorname{gph}\varphi = \operatorname{bd}(\operatorname{epi}\varphi)$. Hence the result of Step 3 ensures the validity of the inclusion

$$N\big((\bar{x},\varphi(\bar{x}));\operatorname{epi}\varphi\big) \subset N\big((\bar{x},\varphi(\bar{x}));\operatorname{gph}\varphi\big),$$

which readily implies (1.31) and completes the proof of the theorem. △

Observe that the inclusion in (1.31) is generally *strict*. To illustrate it, consider the following *example* of a continuous function $\varphi\colon \mathbb{R} \to \mathbb{R}$ with $\varphi(x) := -x^{1/3}$, $x \geq 0$, and $\varphi(x) := 0$, $x < 0$. From Definition 1.1 we calculate the normal cone to the epigraph and graph of this function at the origin by

$$N\big((0,0);\operatorname{epi}\varphi\big) = \big\{(v,0) \in \mathbb{R}^2 \big|\ v \leq 0\big\} \cup \big\{(0,v) \in \mathbb{R}^2 \big|\ v \leq 0\big\}$$

and $N((0,0);\operatorname{gph}\varphi) = N((0,0);\operatorname{epi}\varphi) \cup \mathbb{R}^2_+$; see Fig. 1.10. It shows that $\partial^\infty\varphi(0) = (-\infty,0]$ and $D^*\varphi(0)(0) = (-\infty,\infty)$ with the strict inclusion (1.31).

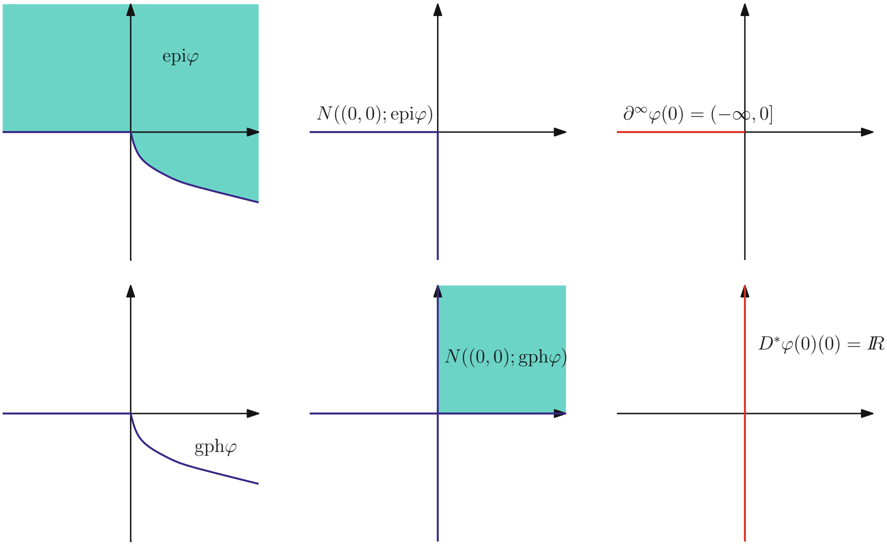


Fig. 1.10 Singular subdifferential vs. coderivative of $\varphi(x) = 0$ if $x < 0$ and $\varphi(x) = -x^{1/3}$ if $x \geq 0$

The precise relationship (1.30) between the coderivative (1.15) and the basic subdifferential (1.24) allows us to deduce subdifferential results from coderivative ones, which is useful in what follows. Let us derive in this way an implementation of Proposition 1.12 in the case of functions.

Corollary 1.24 (Subgradients of Smooth Functions). *Let $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ be of class $\mathcal{C}^1$ around $\bar{x}$. Then we have $\partial\varphi(\bar{x}) = \{\nabla\varphi(\bar{x})\}$.*

Proof. Follows from Theorem 1.23 and Proposition 1.12. △

Note that the reduction of the subgradient set $\partial\varphi(\bar{x})$ to a *singleton* is actually a *characterization* of *strict differentiability* (1.19) for locally *Lipschitzian* functions (1.26); see Theorem 4.17. The elementary functions considered in Example 1.21 demonstrate that both Lipschitz continuity and *strict* differentiability vs. merely differentiability are essential in this characterization.

As shown in Example 1.21(i), the basic subgradient set $\partial\varphi(\bar{x})$ is *nonconvex* for simple functions like $\varphi(x) = -|x|$ at $\bar{x} = 0$. Similarly to the case of normals (and much related to it), we can approximate the subdifferentials (1.24) and (1.25) at $\bar{x}$ by some *convex* sets of subgradients of φ taken at points *nearby*.

1.3.4 Regular Subgradients and ε-Enlargements

Given a function $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ and a point $\bar{x} \in \operatorname{dom}\varphi$, define the collection of *regular subgradients*, or the *presubdifferential*, of φ at $\bar{x}$ by

$$\widehat{\partial}\varphi(\bar{x}) := \Big\{ v \in \mathbb{R}^n \Big|\ \liminf_{x \to \bar{x}} \frac{\varphi(x) - \varphi(\bar{x}) - \langle v, x - \bar{x}\rangle}{\|x - \bar{x}\|} \ge 0 \Big\} \tag{1.33}$$

and for each $\varepsilon > 0$, consider its ε-*enlargement*

$$\widehat{\partial}_\varepsilon\varphi(\bar{x}) := \Big\{ v \in \mathbb{R}^n \Big|\ \liminf_{x \to \bar{x}} \frac{\varphi(x) - \varphi(\bar{x}) - \langle v, x - \bar{x}\rangle}{\|x - \bar{x}\|} \ge -\varepsilon \Big\} \tag{1.34}$$

with $\widehat{\partial}_0\varphi(\bar{x}) = \widehat{\partial}\varphi(\bar{x})$. Note that $\widehat{\partial}\varphi(\bar{x}) = \{\nabla\varphi(\bar{x})\}$ when φ is differentiable (not necessary strictly) at $\bar{x}$ but (1.33) may also reduce to a singleton in the nondifferentiable case, which can be observed from the examples above.

For *convex functions* $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ (i.e., those whose epigraphs are convex sets) , we have the following subgradient descriptions, which show, in particular, that both subgradient sets (1.25) and (1.33) reduce in this case to the classical subdifferential of convex analysis.

Proposition 1.25 (Subgradients and ε-Subgradients of Convex Functions). *Let* $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ *be convex. Then*

$$\widehat{\partial}_\varepsilon\varphi(\bar{x}) = \big\{ v \in \mathbb{R}^n \big|\ \langle v, x - \bar{x}\rangle \le \varphi(x) - \varphi(\bar{x}) + \varepsilon\|x - \bar{x}\| \ \textit{for all}\ \ x \in \mathbb{R}^n \big\}$$

whenever $\bar{x} \in \operatorname{dom}\varphi$ *and* $\varepsilon \ge 0$. *Furthermore, we have representations*

$$\partial\varphi(\bar{x}) = \big\{ v \in \mathbb{R}^n \big|\ \langle v, x - \bar{x}\rangle \le \varphi(x) - \varphi(\bar{x}) \ \textit{for all}\ \ x \in \mathbb{R}^n \big\}. \tag{1.35}$$

$$\partial^\infty\varphi(\bar{x}) = N(\bar{x}; \operatorname{dom}\varphi) = \big\{ v \in \mathbb{R}^n \big|\ \langle v, x - \bar{x}\rangle \le 0 \ \textit{for all}\ \ x \in \operatorname{dom}\varphi \big\}.$$

Proof. Note that the inclusion "⊃" for $\widehat{\partial}_\varepsilon\varphi(\bar{x})$ is obvious. To verify the opposite inclusion, pick an arbitrary subgradient $v \in \widehat{\partial}_\varepsilon\varphi(\bar{x})$ and observe directly from definition (1.34) for any given $\eta > 0$ that the function

$$\vartheta(x) := \varphi(x) - \varphi(\bar{x}) - \langle v, x - \bar{x}\rangle + \big(\varepsilon + \eta\big)\|x - \bar{x}\|$$

attains a local minimum at $\bar{x}$. Since ϑ is convex, $\bar{x}$ is its *global* minimizer, i.e.,

$$\vartheta(x) = \varphi(x) - \varphi(\bar{x}) - \langle v, x - \bar{x}\rangle + \big(\varepsilon + \eta\big)\|x - \bar{x}\| \ge \vartheta(\bar{x}) = 0$$

for all $x \in \mathbb{R}^n$. Taking into account that $\eta > 0$ was chosen arbitrarily, we get the claimed representation of $\widehat{\partial}_\varepsilon\varphi(\bar{x})$ for all $\varepsilon \ge 0$. Furthermore, it follows from the epigraph convexity and the normal cone representation (1.9) that

$$N_{\mathrm{epi}\,\varphi}\big(\bar{x}, \varphi(\bar{x})\big) = \big\{v, \lambda)\big|\ \big\langle(v, \lambda), (x, \alpha) - (\bar{x}, \varphi(\bar{x}))\big\rangle \le 0 \text{ for all } (x, \alpha) \in \mathrm{epi}\,\varphi\big\},$$

which implies by (1.24), (1.25) the formulas for $\partial\varphi(\bar{x})$ and $\partial^\infty\varphi(\bar{x})$. △

It is easy to verify that the sets (1.33) and (1.34) are convex while may be *trivial* for simple nonconvex Lipschitzian functions like $\varphi(x) = -|x|$, where $\widehat{\partial}_\varepsilon\varphi(0) = \emptyset$ for $\varepsilon = 0$ and small $\varepsilon > 0$. On the other hand, we'll see below that these sets considered at points x near $\bar{x}$ can be used for approximating the subdifferential $\partial\varphi(\bar{x})$. Similarly to the case of normals, the regular subgradient collection (1.33) plays a role of the *presubdifferential* in subdifferential theory, along with their ε-subgradient enlargements (1.34). We obviously have

$$\widehat{\partial}_\varepsilon\delta(\bar{x}; \Omega) = \widehat{N}_\varepsilon(\bar{x}; \Omega) \ \text{ whenever } \ \bar{x} \in \Omega \ \text{ and } \ \varepsilon \ge 0$$

for set indicator functions. The next result reveals deeper relationships between (regular) ε-normals and ε-subgradients including the underlying case of $\varepsilon = 0$ most important in what follows. As mentioned above, the norm on $\mathbb{R}^n \times \mathbb{R}$ used in the proof is $\|(x, \alpha)\| = \|x\| + |\alpha|$ by (1.18).

Theorem 1.26 (Geometric Descriptions of Regular Subgradients and Their ε-Enlargements). *Let* $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ *with* $\bar{x} \in \mathrm{dom}\,\varphi$. *Then*

$$\widehat{\partial}_\varepsilon\varphi(\bar{x}) \subset \big\{v \in \mathbb{R}^n\big|\ (v, -1) \in \widehat{N}_\varepsilon\big((\bar{x}, \varphi(\bar{x})); \mathrm{epi}\,\varphi\big)\big\} \ \textit{for all} \ \ \varepsilon \ge 0.$$

Conversely, whenever $0 \le \varepsilon < 1$, *we have the implication*

$$(v, -1) \in \widehat{N}_\varepsilon\big((\bar{x}, \varphi(\bar{x})); \mathrm{epi}\,\varphi\big) \Longrightarrow v \in \widehat{\partial}_{\varepsilon_1}\varphi(\bar{x})$$

with $\varepsilon_1 := \varepsilon(1 + \|v\|)/(1 - \varepsilon)$. *Therefore*

$$\widehat{\partial}\varphi(\bar{x}) = \big\{v \in \mathbb{R}^n\big|\ (v, -1) \in \widehat{N}\big((\bar{x}, \varphi(\bar{x})); \mathrm{epi}\,\varphi\big)\big\}. \tag{1.36}$$

Proof. Pick any $v \in \widehat{\partial}_\varepsilon\varphi(\bar{x})$ and show that $(v, -1) \in \widehat{N}_\varepsilon((\bar{x}, \varphi(\bar{x})); \mathrm{epi}\,\varphi)$ for each $\varepsilon \ge 0$. Indeed, it follows from definition (1.34) that for any $\gamma > 0$ there is a neighborhood U of $\bar{x}$ with

$$\varphi(x) - \varphi(\bar{x}) - \langle v, x - \bar{x}\rangle \ge -(\varepsilon + \gamma)\|x - \bar{x}\| \ \text{ whenever } \ x \in U.$$

This readily gives us for any $x \in U$ and $\alpha \ge \varphi(x)$ that

$$\langle v, x - \bar{x}\rangle + \varphi(\bar{x}) - \alpha \le (\varepsilon + \gamma)\|(x, \alpha) - (\bar{x}, \varphi(\bar{x}))\|,$$

which implies by definition (1.6) with $\varepsilon \geq 0$ for $\Omega = \operatorname{epi} \varphi$ that $v \in \widehat{\partial}_\varepsilon \varphi(\bar{x})$.

To verify the converse implication above, fix $\varepsilon \in [0, 1)$ and assume on the contrary that $v \notin \widehat{\partial}_{\varepsilon_1} \varphi(\bar{x})$ with $\varepsilon_1 \geq 0$ specified in the statement. Then there are $\gamma > 0$ and a sequence $x_k \to \bar{x}$ such that

$$\varphi(x_k) - \varphi(\bar{x}) - \langle v, x_k - \bar{x} \rangle + (\varepsilon_1 + \gamma)\|x_k - \bar{x}\| < 0 \;\text{ for all }\; k \in \mathbb{N}.$$

Letting $\alpha_k := \varphi(\bar{x}) + \langle v, x_k - \bar{x} \rangle - (\varepsilon_1 + \gamma)\|x_k - \bar{x}\|$, observe that $\alpha_k \to \varphi(\bar{x})$ as $k \to \infty$ and that $(x_k, \alpha_k) \in \operatorname{epi} \varphi$ for all $k \in \mathbb{N}$. It implies with the usage of the sum norm (1.18) on the product space that

$$\begin{aligned}\frac{\langle v, x_k - \bar{x} \rangle - \big(\alpha_k - \varphi(\bar{x})\big)}{\|(x_k, \alpha_k) - (\bar{x}, \varphi(\bar{x}))\|} &= \frac{(\varepsilon_1 + \gamma)\|x_k - \bar{x}\|}{\|(x_k - \bar{x}, \langle v, x_k - \bar{x} \rangle - (\varepsilon_1 + \gamma)\|x_k - \bar{x}\|)\|} \\ &\geq \frac{\varepsilon_1 + \gamma}{1 + \|v\| + (\varepsilon_1 + \gamma)} > \frac{\varepsilon_1}{1 + \|v\| + \varepsilon_1} = \varepsilon\end{aligned}$$

for all $k \in \mathbb{N}$ due to $\gamma > 0$ and the choice of ε_1. This clearly implies that $(v, -1) \notin \widehat{N}_\varepsilon((\bar{x}, \varphi(\bar{x})); \operatorname{epi} \varphi)$, which justifies the claimed implication. Representation (1.36) follows by combining the statements above for $\varepsilon = 0$. △

The geometric representation of regular subgradients in (1.36) allows us to deduce their properties from those obtained above for regular normals. The next result establishes in this way a *smooth variational description* of regular subgradients for general extended-real-valued functions.

Theorem 1.27 (Smooth Variational Descriptions of Regular Subgradients). *Let $\varphi: \mathbb{R}^n \to \overline{\mathbb{R}}$ be finite at $\bar{x}$. Then $v \in \widehat{\partial}\varphi(\bar{x})$ if and only if there is a function $\psi: U \to \mathbb{R}$ defined on some neighborhood U of $\bar{x}$ and Fréchet differentiable at $\bar{x}$ such that $\psi(\bar{x}) = \varphi(\bar{x})$, $\nabla\psi(\bar{x}) = v$ and that $\psi(x) - \varphi(x)$ achieves a local maximum on U at $x = \bar{x}$. If furthermore φ is bounded from below on $\mathbb{R}^n$, then we can choose ψ to be concave and smooth on $\mathbb{R}^n$ and such that $\psi(x) - \varphi(x)$ achieves its global maximum on $\mathbb{R}^n$ uniquely at $x = \bar{x}$.*

Proof. The first part of this result follows directly from geometric representation (1.36) of regular normals in Theorem 1.26 and the smooth variational description of regular normals given in Theorem 1.10. To verify the second part, pick any $v \in \widehat{\partial}\varphi(\bar{x})$ and observe that the function

$$\rho(t) := \sup\big\{\varphi(\bar{x}) - \varphi(x) + \langle v, x - \bar{x} \rangle \,\big|\, x \in \bar{x} + t\mathbb{B}\big\}, \quad t \geq 0,$$

satisfies the assumptions of Step 2 in the proof of Theorem 1.10 by the imposed boundedness of φ from below. Having $\tau: [0, \infty) \to [0, \infty)$ constructed therein, we can easily see that the function

$$\psi(x) := -\tau(\|x - \bar{x}\|) - \|x - \bar{x}\|^2 + \varphi(\bar{x}) + \langle v, x - \bar{x} \rangle, \quad x \in \mathbb{R}^n,$$

enjoys all the properties claimed in this corollary. △

1.3.5 Limiting Subdifferential Representations

Next we derive limiting representations of the basic and singular subdifferentials of φ at $\bar{x} \in \operatorname{dom} \varphi$ and present some of their useful consequences.

Theorem 1.28 (Limiting Representations of Basic and Singular Subgradients). *Let $\varphi \colon \mathbb{R}^n \to \overline{\mathbb{R}}$ be finite at $\bar{x}$. Then we have the representations*

$$\partial\varphi(\bar{x}) = \operatorname*{Lim\,sup}_{x \xrightarrow{\varphi} \bar{x}} \widehat{\partial}\varphi(x) = \operatorname*{Lim\,sup}_{\substack{x \xrightarrow{\varphi} \bar{x} \\ \varepsilon \downarrow 0}} \widehat{\partial}_\varepsilon \varphi(x), \tag{1.37}$$

$$\partial^\infty\varphi(\bar{x}) = \operatorname*{Lim\,sup}_{\substack{x \xrightarrow{\varphi} \bar{x} \\ \lambda \downarrow 0}} \lambda\widehat{\partial}\varphi(x) = \operatorname*{Lim\,sup}_{\substack{x \xrightarrow{\varphi} \bar{x} \\ \lambda,\, \varepsilon \downarrow 0}} \lambda\widehat{\partial}\varphi(x). \tag{1.38}$$

Proof. We begin by verifying that the subgradient set $\partial\varphi(\bar{x})$ belongs to the first limit in (1.37) while observing that the inclusion "$\subset$" in the second representation of (1.37) is obvious. Pick any $v \in \partial\varphi(\bar{x})$ and get by definition (1.24) that $(v, -1) \in N((\bar{x}, \varphi(\bar{x})); \operatorname{epi} \varphi)$. Then by the first representation of the normal cone in Theorem 1.6, we find sequences $(x_k, \alpha_k) \xrightarrow{\operatorname{epi} \varphi} (\bar{x}, \varphi(\bar{x}))$ and $(v_k, -\lambda_k) \to (v, -1)$ as $k \to \infty$ such that

$$(v_k, -\lambda_k) \in \widehat{N}\big((x_k, \alpha_k); \operatorname{epi} \varphi\big) \quad \text{for all} \quad k \in \mathbb{N}. \tag{1.39}$$

Suppose without loss of generality that $\lambda_k = 1$ for all k and get $\alpha_k = \varphi(x_k)$ by Exercise 1.62. Then we have from (1.36) that $v_k \in \widehat{\partial}\varphi(x_k)$, which means by (1.1) that the vector v belongs to the outer limit $\operatorname{Lim\,sup} \widehat{\partial}\varphi(x)$ as $x \xrightarrow{\varphi} \bar{x}$.

To proceed further with the proof of (1.37), take v from the rightmost set therein and find some sequences $\varepsilon_k \downarrow 0$, $x_k \xrightarrow{\varphi} \bar{x}$, and $v_k \to v$ satisfying

$$v_k \in \widehat{\partial}_{\varepsilon_k} \varphi(x_k) \quad \text{for all} \quad k \in \mathbb{N}.$$

For any k we get from the first inclusion in Theorem 1.26 that

$$(v_k, -1) \in \widehat{N}_{\varepsilon_k}\big((x_k, \varphi(x_k)); \operatorname{epi} \varphi\big), \qquad k \in \mathbb{N}.$$

Passing now to the limit as $k \to \infty$ gives us by Theorem 1.6 the inclusion $(v, -1) \in N((\bar{x}, \varphi(\bar{x})); \operatorname{epi} \varphi)$, which ensures by (1.24) that $v \in \partial\varphi(\bar{x})$ and thus completes the proof of both representations in (1.37).

To justify the first singular subdifferential representation in (1.38), pick v from the set on the right-hand side therein and find by definition (1.1) sequences $\lambda_k \downarrow 0$, $x_k \xrightarrow{\operatorname{epi} \varphi} \bar{x}$, and $v_k \to v$ as $k \to \infty$ such that $v_k \in \lambda_k \widehat{\partial}\varphi(x_k)$ for all $k \in \mathbb{N}$. This implies by (1.33) and the conic structure of $\widehat{N}(\cdot; \operatorname{epi} \varphi)$ that we have (1.39) with $\alpha_k = \varphi(x_k)$, which ensures by passing to the limit as $k \to \infty$ that $v \in \partial^\infty\varphi(\bar{x})$ due to definition (1.25).

To proceed with verifying the opposite inclusion in (1.38), pick $v \in \partial^\infty\varphi(\bar{x})$ and get $(v,0) \in N((\bar{x},\varphi(\bar{x}));\text{epi}\,\varphi)$. Then Theorem 1.6 generates sequences $(x_k,\alpha_k) \xrightarrow{\text{epi}\,\varphi} (\bar{x},\varphi(\bar{x}))$ and $(v_k,\lambda_k) \to (v,0)$ as $k \to \infty$ such that the inclusions in (1.39) hold. We can put $\alpha_k = \varphi(x_k)$ in (1.39) and easily see as in Proposition 1.17 that $\lambda_k \geq 0$ for all $k \in \mathbb{N}$. There are two cases to consider: either **(a)** $\lambda_k > 0$ or **(b)** $\lambda_k = 0$ along some subsequence of $k \to \infty$. In case (a) we have $v_k \in \lambda_k\widehat{\partial}\varphi(x_k)$ and thus conclude that v belongs to the outer limit on the right-hand side of (1.38). Case (b) reduces to (a) by showing that in this case the sequence $\{v_k\}$ can be slightly adjusted so that there are $(\widetilde{v}_k,-\widetilde{\lambda}_k) \in \widehat{N}((x_k,\varphi(x_k));\text{epi}\,\varphi)$ with $\widetilde{\lambda}_k \downarrow 0$ and $\widetilde{v}_k \to v$ as $k \to \infty$. The proof of this adjustment is technically involved and is omitted here; see [678, Theorem 8.9] and [522, Lemma 2.37] for different detailed arguments. The second representation in (1.38) is justified similarly to the case of (1.37). △

Note that the second representations in (1.37) and (1.38) justify the *stability* of the limiting representation of $\partial\varphi(\bar{x})$ with respect to the presubdifferential enlargement. Such a stability is clearly related to that in the normal cone representations of Theorem 1.6. Let us demonstrate the importance of it in the proof of the following useful property of singular subgradients.

Proposition 1.29 (Singular Subgradients Under Lipschitzian Additions). *Let* $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ *be finite at* $\bar{x} \in \text{dom}\,\varphi$*, and let* $\psi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ *be locally Lipschitzian around this point. Then*

$$\partial^\infty(\varphi+\psi)(\bar{x}) = \partial^\infty\varphi(\bar{x}).$$

Proof. Given $v \in \partial^\infty(\varphi+\psi)(\bar{x})$, find by definition (1.25) sequences $\gamma_k \downarrow 0$, $(x_k,\alpha_k) \xrightarrow{\text{epi}(\varphi+\psi)} (\bar{x},(\varphi+\psi)(\bar{x}))$, $v_k \to v$, $\nu_k \to 0$, and $\eta_k \downarrow 0$ such that

$$\langle v_k, x-x_k\rangle + \nu_k(\alpha-\alpha_k) \leq \gamma_k(\|x-x_k\|+|\alpha-\alpha_k|)$$

for all $(x,\alpha) \in \text{epi}\,(\varphi+\psi)$ with $x \in x_k+\eta_k\mathbb{B}$ and $|\alpha-\alpha_k| \leq \eta_k$ as $k \in \mathbb{N}$. Take a Lipschitz constant $\ell > 0$ of ψ around $\bar{x}$ from (1.26) and denote $\widetilde{\eta}_k := \eta_k/2(\ell+1)$ and $\widetilde{\alpha}_k := \alpha_k - \psi(x_k)$. Then $(x_k,\widetilde{\alpha}_k) \xrightarrow{\text{epi}\varphi} (\bar{x},\varphi(\bar{x}))$ and

$$(x,\alpha+\psi(x)) \in \text{epi}\,(\varphi+\psi), \quad |(\alpha+\psi(x))-\alpha_k| \leq \eta_k$$

whenever $(x,\alpha) \in \text{epi}\,\varphi$, $x \in x_k+\widetilde{\eta}_k\mathbb{B}$, and $|\alpha-\widetilde{\alpha}_k| \leq \widetilde{\eta}_k$. Hence

$$\langle v_k, x-x_k\rangle + \nu_k(\alpha-\widetilde{\alpha}_k) \leq \varepsilon_k(\|x-x_k\|+|\alpha-\widetilde{\alpha}_k|) \text{ with } \varepsilon_k := \gamma_k(1+\ell)+|\nu_k|\ell$$

for any $(x,\alpha) \in \text{epi}\,\varphi$ with $x \in x_k+\widetilde{\eta}_k\mathbb{B}$ and $|\alpha-\widetilde{\alpha}_k| \leq \widetilde{\eta}_k$. This yields $(v_k,\nu_k) \in \widehat{N}_{\varepsilon_k}((x_k,\widetilde{\alpha}_k);\text{epi}\,\varphi)$ for all $k \in \mathbb{N}$, and so $(v,0) \in N((\bar{x},\varphi(\bar{x}));\text{epi}\,\varphi)$ since $\varepsilon_k \downarrow 0$ as $k \to \infty$. Thus we get the inclusion "$\subset$" in the statement above. Applying it to the sum $\varphi = (\varphi+\psi)+(-\psi)$ gives us $\partial^\infty\varphi(\bar{x}) \subset \partial^\infty(\psi+\varphi)(\bar{x})$, which justifies the claimed equality and completes the proof. △

It is easy to observe that the *convex* set of regular subgradients (1.33) for any extended-real-valued function φ admits the *dual representation*

$$\widehat{\partial}\varphi(\bar{x}) = \left\{ v \in \mathbb{R}^n \,\middle|\, \langle v, w\rangle \le d\varphi(\bar{x}; w) \;\text{ for all }\; w \in \mathbb{R}^n \right\} \tag{1.40}$$

via the *contingent derivative* of φ at $\bar{x} \in \operatorname{dom}\varphi$ in the *direction* w defined by

$$d\varphi(\bar{x}; w) := \inf\left\{ \nu \in \mathbb{R} \,\middle|\, (w, \nu) \in T\big((\bar{x}, \varphi(\bar{x})); \operatorname{epi}\varphi\big) \right\} \tag{1.41}$$

geometrically in terms of the contingent cone (1.11) to the epigraph. This is similar to the duality relationship between the prenormal and contingent cones to closed sets in Proposition 1.9 It follows directly from the definitions that $\operatorname{epi} d\varphi(\bar{x}; \cdot) = T_{\operatorname{epi}\varphi}(\bar{x}, \varphi(\bar{x}))$ and that $d\varphi(\bar{x}; w)$ can be described analytically via the lower limit of difference quotient

$$d\varphi(\bar{x}; w) = \liminf_{\substack{z \to w \\ t \downarrow 0}} \frac{\varphi(\bar{x} + tz) - \varphi(\bar{x})}{t}. \tag{1.42}$$

Observe that we can equivalently let $z = w$ in (1.42) if φ is locally Lipschitzian around $\bar{x}$. Note also that our basic subdifferential (1.24), being *nonconvex*, cannot be generated in the duality scheme of type (1.40) by *any* directional derivative. On the other hand, the approximation results of Theorem 1.28 show that it can be done in the *limiting procedure*.

We'll see in Chapters 2–4 that, in spite of (actually due to) their nonconvexity, the basic and singular subdifferentials as well as the normal cone and coderivative associated with them enjoy comprehensive calculus rules and other properties crucial for applications, while their regular counterparts are inadequate in themselves for a satisfactory theory and applications.

Let us first present some simple albeit important properties that are shared by basic and regular subgradients.

Proposition 1.30 (Elementary Rules for Basic and Regular Subgradients). *Let $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ be finite at $\bar{x}$. The following assertions hold:*

(i) (GENERALIZED FERMAT RULE) *If $\bar{x}$ is a local minimizer of φ, then*

$$0 \in \widehat{\partial}\varphi(\bar{x}) \;\; \textit{and} \;\; 0 \in \partial\varphi(\bar{x}).$$

These conditions agree and are sufficient for global minima when φ is convex.

(ii) (SUM RULES WITH DIFFERENTIABLE ADDITIONS) *Let $\psi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ be differentiable at $\bar{x}$. Then we have*

$$\widehat{\partial}(\psi + \varphi)(\bar{x}) = \nabla\psi(\bar{x}) + \widehat{\partial}\varphi(\bar{x}).$$

If furthermore ψ is of class $\mathcal{C}^1$ around this point, then

$$\partial(\psi + \varphi)(\bar{x}) = \nabla\psi(\bar{x}) + \partial\varphi(\bar{x}).$$

Proof. When $\bar{x}$ is a local minimizer of φ, we get directly from definition (1.33) that $v = 0$ is a regular subgradient of φ. The second inclusion in (i) follows from that of $\widehat{\partial}\varphi(\bar{x}) \subset \partial\varphi(\bar{x})$, which in turn is a consequence of representation (1.37) in Theorem 1.28. If φ is convex, the sets $\widehat{\partial}\varphi(\bar{x})$ and $\partial\varphi(\bar{x})$ agree with each other by Proposition 1.25, and the condition $0 \in \partial\varphi(\bar{x})$ ensures that $\bar{x}$ is a global minimizer of φ by the subdifferential representation therein.

The inclusion "$\subset$" in the rule for $\widehat{\partial}(\psi + \varphi)(\bar{x})$ is verified directly by definition. The opposite one follows from it by applying to $\varphi = (\psi + \varphi) + (-\psi)$. To obtain the sum rule for basic subgradients, we pass to the limit from its regular counterpart at points nearby with the usage of Theorem 1.28. △

The limiting representation (1.37) of basic subgradients via regular ones is convenient for their calculations in multidimensional spaces. The next example illustrates this for two Lipschitz continuous functions on $\mathbb{R}^2$.

Example 1.31 (Subdifferential Calculations for Lipschitzian Functions).
(i) Consider first the function $\varphi\colon \mathbb{R}^2 \to \mathbb{R}$ defined by

$$\varphi(x_1, x_2) := |x_1| - |x_2| \;\text{ for }\; (x_1, x_2) \in \mathbb{R}^2,$$

which is Lipschitz continuous on $\mathbb{R}^2$ and differentiable at every $(x_1, x_2) \in \mathbb{R}^2$ with both nonzero components x_1, x_2. We have

$$\nabla\varphi(x_1, x_2) \in \big\{(1, 1),\ (1, -1),\ (-1, 1),\ (-1, -1)\big\}$$

for all such (x_1, x_2). It is easy to calculate regular subgradients of φ at any $(x_1, x_2) \in \mathbb{R}^2$ by definition (1.33):

$$\widehat{\partial}\varphi(x_1, x_2) = \begin{cases} (1, -1) & \text{if } x_1 > 0,\ x_2 > 0, \\ (-1, -1) & \text{if } x_1 < 0,\ x_2 > 0, \\ (-1, 1) & \text{if } x_1 < 0,\ x_2 < 0, \\ (1, 1) & \text{if } x_1 > 0,\ x_2 < 0, \\ \big\{(v, -1)\big|\, -1 \le v \le 1\big\} & \text{if } x_1 = 0,\ x_2 > 0, \\ \big\{(v, 1)\big|\, -1 \le v \le 1\big\} & \text{if } x_1 = 0,\ x_2 < 0, \\ \emptyset & \text{if } x_2 = 0. \end{cases}$$

Employing Theorem 1.28 gives us the basic subdifferential (see Fig. 1.11)

$$\partial\varphi(0, 0) = \big\{(v, 1)\big|\, -1 \le v \le 1\big\} \cup \big\{(v, -1)\big|\, -1 \le v \le 1\big\}.$$

(ii) Consider next the more complicated function:

$$\varphi(x_1, x_2) := \big|\, |x_1| + x_2 \big| \;\text{ for }\; (x_1, x_2) \in \mathbb{R}^2,$$

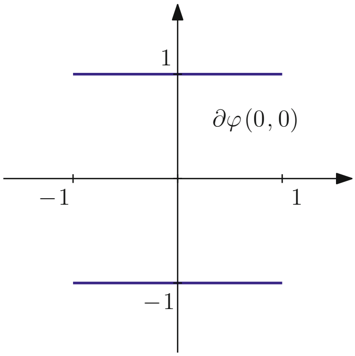


Fig. 1.11 Basic subdifferential of $\varphi(x_1, x_2) = |x_1| - |x_2|$

which is also Lipschitz continuous on $\mathbb{R}^2$. Based on their definition (1.33), we calculate regular subgradients of φ at any $x \in \mathbb{R}^2$ by

$$\widehat{\partial}\varphi(x_1, x_2) = \begin{cases} (1, 1) & \text{if } x_1 > 0,\ x_1 + x_2 > 0, \\ (-1, -1) & \text{if } x_1 > 0,\ x_1 + x_2 < 0, \\ (-1, 1) & \text{if } x_1 < 0,\ x_1 - x_2 < 0, \\ (1, -1) & \text{if } x_1 < 0,\ x_1 - x_2 > 0, \\ \big\{(v, 1)\big| -1 \le v \le 1\big\} & \text{if } x_1 = 0,\ x_2 > 0, \\ \big\{(v, v)\big| -1 \le v \le 1\big\} & \text{if } x_1 > 0,\ x_1 + x_2 = 0, \\ \big\{(v, -v)\big| -1 \le v \le 1\big\} & \text{if } x_1 < 0,\ x_1 - x_2 = 0, \\ \big\{(v_1, v_2)\big|\ |v_1| \le v_2 \le 1\big\} & \text{if } x_1 = 0,\ x_2 = 0, \\ \emptyset & \text{if } x_1 = 0,\ x_2 < 0. \end{cases}$$

By Theorem 1.28 we then calculate (see Fig. 1.12)

$$\partial\varphi(0, 0) = \big\{(v_1, v_2)\big|\ |v_1| \le v_2 \le 1\big\} \cup \big\{(v_1, v_2)\big|\ v_2 = -|v_1|, -1 \le v_1 \le 1\big\}.$$

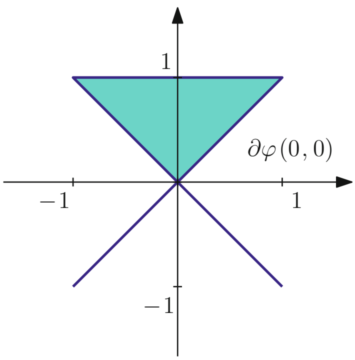


Fig. 1.12 Basic subdifferential of $\varphi(x_1, x_2) = \big|\ |x_1| + x_2\big|$.

Now we apply both representations of basic subgradients in (1.37) of Theorem 1.28 to derive the important *scalarization formula* for expressing the coderivative (1.15) of a single-valued Lipschitzian mapping $f\colon \mathbb{R}^n \to \mathbb{R}^m$ via the subdifferential (1.24) of the scalarization $\langle v, f\rangle(x) := \langle v, f(x)\rangle$, $x \in \mathbb{R}^n$.

Theorem 1.32 (Coderivative Scalarization). *Let $f\colon \mathbb{R}^n \to R^m$ be continuous around $\bar{x}$. Then we have the inclusion*

$$\partial\langle v, f\rangle(\bar{x}) \subset D^*f(\bar{x})(v) \ \textit{ for all } \ v \in \mathbb{R}^m.$$

If in addition f is locally Lipschitzian around $\bar{x}$, then

$$D^*f(\bar{x})(v) = \partial\langle v, f\rangle(\bar{x}) \ \textit{ for all } \ v \in \mathbb{R}^m.$$

Proof. Picking any $u \in \partial\langle v, f\rangle(\bar{x})$ and using the first representation in (1.37) give us sequences $x_k \to \bar{x}$ and $u_k \to u$ such that $u_k \in \widehat{\partial}\langle v, f\rangle(x_k)$ for $k \in \mathbb{N}$. By definition (1.33) for each k, there exist a neighborhood U_k of x_k and a number $\gamma_k > 0$ satisfying the inequality

$$\langle v, f\rangle(x) - \langle v, f\rangle(x_k) - \langle u_k, x - x_k\rangle \ge -\gamma_k\|x - x_k\| \ \text{ when } \ x \in U_k,$$

which ensures in turn the relationship

$$\limsup_{x \to x_k} \frac{\langle u_k, x - x_k\rangle - \langle v, f(x) - f(x_k)\rangle}{\|(x - x_k, f(x) - f(x_k))\|} \le \gamma_k.$$

Hence $(u_k, -v) \in \widehat{N}_{\gamma_k}((x_k, f(x_k)); \operatorname{gph} f)$ for each $k \in \mathbb{N}$, which gives us $u \in D^*f(\bar{x})(v)$ by Theorem 1.6 and the coderivative definition (1.15).

To prove the opposite inclusion, pick $u \in D^*f(\bar{x})(v)$ and by Theorem 1.6 find $x_k \to \bar{x}$, $u_k \to u$, and $v_k \to v$ such that $(u_k, -v_k) \in \widehat{N}(x_k, f(x_k)); \operatorname{gph} f)$ for $k \in \mathbb{N}$. Hence there exist $\eta_k \downarrow 0$ and $\gamma_k \downarrow 0$ with

$$\langle u_k, x - x_k\rangle - \langle v_k, f(x) - f(x_k)\rangle \le \gamma_k(1 + \ell)\|x - x_k\| \ \text{ for all } \ x \in x_k + \eta_k\mathbb{B},$$

where $\ell > 0$ is a Lipschitz modulus (1.26) for f around $\bar{x}$. This yields

$$u_k \in \widehat{\partial}_{\varepsilon_k}\langle v, f\rangle(x_k) \ \text{ with } \ \varepsilon_k := \gamma_k(1 + \ell) + \ell\|v_k - v\| \downarrow 0$$

and gives us by Theorem 1.28 that $u \in \partial\langle v, f\rangle(\bar{x})$. △

1.3.6 Subgradients of the Distance Function

We conclude this section with calculating the basic subdifferential of the *distance function* $d_\Omega(x)$ from (1.2) associated with a nonempty (locally closed) set Ω. This function is intrinsically nonsmooth while being globally Lipschitzian on $\mathbb{R}^n$ with modulus $\ell = 1$. Subdifferential properties of d_Ω at the given point $\bar{x}$ depend on the location of $\bar{x}$: either *in-set* $\bar{x} \in \Omega$ or *out-of-set* $\bar{x} \notin \Omega$. The following theorem presents formulas for calculating regular and basic subgradients at both in-set and out-of-set points.

Theorem 1.33 (Subdifferentiation of the Distance Function at In-Set and Out-of-Set Points). *For the distance function $d_\Omega(x)$, the following hold:*

(i) *If $\bar{x} \in \Omega$, then we have that*

$$\widehat{\partial} d_\Omega(\bar{x}) = \widehat{N}_\Omega(\bar{x}) \cap \mathbb{B} \;\; \text{and} \;\; \partial d_\Omega(\bar{x}) = N_\Omega(\bar{x}) \cap \mathbb{B}. \tag{1.43}$$

(ii) *If $\bar{x} \notin \Omega$, then we have via the Euclidean projector Π_Ω that*

$$\widehat{\partial} d_\Omega(\bar{x}) = \begin{cases} \dfrac{\bar{x} - \bar{w}}{\|\bar{x} - \bar{w}\|} & \text{if } \Pi_\Omega(\bar{x}) = \{\bar{w}\}, \\ \emptyset & \text{otherwise}; \end{cases} \qquad \partial d_\Omega(\bar{x}) = \frac{\bar{x} - \Pi_\Omega(\bar{x})}{d_\Omega(\bar{x})}. \tag{1.44}$$

Proof. We split the proof into several major steps of their own interest.
Step 1: *For any $\bar{x} \in \Omega$, the first formula in* (1.43) *holds.*

Indeed, picking $v \in \widehat{\partial} d_\Omega(\bar{x})$ with $\bar{x} \in \Omega$ gives us by (1.33) that

$$0 \le \liminf_{x \xrightarrow{\Omega} \bar{x}} \frac{d_\Omega(x) - d_\Omega(\bar{x}) - \langle v, x - \bar{x}\rangle}{\|x - \bar{x}\|} = -\limsup_{x \xrightarrow{\Omega} \bar{x}} \frac{\langle v, x - \bar{x}\rangle}{\|x - \bar{x}\|}, \tag{1.45}$$

which shows by (1.6) that $v \in \widehat{N}_\Omega(\bar{x})$. Furthermore, the Lipschitz continuity of d_Ω with constant $\ell = 1$ immediately implies that

$$\limsup_{x \to \bar{x}} \frac{\langle v, x - \bar{x}\rangle}{\|x - \bar{x}\|} \le 1, \quad \text{i.e.,} \quad \|v\| \le 1,$$

and so $\widehat{\partial} d_\Omega(\bar{x}) \subset \widehat{N}_\Omega(\bar{x}) \cap \mathbb{B}$. To justify the opposite inclusion, take any $v \in \widehat{N}_\Omega(\bar{x}) \cap \mathbb{B}$ and observe from (1.45) that it remains to consider the underlying "liminf" therein for $x \to \bar{x}$ with $x \notin \Omega$. To proceed, fix $x \notin \Omega$ with $d_\Omega(x) > 0$ and find $u \in \Omega$ such that

$$0 < \|x - u\| \le d_\Omega(x) + \|x - \bar{x}\|^2.$$

Then for any x sufficiently close to $\bar{x}$, we have

$$\|u - \bar{x}\| \le \|x - u\| + \|x - \bar{x}\| \le d_\Omega(x) + \|x - \bar{x}\|^2 \le 3\|x - \bar{x}\|. \tag{1.46}$$

This, with taking the estimates in (1.46) and $\|v\| \le 1$ into account, allows us to derive the following chain of inequalities:

$$\liminf_{\substack{x\to\bar{x}\\ x\notin\Omega}}\frac{d_\Omega(x)-d_\Omega(\bar{x})-\langle v,x-\bar{x}\rangle}{\|x-\bar{x}\|}\ge\liminf_{\substack{x\to\bar{x}\\ x\notin\Omega}}\frac{\|x-u\|-\|x-\bar{x}\|^2-\langle v,x-\bar{x}\rangle}{\|x-\bar{x}\|}$$

$$\ge\liminf_{\substack{x\to\bar{x}\\ x\notin\Omega}}\Big[\frac{(1-\|v\|)\cdot\|x-u\|}{\|x-\bar{x}\|}-\frac{\langle v,u-\bar{x}\rangle}{\|x-\bar{x}\|}\Big]$$

$$\ge-\limsup_{\substack{x\to\bar{x}\\ x\notin\Omega}}\frac{\langle v,u-\bar{x}\rangle}{\|x-\bar{x}\|}\ge\min\Big\{0,-\limsup_{u\xrightarrow{\Omega}\bar{x}}\frac{3\langle v,u-\bar{x}\rangle}{\|u-\bar{x}\|}\Big\}\ge 0,$$

Together with the equality in (1.45), it shows that $\widehat{N}_\Omega(\bar{x})\cap\mathbb{B}\subset\widehat{\partial}d_\Omega(\bar{x})$ and thus ensures the validity of the first formula in (1.43).

Step 2: *For any $\bar{x}\notin\Omega$ and $\bar{w}\in\Pi_\Omega(\bar{x})$, we have the inclusion*

$$\widehat{\partial}d_\Omega(\bar{x})\subset\widehat{N}_\Omega(\bar{w})\cap\mathbb{B}.$$

To verify it, pick $v\in\widehat{\partial}d_\Omega(\bar{x})$ and deduce from Theorem 1.22 that $\|v\|\le 1$. It follows from the definitions that for any $\gamma>0$ there is $\nu>0$ such that

$$\langle v,x-\bar{x}\rangle\le d_\Omega(x)-d_\Omega(\bar{x})+\gamma\|x-\bar{x}\|=d_\Omega(x)-\|\bar{x}-\bar{w}\|+\gamma\|x-\bar{x}\|$$

if $\|x-\bar{x}\|<\nu$. Fix $w\in\Omega$ with $\|w-\bar{w}\|<\nu$ and observe by using $\|(w-\bar{w}+\bar{x})-\bar{x}\|<\nu$ and $d_\Omega(w-\bar{w}+\bar{x})\le\|w-\bar{w}+\bar{x}-w\|=\|\bar{x}-\bar{w}\|$ that

$$\langle v,w-\bar{w}\rangle\le d_\Omega(w-\bar{w}+\bar{x})-\|\bar{x}-\bar{w}\|+\gamma\|w-\bar{w}\|\le\gamma\|w-\bar{w}\|,$$

which shows that $v\in\widehat{N}_\Omega(\bar{w})$ and thus justified the claimed inclusion.

Step 3: *For any $\bar{x}\in\Omega$ the second formula in (1.43) holds.*

Indeed, take $v\in\partial d_\Omega(\bar{x})$ and find by the subdifferential construction some sequences $x_k\to\bar{x}$, $v_k\to v$ with $v_k\in\widehat{\partial}d_\Omega(x_k)$ as $k\in\mathbb{N}$. Picking $w_k\in\Pi_\Omega(x_k)$ for large k, we get that $w_k\to\bar{x}$ and $v_k\in\widehat{N}_\Omega(w_k)\cap\mathbb{B}$ by Step 2. It tells us that $v\in N_\Omega(\bar{x})\cap\mathbb{B}$ by passing to the limit as $k\to\infty$. To prove next the opposite inclusion, fix any $v\in N_\Omega(\bar{x})\cap\mathbb{B}$ and find sequences $x_k\to\bar{x}$ and $v_k\to v$ such that $x_k\in\Omega$ and $v_k\in\widehat{N}_\Omega(x_k)$ for all $k\in\mathbb{N}$. Define

$$w_k:=\frac{v_k}{\max\big\{\|v_k\|,1\big\}},\quad k\in\mathbb{N},$$

and observe that $w_k\in\mathbb{B}$, $w_k\in\widehat{N}_\Omega(x_k)$, and thus $w_k\in\widehat{\partial}d_\Omega(x_k)$ by Step 1. Since the sequence $\{w_k\}$ also converges to v, we get $v\in\partial d_\Omega(\bar{x})$ and therefore complete the proof of assertion (i) of the theorem.

Step 4: *For any $\bar{x}\notin\Omega$ the contingent derivative of the distance function $\varphi(x)=d_\Omega(x)$ at $\bar{x}$ in the direction $z\in\mathbb{R}^n$ admits the representation*

$$d\varphi(\bar{x})(z)=\min\Big\{\frac{\langle\bar{x}-\bar{w},z\rangle}{\|\bar{x}-\bar{w}\|}\,\Big|\,\bar{w}\in\Pi_\Omega(\bar{w})\Big\}.\tag{1.47}$$

To verify this, we use the equivalent representation

$$d\varphi(\bar{x}; z) = \liminf_{t \downarrow 0} \frac{\varphi(\bar{x} + tz) - \varphi(\bar{x})}{t} \tag{1.48}$$

of the contingent derivative (1.41) of a locally Lipschitzian function that easily follows from the analytic description (1.42). For $\varphi(x) = d_\Omega(x)$ and any projection $\bar{w} \in \Pi_\Omega(\bar{x})$, we deduce from (1.48) and the differentiability of the norm function $\psi(x) := \|x\|$ at $\bar{x} \neq 0$ with $\nabla\psi(x) = \frac{x}{\|x\|}$ that

$$d\varphi(\bar{x}; z) \leq \liminf_{t \downarrow 0} \frac{\|\bar{x} + tz - \bar{w}\| - \|\bar{x} - \bar{w}\|}{t} = \frac{\langle \bar{x} - \bar{w}, z\rangle}{\|\bar{x} - \bar{w}\|},$$

and thus we get the inequality "$\leq$" in (1.47). To justify the opposite inequality in (1.47), fix $z \in \mathbb{R}^n$, take a sequence of $t_k \downarrow 0$ for which the limit in (1.48) is realized as $\varphi(x) = d_\Omega(x)$, and select $w_k \in \Pi_\Omega(\bar{x} + t_k z)$ for large k. Since

$$d_\Omega(\bar{x} + t_k z) = \|\bar{x} + t_k z - w_k\| \leq d_\Omega(\bar{x}) + t_k\|z\| \to d_\Omega(\bar{x}),$$

we may assume that $w_k \to \bar{w}$ as $k \to \infty$ for some $\bar{w} \in \Pi_\Omega(\bar{x})$. By $w_k \in \Omega$ we have $\|\bar{x} - w_k\| \geq \|\bar{x} - \bar{w}\|$ and so

$$\frac{d_\Omega(\bar{x} + t_k z) - d_\Omega(\bar{x})}{t_k} \geq \frac{\|\bar{x} + t_k z - w_k\| - \|\bar{x} - w_k\|}{t_k}.$$

This yields the inequality "$\geq$" in (1.47) by $\langle \psi(\bar{x}), x - \bar{x}\rangle \leq \psi(x) - \psi(\bar{x})$ due the convexity of the norm function $\psi = \|x\|$.

Step 5: *For any $\bar{x} \notin \Omega$ we have both formulas in* (1.44).

It follows from the duality correspondence of Proposition 1.9 between the regular normal cone and the contingent cone and from the above formula (1.48) for $\varphi(x) = d_\Omega(x)$ that

$$\widehat{\partial} d_\Omega(\bar{x}) = \left\{ v \in \mathbb{R}^n \,\middle|\, \langle v, z\rangle \leq \liminf_{t \downarrow 0} \frac{d_\Omega(\bar{x} + tz) - d_\Omega(\bar{x})}{t} \;\text{ for all }\; z \in \mathbb{R}^n \right\}.$$

Combining this with (1.47) tells us that $v \in \widehat{\partial} d_\Omega(\bar{x})$ if and only if

$$\langle v, z\rangle \leq \left\langle \frac{\bar{x} - \bar{w}}{\|\bar{x} - \bar{w}\|}, z \right\rangle \;\text{ for all }\; z \in \mathbb{R}^n,\; \bar{w} \in \Pi_\Omega(\bar{x}),$$

which implies the first formula in (1.44). We can derive the second formula therein by using the first one, representation (1.37), and the definitions. △

Observe that the formulation and proof of Theorem 1.33 are more involved in the out-of-set case in comparison with the in-set one. Let us develop another approach to subdifferentiation of the distance function at out-of-set points $\bar{x} \notin \Omega$ involving the *ρ-enlargement* of Ω *relative to $\bar{x}$* defined by

$$\Omega(\rho) := \left\{ x \in \mathbb{R}^n \,\middle|\, d_\Omega(x) \leq \rho \right\} \;\text{ with }\; \rho := d_\Omega(\bar{x}). \tag{1.49}$$

Note that the ρ-enlargement of Ω is always closed for any $\rho \geq 0$, even when Ω is not. Furthermore, $\Omega(\rho) = \Omega + \rho\mathbb{B}$ if Ω is closed.

First we present a useful result on calculating regular subgradients of d_Ω at $\bar{x} \notin \Omega$ via regular normals to the ρ-enlargement (1.49) at this point.

Lemma 1.34 (Regular Subgradients of the Distance Function via Regular Normals to Enlargements). *For any $\bar{x} \notin \Omega \subset \mathbb{R}^n$, we have*

$$\widehat{\partial} d_\Omega(\bar{x}) = \widehat{N}\big(\bar{x}; \Omega(\rho)\big) \cap \big\{v \in \mathbb{R}^n \big|\ \|v\| = 1\big\}. \tag{1.50}$$

Proof. We start by checking the representation

$$d_{\Omega(\rho)}(x) = d_\Omega(x) - \rho \ \text{ for any } \ x \notin \Omega(\rho) \ \text{ and } \ \rho > 0. \tag{1.51}$$

To proceed, fix $x \notin \Omega(\rho)$ and take any $u \in \Omega(\rho)$ with $d_\Omega(u) \leq \rho$. Then for every $\gamma > 0$, there is $u_\gamma \in \Omega$ satisfying

$$\|u - u_\gamma\| \leq d_\Omega(u) + \gamma \leq \rho + \gamma,$$

which yields in turn the estimates

$$\|u - x\| \geq \|u_\gamma - x\| - \|u_\gamma - u\| \geq d_\Omega(x) - \|u_\gamma - u\| \geq d_\Omega(x) - \rho - \gamma.$$

Since the estimate $\|u - x\| \geq d_\Omega(x) - \rho - \gamma$ holds for all $u \in \Omega(\rho)$ and all $\gamma > 0$, we obtain the inequality

$$d_{\Omega(\rho)}(x) \geq d_\Omega(x) - \rho.$$

To verify the opposite inequality in (1.51), consider the continuous function

$$\varphi(t) := d_\Omega\big(tx + (1-t)u\big), \quad t \in [0, 1],$$

for a fixed point $u \in \Omega$. Since $\varphi(0) = 0$ and $\varphi(1) > \rho$, there is $t_0 \in (0, 1)$ with $\varphi(t_0) = \rho$ by the classical intermediate value theorem. Putting $z := t_0 x + (t - t_0)u$, we have $d_\Omega(z) = \rho$ and $\|x - u\| = \|x - z\| + \|v - u\|$. Hence

$$\|x - u\| \geq \|x - z\| + d_\Omega(z) = \|x - z\| + \rho$$

by $u \in \Omega$ and $z \in \Omega(\rho)$, which ensures the validity of (1.51) .

Using this representation of $d_{\Omega(\rho)}$, we justify now representation (1.50) starting with the inclusion "$\subset$" therein. Pick any $v \in \widehat{\partial} d_\Omega(\bar{x})$ and fix $\gamma > 0$. Then by the construction of regular subgradients, there is $\nu > 0$ for which

$$\langle v, x - \bar{x}\rangle \leq d_\Omega(x) - d_\Omega(\bar{x}) + \gamma\|x - \bar{x}\| \ \text{ whenever } \ x \in \bar{x} + \nu\mathbb{B}.$$

It ensures that $\langle v, x - \bar{x}\rangle \leq \gamma\|x - \bar{x}\|$ for all $x \in (\bar{x} + \nu\mathbb{B}) \cap \Omega(\rho)$ by virtue of $d_\Omega(x) - d_\Omega(\bar{x}) \leq 0$ as $x \in \Omega(\rho)$ with $\rho = d_\Omega(\bar{x})$. This yields $v \in \widehat{N}(\bar{x}; \Omega(\rho))$.

Let us show that $\|v\| = 1$ whenever $v \in \widehat{\partial} d_\Omega(\bar{x})$. Use again the definition of regular subgradients of d_Ω at $\bar{x}$ with γ and ν, put

$$r := \min\Big\{1, \gamma, \frac{\nu}{1 + d_\Omega(\bar{x})}\Big\},$$

and choose $x_r \in \Omega$ so that $\|\bar{x} - x_r\| \le d_\Omega(\bar{x}) + r^2$. For $x := \bar{x} + r(x_r - \bar{x})$, we obviously have the estimates

$$\|x - \bar{x}\| \le r\|\bar{x} - x_r\| \le r d_\Omega(\bar{x}) + r^2 \le r\big(1 + d_\Omega(\bar{x})\big) \le \nu,$$

which lead us to the relationships

$$\begin{aligned}\langle v, x - \bar{x}\rangle &\le \|x - \bar{x}\| - \|\bar{x} - x_r\| + r^2 + \gamma r\|\bar{x} - x_r\|\\ &= -r\|\bar{x} - x_r\| + r^2 + \varepsilon r\|\bar{x} - x_r\|.\end{aligned}$$

Taking into account the above choice of x tells us that

$$\langle v, x_r - \bar{x}\rangle \le -\|\bar{x} - x_r\| + \varepsilon(1 + \|\bar{x} - x_r\|),$$

which readily ensures the estimates

$$\frac{\langle v, \bar{x} - x_r\rangle}{\|\bar{x} - x_r\|} \ge 1 - \gamma\Big(1 + \frac{1}{\|\bar{x} - x_r\|}\Big) \ge 1 - \gamma\Big(1 + \frac{1}{d_\Omega(\bar{x})}\Big),$$

and thus $\|v\| \ge 1$. Since $\|v\| \le 1$ by the Lipschitz continuity of d_Ω with modulus $\ell = 1$, we conclude that $\|v\| = 1$ and get the inclusion "$\subset$" in (1.50).

To justify the opposite inclusion in (1.50), fix $v \in \widehat{N}(\bar{x}; \Omega(\rho))$ with $\|v\| = 1$ and then take arbitrary $\gamma > 0$ and $\eta \in (0, 1)$. By the first relationship in (1.43), we get $v \in \widehat{\partial} d_{\Omega(\rho)}(\bar{x})$, and hence there is $\nu_1 > 0$ such that

$$\langle v, x - \bar{x}\rangle \le d_{\Omega(\rho)}(x) - d_{\Omega(\rho)}(\bar{x}) + \gamma\|x - \bar{x}\| \ \text{ whenever } \ x \in \bar{x} + \nu_1\mathbb{B}.$$

It follows from the representation of $d_{\Omega(\rho)}$ established above that

$$\langle v, x - \bar{x}\rangle \le d_\Omega(x) - d_\Omega(\bar{x}) + \gamma\|x - \bar{x}\| \ \text{ whenever } \ x \in \big(\bar{x} + \nu_1\mathbb{B}\big) \setminus \Omega(\rho).$$

On the other hand, the inclusion $v \in \widehat{N}(\bar{x}; \Omega(\rho))$ implies the existence of $\nu_2 > 0$ ensuring the estimate

$$\langle v, x - \bar{x}\rangle \le (\gamma/2)\|x - \bar{x}\| \ \text{ for all } \ x \in \big(\bar{x} + \nu_2\mathbb{B}\big) \cap \Omega(\rho).$$

Since $\|v\| = 1$, we choose $u \in \mathbb{R}^n$ such that $\|u\| = 1$ and $\langle v, u\rangle \ge 1 - \eta$. Fix $\nu_3 \in (0, \nu_2/2)$ and $x \in (\bar{x} + \nu_3\mathbb{B}) \cap \Omega(\rho)$ and put $\sigma_x := d_\Omega(\bar{x}) - d_\Omega(x) \ge 0$. Then $x + \sigma_x u \in \Omega(\rho) \cap (\bar{x} + \nu\mathbb{B})$ due to

$$d_\Omega(x + \sigma_x u) \le d_\Omega(x) + \sigma_x = d_\Omega(\bar{x}) = \rho \ \text{ and}$$

$$\|x+\sigma_x u-\bar{x}\| \le \|x-\bar{x}\|+\sigma_x \le 2\|x-\bar{x}\| \le 2\nu_3 \le \nu_2,$$

which implies that $\langle v, x+\sigma_x u-\bar{x}\rangle \le \gamma\|x-\bar{x}\|$ and hence

$$\begin{aligned}\langle v, x-\bar{x}\rangle &= \langle v, x+\sigma_x u-\bar{x}\rangle - \langle v, \sigma_x u\rangle \le \gamma\|x-\bar{x}\| - \sigma_x(1-\eta)\\ &\le \gamma\|x-\bar{x}\| + \big(d_\Omega(x)-d_\Omega(\bar{x})\big)(1-\eta).\end{aligned}$$

Since $\eta>0$ was chosen arbitrarily, we have

$$\langle v, x-\bar{x}\rangle \le \gamma\|x-\bar{x}\| + d_\Omega(x)-d_\Omega(\bar{x}) \;\text{ whenever }\; x\in(\bar{x}+\nu_3\mathbb{B})\cap\Omega(\rho),$$

and therefore the latter holds for all $x\in\bar{x}+\nu\mathbb{B}$ with $\nu:=\min\{\nu_1,\nu_3\}$. Thus we get $v\in\widehat{\partial}d_\Omega(\bar{x})$ and complete the proof of the lemma. △

The obtained result (1.50) justifies an exact counterpart of the *first relationship* in (1.43) in the case of regular subgradients of the distance function at out-of set points and regular normals to the enlargement (1.49). A natural question arises about the validity of a corresponding counterpart of the *second relationship* in (1.43) for basic subgradients and normals. The following simple example in $\mathbb{R}^2$ shows that the answer is *negative* for the crucial inclusion

$$\partial d_\Omega(\bar{x}) \subset N\big(\bar{x};\Omega(\rho)\big)\cap\mathbb{B} \;\text{ with }\; \rho=d_\Omega(\bar{x})>0. \tag{1.52}$$

Example 1.35 (Basic Subgradients of the Distance Function Are Not Represented via Basic Normals to Enlargements). Consider the set

$$\Omega := \big\{(x_1,x_2)\in\mathbb{R}^2\big|\; x_1^2+x_2^2\ge 1\big\}$$

with $\bar{x}=(0,0)\notin\Omega$. In this case $d_\Omega(\bar{x})=1$, $\Omega(\rho)=\Omega+\rho\mathbb{B}=\mathbb{R}^2$ with $\rho=1$, and hence $N\big(\bar{x};\Omega(\rho)\big)=\{0\}$. On the other hand, it is easy to see that

$$d_\Omega(x_1,x_2)=1-\sqrt{x_1^2+x_2^2},$$

and so $\partial d_\Omega(\bar{x})=S_{\mathbb{R}^2}$. This shows that inclusion (1.52) fails.

To establish a correct relationship between subgradients of the distance function at out-of-set points and basic normals to the enlargement (1.49), we need to *narrow* the collection of basic subgradients from $\partial d_\Omega(\bar{x})$ at $\bar{x}\notin\Omega$. It is done below by observing that the limiting procedure employed for this purpose employs regular subgradients of the distance function not at all the points $x_k\to\bar{x}$ but only at those where the function values are *to the right* of $d(\bar{x};\Omega)$. In this way we arrive at the following *right-sided* limiting subdifferential of an extended-real-valued function that, along with its modification, is useful for various applications; see more discussions in Section 1.5.

Definition 1.36 (Right-Sided Subdifferential). *Given $\varphi\colon\mathbb{R}^n\to\overline{\mathbb{R}}$ finite at $\bar{x}$, define the* RIGHT-SIDED LIMITING SUBDIFFERENTIAL *of φ at $\bar{x}$ by*

$$\partial_{\geq}\varphi(\bar{x}) := \operatorname*{Lim\,sup}_{x \overset{\varphi+}{\to} \bar{x}} \widehat{\partial}\varphi(x), \tag{1.53}$$

where $x \overset{\varphi+}{\to} \bar{x}$ *means that* $x \to \bar{x}$ *with* $\varphi(x) \to \varphi(\bar{x})$ *and* $\varphi(x) \geq \varphi(\bar{x})$.

It follows directly from the construction in (1.53) that

$$\widehat{\partial}\varphi(\bar{x}) \subset \partial_{\geq}\varphi(\bar{x}) \subset \partial\varphi(\bar{x})$$

while $\partial_{\geq}\varphi(\bar{x})$, in contrast to $\partial\varphi(\bar{x})$, may be empty for simple nonsmooth Lipschitzian functions as in Example 1.35. Observe the following useful properties. The obtained result (1.50) justifies an exact counterpart of the *first relationship* in (1.43) in the case of regular subgradients of the

Proposition 1.37 (Some Properties of the Right-Sided Subdifferential). *Let* $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ *be finite at* $\bar{x}$.

(i) *If* $\bar{x}$ *is a local minimizer of* φ*, then*

$$\partial_{\geq}\varphi(\bar{x}) = \partial\varphi(\bar{x}), \quad \textit{and so} \;\; 0 \in \partial_{\geq}\varphi(\bar{x}).$$

(ii) *We have the stability property with respect to* ε*-enlargements:*

$$\partial_{\geq}\varphi(\bar{x}) = \operatorname*{Lim\,sup}_{\substack{x \overset{\varphi+}{\to} \bar{x} \\ \varepsilon \downarrow 0}} \widehat{\partial}_{\varepsilon}\varphi(x). \tag{1.54}$$

Proof. Property (i) follows from (1.53) and the definition of local minimizers. To verify (ii), we proceed as in the proof of (1.37) in Theorem 1.28. △

Now we are ready to establish relationships between the right-sided subgradients of the distance function and basic normals to enlargements.

Theorem 1.38 (Right-Sided Subgradients of the Distance Function and Basic Normals at Out-of-Set Points). *Given a set* $\emptyset \neq \Omega \subset \mathbb{R}^n$ *and a point* $\bar{x} \notin \Omega$*, denote* $\rho := d_\Omega(\bar{x})$ *and consider the* ρ*-enlargement* $\Omega(\rho)$ *of* Ω *defined in* (1.49)*. Then the following relationships hold:*

$$\partial_{\geq} d_\Omega(\bar{x}) \subset \big[N\big(\bar{x}; \Omega(\rho)\big) \cap \mathbb{B}\big] \setminus \{0\}, \tag{1.55}$$

$$N\big(\bar{x}; \Omega(\rho)\big) = \bigcup_{\lambda \geq 0} \lambda \partial_{\geq} d_\Omega(\bar{x}). \tag{1.56}$$

Proof. To verify (1.55), pick any $v \in \partial_{\geq} d_\Omega(\bar{x})$ and by (1.53) find $x_k \to \bar{x}$ with $d_\Omega(x_k) \geq d_\Omega(\bar{x})$ and $v_k \to v$ satisfying $v_k \in \widehat{\partial} d_\Omega(x_k)$, $k \in \mathbb{N}$. It follows from Lemma 1.34 that $\|v_k\| = 1$ when k sufficiently large, and so $\|v\| = 1$. Denote for convenience $\Omega(\bar{x}) := \Omega(\rho)$ with $\rho = d_\Omega(\bar{x})$ and consider the following two cases, which fully cover the situation: **(a)** There is a subsequence of $\{x_k\}$ such that

$d_\Omega(x_k) = d_\Omega(\bar{x})$ along this subsequence. **(b)** Otherwise. Since $d_\Omega(x_k) > d_\Omega(\bar{x})$ in this case, we have that $x_k \notin \Omega(\bar{x})$ for $k \in \mathbb{N}$ sufficiently large.

In case (a) we get from Lemma 1.34 that $v_k \in \widehat{N}(x_k; \Omega(\bar{x}))$ along this subsequence and then arrive at (1.55) by passing to the limit as $k \to \infty$.

Considering case (b), recall by (1.51) that

$$d_\Omega(x) = d_\Omega(\bar{x}) + d_{\Omega(\bar{x})}(x) \quad \text{whenever} \quad x \notin \Omega(\bar{x}).$$

Therefore for every $k \in \mathbb{N}$, we have the conditions

$$v_k \in \widehat{\partial} d_\Omega(x_k) = \widehat{\partial}\big[d_\Omega(\bar{x}) + d_{\Omega(\bar{x})}\big](x_k) = \widehat{\partial} d_{\Omega(\bar{x})}(x_k)$$

along the sequence under consideration. Denoting $\varepsilon_k := \|x_k - \bar{x}\|$, deduce from the proof of Theorem 1.33(i) the existence of $\{\widetilde{x}_k\} \subset \Omega(\bar{x})$ such that

$$\|\widetilde{x}_k - x_k\| \le d_{\Omega(\bar{x})}(x_k) \le \varepsilon_k \quad \text{and} \quad v_k \in \widehat{N}\big(\widetilde{x}_k; \Omega(\bar{x})\big), \qquad k \in \mathbb{N}.$$

which yields $v \in N(\bar{x}; \Omega(\bar{x}))$ by passing to the limit as $k \to \infty$ and thus completes the verification of inclusion (1.55).

Observe that the inclusion "$\supset$" in (1.56) follows directly from (1.55). To verify the opposite inclusion therein, pick $v \in N\big(\bar{x}; \Omega(\bar{x})\big)$ and suppose that $v \neq 0$; the alternative case is trivial. Then there are some sequences $x_k \to \bar{x}$ with $x_k \in \Omega(\bar{x})$ and $v_k \to v$ such that $v_k \in \widehat{N}(x_k; \Omega(\bar{x}))$ for all $k \in \mathbb{N}$. Since $\|v_k\| > 0$ when k is sufficiently large, we deduce from Lemma 1.34 that

$$v_k \in \|v_k\| \widehat{\partial} d_\Omega(x_k) \quad \text{as} \quad k \to \infty.$$

Note that $d_\Omega(x_k) \le \rho$ by the choice of $x_k \in \Omega(\bar{x})$, while the strict inequality $d_\Omega(x_k) < \rho$ is not possible for large k due to $0 \neq v_k \in \widehat{N}(x_k; \Omega(\bar{x}))$. Selecting now a convergent subsequence of $\|v_k\|$ and using Definition 1.36 of the right-sided subdifferential, we find $\lambda > 0$ such that $v \in \lambda \partial_{\ge} d_\Omega(\bar{x})$, which justifies (1.54) and completes the proof of the theorem. $\triangle$

1.4 Exercises for Chapter 1

Exercise 1.39 (Properties of Generalized Normals).

(i) Show that the normal cone $N(\bar{x}; \Omega)$ in definition (1.4) can change if another norm on $\mathbb{R}^n$ is used instead of the Euclidean one, even for convex sets Ω.

(ii) Show that the collection of regular normal $\widehat{N}(\bar{x}; \Omega)$ defined by (1.5) in arbitrary Banach spaces is invariant with respect to any equivalent norm on this space. Is it true for $\widehat{N}_\varepsilon(\bar{x}; \Omega)$ defined in (1.6) as $\varepsilon > 0$?

(iii) Verify the decreasing property

$$\widehat{N}_\varepsilon(\bar{x}; \Omega_1) \subset \widehat{N}_\varepsilon(\bar{x}; \Omega_2) \quad \text{whenever} \quad \bar{x} \in \Omega_2 \subset \Omega_1 \quad \text{and} \quad \varepsilon \ge 0.$$

Does this hold for the normal cone $N(\bar{x}; \Omega)$ defined by (1.4) or by using (1.7)?

Exercise 1.40 (Sequential vs. Topological Weak* Outer Limits). Let $F\colon X \rightrightarrows X^*$ be a set-valued mapping between a Banach space X and its dual X^*. The *sequential weak* outer limit* of F as $x \to \bar{x}$ is defined by

$$\mathop{\mathrm{Lim\,sup}}_{x\to\bar{x}} F(x) := \Big\{x^* \in X^* \;\Big|\; \exists \text{ sequences } x_k \to \bar{x} \text{ and } x_k^* \xrightarrow{w^*} x^* \text{ with } x_k^* \in F(x_k) \text{ for all } k \in \mathbb{N}\Big\}. \tag{1.57}$$

The *topological weak* outer limit* of F as $x \to \bar{x}$ is defined in scheme (1.57) by replacing the weak* convergence of sequence $x_k^* \to x^*$ by that of nets. Both limits reduce to the Painlevé-Kuratowski outer limit (1.1) if X is finite-dimensional.

(i) Give an example where the topological weak* outer limit of some mapping F at $\bar{x}$ is strictly larger than the sequential weak* outer limit of F at this point.

(ii) Show that the conclusion of (i) holds also in the case where the weak* convergence of nets in the definition of the topological outer limit is replaced by the weak* convergence of *bounded* nets.

Exercise 1.41 (Asplund Spaces). A Banach space X is called *Asplund* (or it has the *Asplund property*) if every convex continuous function $\varphi\colon U \to \mathbb{R}$ on an open convex set $U \subset X$ is Fréchet differentiable on a dense subset of U. Show that

(i) The Asplund property of X is equivalent to the Fréchet differentiability of every convex continuous function $\varphi\colon X \to \mathbb{R}$ at some point of X.

(ii) The space X is Asplund if and only if for every separable subspace $Z \subset X$ its dual subspace $Z^* \subset X^*$ is separable as well.

(iii) If X is Asplund, then unit ball $\mathbb{B}^* \subset X^*$ is weak* sequentially compact.

(iv) The product $X \times Y$ of two Asplund spaces is Asplund.

Hint: Consult the books [255, 522, 638] and the references therein.

Exercise 1.42 (Representation of ε-Normals). Consider the following statement: Given a (locally closed) set $\Omega \subset X$ with $\bar{x} \in \Omega$ and given any numbers $\varepsilon \ge 0$ and $\gamma > 0$, we have the inclusion

$$\widehat{N}_\varepsilon(\bar{x};\Omega) \subset \bigcup \Big\{\widehat{N}(x;\Omega)\Big|\; x \in \Omega \cap (\bar{x}+\gamma\mathbb{B})\Big\} + (\varepsilon+\gamma)\mathbb{B}^*,$$

where the sets of ε-normals in X^* are defined as in (1.6) by using the canonical pairing $\langle x^*, x\rangle$ between X and X^*.

(i) Deduce this statement from the proof of Theorem 1.6 for $X = \mathbb{R}^n$.

(ii) Verify this statement in the case where X is an Asplund space and compare it with the proof of [522, Theorem 2.34] based on the variational result (fuzzy sum rule from the extremal principle) formulated below in Exercise 2.26.

Exercise 1.43 (Basic Normals in Banach and Asplund Spaces). Let $\Omega \subset X$ be a subset of a Banach space with $\bar{x} \in \Omega$. The (basic, limiting) *normal cone* to Ω at $\bar{x}$ is defined by via the sequential weak* outer limit (1.57) by ,

$$N(\bar{x};\Omega) := \mathop{\mathrm{Lim\,sup}}_{\substack{x\xrightarrow{\Omega}\bar{x}\\ \varepsilon\downarrow 0}} \widehat{N}_\varepsilon(x;\Omega) = \Big\{\, x^* \in X^* \Big|\, \exists \text{ seqs. } \varepsilon_k \downarrow 0,\ x_k \xrightarrow{\Omega} \bar{x},\ x_k^* \xrightarrow{w^*} x^* \text{ with } x_k^* \in \widehat{N}_{\varepsilon_k}(x_k;\Omega)\Big\}, \tag{1.58}$$

(i) Show that the normal cone (1.58) can be equivalently represented as

$$N(\bar{x};\Omega) = \mathop{\mathrm{Lim\,sup}}_{x\xrightarrow{\Omega}\bar{x}} \widehat{N}(x;\Omega) = \Big\{\, x^* \in X^* \Big|\, \exists \text{ seqs. } x_k \xrightarrow{\Omega} \bar{x},\ x_k^* \xrightarrow{w^*} x^* \text{ with } x_k^* \in \widehat{N}(x_k;\Omega)\Big\} \tag{1.59}$$

if X is Asplund. *Hint:* Use the results from Exercise 1.42(ii) and Exercise 1.41(iii).

(ii) Give an example showing that set (1.58) may be strictly larger than (1.59) for closed sets in non-Asplund spaces.

(iii) Give examples in both Asplund and non-Asplund settings of Banach spaces showing that replacing the sequential weak* convergence in (1.58) and (1.59) by the weak* convergence of bounded nets results in strictly larger sets.

Exercise 1.44 (Robustness of Generalized Normals in Finite and Infinite Dimensions). Let $\emptyset \neq \Omega \subset X$ be an arbitrary (closed) subset of a Banach space X.

(i) Does the robustness property of Proposition 1.3 hold for the prenormal cone $\widehat{N}(\cdot;\Omega)$ in finite-dimensional spaces?

(ii) Give an example demonstrating that the robustness property fails in $\mathbb{R}^n$ for the convexified normal cone defined in (1.61), which can be represented as

$$\overline{N}(\bar{x};\Omega) := \text{clco}\, N(\bar{x};\Omega), \quad \bar{x} \in \Omega \subset \mathbb{R}^n. \tag{1.60}$$

Hint: Verify first the representation in (1.60) and compare it with Exercise 4.36(iii).

(iii) Show that Proposition 1.3 doesn't generally hold even for cones Ω in Hilbert spaces X; compare it with [522, Example 1.7].

(iv) Give sufficient conditions for robustness of $N(\cdot;\Omega)$ in infinite dimensions. *Hint:* Compare the latter with [522, Theorem 62].

Exercise 1.45 (Normals to Products of Sets in Banach Spaces). Let $\Omega_1 \subset X_1$ and $\Omega_2 \subset X_2$ be nonempty subsets of Banach spaces.

(i) Does a counterpart of Proposition 1.4 hold for regular normals?

(ii) Establish corresponding relationships for ε-normals to products of sets.

(iii) Show that Proposition 1.4 holds for basic normals defined by (1.58).

(iv) Does a counterpart of the product formula from Proposition 1.4 hold for the convexified normal cone (1.60)?

Exercise 1.46 (Convexified Normal Cone to Lipschitzian Manifolds). A set $\Omega \subset \mathbb{R}^q$ is called a *Lipschitzian manifold* of dimension $d \le q$ around $\bar{z} \in \Omega$ if there is $f\colon \mathbb{R}^n \to \mathbb{R}^m$ locally Lipschitzian around $\bar{x}$ such that $\bar{z} = (\bar{x}, f(\bar{x}))$ and the set Ω is locally homeomorphic around $\bar{z}$ to the graph of f. The set Ω is *strictly smooth* at $\bar{z}$ if f can be selected as strictly differentiable (1.19) at $\bar{x}$.

(i) Show that, besides graphs of locally Lipschitzian mappings, Lipschitzian manifolds include graphs *maximal monotone* operators as in (4.27), *subgradient* mappings for *convex* and more generally *prox-regular* functions $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ as in Definition 3.27, etc. *Hint:* Compare with [676, 678].

(ii) Prove that the convexified normal cone (1.60) to a Lipschitzian manifold $\Omega \subset \mathbb{R}^q$ around $\bar{z}$ of dimension d is not a one-sided cone but a *linear subspace* of dimension *greater* than $q - d$, which *equals* to $q - d$ if and only if Ω is strictly smooth at $\bar{z}$. *Hint:* Compare with the proof in [676] while simplifying it by using dual/normal vs. primal/tangent arguments similarly to those in [522, Theorem 3.62].

(iii) Derive a Banach space extension of the "subspace property" result from (i) for the *Clarke normal cone* defined by the dual correspondence

$$\overline{N}(\bar{x};\Omega) := \overline{T}(\bar{x};\Omega)^* = \big\{x^* \in X^* \big|\, \langle x^*, x\rangle \le 0 \ \text{ for all } \ x \in \overline{T}(\bar{x};\Omega)\big\} \tag{1.61}$$

via his (always convex) *regular tangent cone* to Ω at $\bar{x} \in \Omega$ given by

$$\overline{T}(\bar{x};\Omega) := \Big\{w \in X \Big|\, \forall \text{ seqs. } t_k \downarrow 0, x_k \xrightarrow{\Omega} \bar{x}\ \exists\, z_k \xrightarrow{\Omega} \bar{x} \text{ with } \frac{z_k - x_k}{t_k} \to w\Big\}.$$

Hint: Proceed as in the proof of [522, Theorem 3.62].

Exercise 1.47 (Basic Normals in Hilbert Spaces). Let X be a Hilbert space, and let $\Omega \subset X$ with $\bar{x} \in \Omega$. Establish an appropriate counterpart of representation (1.4) of the basic normal cone defined in (1.59), where the "Limsup" in (1.4) and (1.59) are taken with respect to the weak topology of $X^* = X$. *Hint*: Use the projection descriptions from [167, Proposition 1.1.3].

Exercise 1.48 (Normal-Tangent Relationships). Let X be a Banach space, and let $\Omega \subset X$ with $\bar{x} \in \Omega$. Define the *contingent cone* $T(\bar{x};\Omega)$ to Ω at $\bar{x}$ as in (1.11) with the outer limit taken in the norm topology of X. The *weak contingent cone* $T_W(\bar{x};\Omega)$ is the collections of $w \in X$ such that there are sequences $\{x_k\} \subset \Omega$ and $\{\alpha_k\} \subset \mathbb{R}_+$ with $x_k \to \bar{x}$ strongly in X and $\alpha_k(x_k - \bar{x}) \to w$ weakly in X as $k \to \infty$.

(i) Prove the duality relationship

$$\widehat{N}(\bar{x};\Omega) \subset T_W^*(\bar{x};\Omega) := \big\{x^* \in X^* \,|\, \langle x^*, w\rangle \le 0 \ \text{ for all } \ w \in T_W(\bar{x};\Omega)\big\}, \tag{1.62}$$

where the equality holds if X is reflexive. *Hint:* Compare with [522, Theorem 1.10].

(ii) Give an example where $T(\bar{x};\Omega) \ne T_W(\bar{x};\Omega)$ in the case of reflexive spaces, and so the equality in (1.62) fails if $T_W(\bar{x};\Omega)$ is replaced by $T(\bar{x};\Omega)$.

(iii) Do we have the converse duality $\widehat{N}^*(\bar{x};\Omega) = T(\bar{x};\Omega)$ in $\mathbb{R}^n$?

(iv) Obtain relationships between $T(\bar{x};\Omega)$, $T_W(\bar{x};\Omega)$, and $\overline{T}(\bar{x};\Omega)$ in finite and infinite dimensions. *Hint:* See [522, Theorem 1.9] and the references therein.

(v) Show that, along with the duality construction (1.61), the converse duality $\overline{N}^*(\bar{x};\Omega) = \overline{T}(\bar{x};\Omega)$ holds in arbitrary Banach spaces.

Exercise 1.49 (Normals to Contingent Cones). For any $\Omega \subset \mathbb{R}^n$ and $\bar{x} \in \Omega$, we have the following relationships:

(i) $\widehat{N}(\bar{x};\Omega) = \widehat{N}\big(0; T(\bar{x};\Omega)\big)$.

(ii) $N(0; T(\bar{x};\Omega)) \subset N(\bar{x};\Omega)$. *Hint:* Compare with the results and proofs in [678, Proposition 6.27] and [568, Corollary 6.5]

(iii) Give an example showing that the inclusion in (ii) is strict in $\mathbb{R}^2$.

(iii) Do the relationships in (i) and (ii) hold in infinite dimensions?

Exercise 1.50 (Boundary Points and Convex Separation).

(i) Derive the classical convex separation theorem in $\mathbb{R}^n$ from Proposition 1.2.

(ii) Give an example of the failure of Proposition 1.2 in infinite dimensions.

(iii) Derive sufficient conditions for the validity of Proposition 1.2 for closed convex and nonconvex sets in Hilbert spaces.

Exercise 1.51 (Variational Characterization of Regular Normals). Following the proof of Theorem 1.10, clarify that:

(i) Assertion (i) therein holds in any Banach spaces.

(ii) Assertion (ii) therein holds in *Fréchet smooth spaces*, i.e., such Banach (actually Asplund) spaces where there is an equivalent norm (renorming) Fréchet differentiable at every nonzero point. Is the Fréchet smooth property of Banach spaces *necessary* for the validity of the smooth variational description in (ii)?

(iii) It is said that a Banach space X admits an $\mathcal{S}$*-smooth bump function* of a given class $\mathcal{S}$ if there is $b\colon X \to \mathbb{R}$ such that $b(\cdot) \in \mathcal{S}$, $b(x_0) \ne 0$ for some $x_0 \in X$, and $b(x) = 0$ whenever x lies outside a ball in X. Let $\mathcal{S}$ stand either for the class of Fréchet smooth and Lipschitz continuous functions or for the class of $\mathcal{C}^1$-smooth and Lipschitz continuous functions on X. Show that the existence of $\mathcal{S}$-smooth bump functions on X ensures the descriptions of regular normals to any set $\Omega \subset X$ as in assertion (ii) while replacing Fréchet smooth and concave functions therein by $\mathcal{S}$-smooth functions of the aforementioned classes. Is the existence of $\mathcal{S}$-smooth bump functions on X necessary for such descriptions?

Hint: Compare this with [257, Theorems 4.1 and 4.2] and [523, Theorem 1.30].

Exercise 1.52 (Strictly Differentiable Mappings). Let $f: X \to Y$ be a mapping between Banach spaces, and let $\bar{x} \in X$.

(i) Show that the strict differentiability of f at $\bar{x}$ yields the local Lipschitz continuity of f around this point.

(ii) Give an example of a continuous function $f: \mathbb{R} \to \mathbb{R}$, which is Fréchet differentiable at $\bar{x}$ as in (1.12) but not strictly differentiable at this point.

(iii) Give an example of a Lipschitz continuous function $f: \mathbb{R} \to \mathbb{R}$, which is strictly differentiable at $\bar{x}$ but not of class $\mathcal{C}^1$ around this point.

Exercise 1.53 (Adjoints to Surjective Linear Operators). Let $A: X \to Y$ be a linear bounded operator between Banach spaces, and let $A^*: Y^* \to X^*$ be the adjoint operator to A. Assume that A is *surjective* ($AX = Y$), which reduces to the full rank $m \le n$ of A when $X = \mathbb{R}^n$, $Y = \mathbb{R}^m$. Then for any $y^* \in Y^*$, we have

$$\|A^*y^*\| \ge \kappa\|y^*\| \text{ with } \kappa = \inf\Big\{\|A^*y^*\| \;\Big|\; \|y^*\| = 1\Big\} \in (0, \infty).$$

In particular, A^* is *injective*, i.e., $A^*y_1^* \neq A^*y_2^*$ if $y_1^* \neq y_2^*$.

Hint: Use the classical open mapping theorem; cf. [522, Lemma 1.18].

Exercise 1.54 (Normals to Inverse Images of Sets Under Differentiable Mappings). Let $f: X \to Y$ be a mapping between Banach spaces that is strictly differentiable at $\bar{x}$ as in (1.19) and such that the derivative operator $\nabla f(\bar{x}): X \to Y$ is surjective, and let $\Theta \subset Y$ with $\bar{y} := f(\bar{x}) \in \Theta$.

(i) Show that $\widehat{N}(\bar{x}; f^{-1}(\Theta)) = \nabla f(\bar{x})^*\widehat{N}(\bar{y}; \Theta)$. Is the surjectivity of $\nabla f(\bar{x})$ essential here? Is it possible to replace the strict differentiability of f at $\bar{x}$ by its Fréchet differentiability at this point if $\dim Y < \infty$?

(ii) Verify the basic normal formula

$$N\big(\bar{x}; f^{-1}(\Theta)\big) = \nabla f(\bar{x})^* N(\bar{y}; \Theta).$$

Is the strict differentiability of f at $\bar{x}$ essential here in the case of $\dim Y < \infty$?

Hint: Compare it with the proofs of [522, Theorems 1.14 and 1.17] and simplify them in the case of finite-dimensional spaces.

Exercise 1.55 (Normal Regularity of Sets). A subset $\Omega \subset X$ of a Banach space is *normally regular* at $\bar{x} \in \Omega$ if $N(\bar{x}; \Omega) = \widehat{N}(\bar{x}; \Omega)$.

(i) Show that every convex set is normally regular at each of its point.

(ii) Consider that preimage $\Omega := f^{-1}(\Theta)$ of $\Theta \subset Y$ under a mapping $f: X \to Y$ between Banach spaces and assume that f is strictly differentiable at $\bar{x} \in \Omega$ with the surjective derivative $\nabla f(\bar{x})$. Verify that Ω is normally regular at $\bar{x}$ if and only if Θ is normally regular at $\bar{y} := f(\bar{x})$. *Hint:* Use the results of Exercises 1.54 and 1.53.

(iii) Let $\Omega \subset \mathbb{R}^n$ be a Lipschitzian manifold around $\bar{x} \in \Omega$. Show that the set Ω is normally regular at $\bar{x}$ if and only if it is strictly smooth at this point. *Hint:* Employ the results of Exercise 1.46(ii).

Exercise 1.56 (Coderivatives of Mappings Between Banach Spaces). Let $F: X \rightrightarrows Y$ be a set-valued mapping between Banach spaces, and let $(\bar{x}, \bar{y}) \in \operatorname{gph} F$.

(i) The *normal coderivative* $D_N^* F(\bar{x}, \bar{y}): Y^* \rightrightarrows X^*$ of F at $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ is defined by scheme (1.15) by using the normal cone (1.58) to the graph $\Omega = \operatorname{gph} F$ at this point, and thus it admits the weak* sequential limiting representation

$$D_N^* F(\bar{x}, \bar{y})(\bar{y}^*) = \mathop{\mathrm{Lim\,sup}}_{\substack{(x,y)\xrightarrow{\operatorname{gph} F}(\bar{x},\bar{y}) \\ y^*\xrightarrow{w^*}\bar{y}^*,\, \varepsilon\downarrow 0}} \widehat{D}_\varepsilon^* F(x, y)(y^*), \quad \bar{y}^* \in Y^*, \tag{1.63}$$

via the ε-*coderivative* mapping $(x, y, y^*, \varepsilon) \mapsto \widehat{D}^*_\varepsilon F(x, y)(y^*)$ given by

$$\widehat{D}^*_\varepsilon F(x, y)(y^*) := \big\{x^* \in X^* \big| \ (x^*, -y^*) \in \widehat{N}_\varepsilon\big((x, y); \operatorname{gph} F\big)\big\}, \quad y^* \in Y^*. \tag{1.64}$$

Show that ε can be equivalently dropped in (1.63), i.e., $\widehat{D}^*_\varepsilon$ can be replaced by the precoderivative/regular coderivative $\widehat{D}^*$ as in finite dimensions (1.17), provided that the spaces X and Y are Asplund. *Hint:* Use Exercises 1.41(iv) and 1.43(i).

(ii) The *mixed coderivative* $D^*_M F(\bar{x}, \bar{y})\colon Y^* \rightrightarrows X^*$ of F at $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ is

$$D^*_M F(\bar{x}, \bar{y})(\bar{y}^*) = \operatorname*{Lim\,sup}_{\substack{(x,y)\xrightarrow{\operatorname{gph} F}(\bar{x},\bar{y})\\ y^*\xrightarrow{\|\cdot\|}\bar{y}^*,\ \varepsilon\downarrow 0}} \widehat{D}^*_\varepsilon F(x, y)(y^*), \quad \bar{y}^* \in Y^*, \tag{1.65}$$

i.e., it is defined by replacing the weak* convergence $y^* \xrightarrow{w^*} \bar{y}^*$ with the norm convergence $\|y^* - \bar{y}^*\| \to 0$ in Y^*. Show similarly to (i) that ε can be equivalently dropped in (1.65) when both spaces X and Y are Asplund. Furthermore, give an example showing that the sets in (1.65) may be strictly smaller than those in (1.63) for each y^* even for Lipschitz continuous mappings $F = f\colon \mathbb{R} \to Y$ with values in Hilbert spaces. *Hint:* Compare with [522, Example 1.35].

Exercise 1.57 (Coderivatives of Differentiable Mappings). Let $F = f\colon X \to Y$ be a single-valued mapping between Banach spaces, and let $\bar{x} \in X$.

(i) Assume that f is Fréchet differentiable at $\bar{x}$, i.e., (1.19) holds with $z = \bar{x}$. Verify the regular coderivative representation

$$\widehat{D}^* f(\bar{x})(y^*) = \big\{\nabla f(\bar{x})^* y^*\big\} \ \text{ for all } \ y^* \in Y^*.$$

(ii) Assume that f is strictly differentiable at $\bar{x}$ as in (1.19). Show that

$$D^*_N f(\bar{x})(y^*) = D^*_M f(\bar{x})(y^*) = \big\{\nabla f(\bar{x})^* y^*\big\} \ \text{ for all } \ y^* \in Y^*.$$

(iii) Is the strict differentiability assumption essential for the coderivative representations in (ii)? Is it necessary for the validity of these representations?

Hint: To justify (i), proceed as in the proof of Proposition 1.12. The proof of (ii) requires the careful usage of the strict derivative definition; cf. [522, Theorem 1.38].

Exercise 1.58 (Coderivatives of Convex-Graph and Convex-Valued Multifunctions Between Banach Spaces). Let $F\colon X \rightrightarrows Y$ be a set-valued mapping between Banach spaces, and let $(\bar{x}, \bar{y}) \in \operatorname{gph} F$.

(i) Assume that F is convex-graph and check that for all $y^* \in Y^*$, we have

$$\begin{aligned}\widehat{D}^* F(\bar{x}, \bar{y})(y^*) &= D^*_M F(\bar{x}, \bar{y})(y^*) = D^*_N F(\bar{x}, \bar{y})(y^*)\\ &= \Big\{x^* \in X^* \Big| \ \langle x^*, \bar{x}\rangle - \langle y^*, \bar{y}\rangle = \max_{(x,y)\in\operatorname{gph} F} \big[\langle x^*, x\rangle - \langle y^*, y\rangle\big]\Big\}.\end{aligned}$$

(ii) Assume that F is convex-valued around $\bar{x}$ and inner semicontinuous at $\bar{x}$; the latter is defined in (1.20) without any change in Banach spaces. Show that the result of Theorem 1.15 holds for both normal and mixed coderivatives.

Exercise 1.59 (Coderivatives of Indicator Mappings). Given Banach spaces X and Y, consider a nonempty set $\Omega \subset X$ and define the *indicator mapping* $\Delta\colon X \to Y$ of the set Ω relative to the range space Y by

$$\Delta(x; \Omega) := \begin{cases} 0 \in Y & if \quad x \in \Omega,\\ \emptyset & if \quad x \notin \Omega.\end{cases}$$

Check that for any $\bar{x} \in \Omega$ and $y^* \in Y^*$, we have

$$\widehat{D}^*_\varepsilon \Delta(\bar{x}; \Omega)(y^*) = \widehat{N}_\varepsilon(\bar{x}; \Omega), \quad \varepsilon \geq 0;$$

$$D^*_N \Delta(\bar{x}; \Omega)(y^*) = D^*_M \Delta(\bar{x}; \Omega)(y^*) = N(\bar{x}; \Omega).$$

Exercise 1.60 (Graphical Regularity of Mappings). Let $F\colon X \rightrightarrows Y$ be a set-valued mapping between Banach spaces, and let $(\bar{x}, \bar{y}) \in \operatorname{gph} F$.

(i) F is *N-regular* at $(\bar{x}, \bar{y})$ if $D^*_N F(\bar{x}, \bar{y})(y^*) = \widehat{D}^* F(\bar{x}, \bar{y})(y^*)$ for all $y^* \in Y^*$. Indicate classes of mappings that are N-regular and show that this property fails, in particular, for any $f\colon \mathbb{R}^n \to \mathbb{R}^m$, which is locally Lipschitzian around $\bar{x}$ but not strictly differentiable at this point. *Hint:* Use results from previous exercises.

(ii) F is *M-regular* at $(\bar{x}, \bar{y})$ if $D^*_M F(\bar{x}, \bar{y})(y^*) = \widehat{D}^* F(\bar{x}, \bar{y})(y^*)$ for all $y^* \in Y^*$. Construct a mapping that is M-regular but not N-regular at a given point.

(iii) Let $F = f\colon \mathbb{R}^n \to \mathbb{R}^m$ be Lipschitz continuous around $\bar{x}$. Show that f is graphically regular at $\bar{x}$ if and only if it is strictly differentiable at this point. *Hint*: Use the subspace property of the convexified normal cone from Exercise 1.46(ii) and compare with the proof of [522, Theorem 1.46].

(iv) Consider another approach to the result in (iii) and its infinite-dimensional extensions based on the coderivative scalarization as in [522, Subsection 3.2.4].

Exercise 1.61 (Coderivative Chain Rules with Surjective Derivatives of Inner Mappings). Let $g\colon X \to Y$ and $F\colon Y \rightrightarrows Z$ be mappings between Banach spaces, and let $\bar{z} \in (F \circ g)(\bar{x})$. Assume that g is strictly differentiable at $\bar{x}$ with the surjective derivative $\nabla g(\bar{x})$. Then the following hold:

$$\widehat{D}^*(F \circ g)(\bar{x}, \bar{z}) = \nabla g(\bar{x})^* \widehat{D}^* F\big(g(\bar{x}), \bar{z}\big),$$

$$D^*(F \circ g)(\bar{x}, \bar{z}) = \nabla g(\bar{x})^* D^* F\big(g(\bar{x}), \bar{z}\big)$$

for both $D^* = D^*_N, D^*_M$. Moreover, $F \circ g$ is N-regular (resp. M-regular) at $(\bar{x}, \bar{z})$ if and only if F has the corresponding regularity property at $(g(\bar{x}), \bar{z})$. *Hint:* Apply the results from Exercises 1.54 and 1.53; see [522, Theorem 1.66] for more details.

Exercise 1.62 (Slanted Regular Normals to Epigraphs). Let X be Banach, and let $\varphi\colon X \to \overline{\mathbb{R}}$ be l.s.c. around $\bar{x} \in \operatorname{dom} \varphi$. Show that the inclusion $(v, -\lambda) \in \widehat{N}(x, \alpha); \operatorname{epi} \varphi)$ with $\lambda > 0$ implies that $\alpha = \varphi(\bar{x})$. *Hint:* Proceed by the definitions by using arguments similar to those in Step 4 of Theorem 1.23.

Exercise 1.63 (ε-Subgradients of Locally Lipschitzian Functions). Let X be a Banach space, $\varphi\colon X \to \overline{\mathbb{R}}$ be locally Lipschitzian around $\bar{x}$ with modulus $\ell \geq 0$, and $\widehat{\partial}_\varepsilon \varphi(\bar{x})$ be the *ε-subdifferential* of φ at $\bar{x}$ defined as in (1.34) for any $\varepsilon \geq 0$.

(i) Show that there is $\eta > 0$ such that

$$\|x^*\| \leq \ell + \varepsilon \ \text{ for all } \ x^* \in \widehat{\partial}_\varepsilon \varphi(x), \quad x \in \bar{x} + \eta \mathbb{B}.$$

(ii) Show that there is $\eta > 0$ such that

$$\|x^*\| \leq \varepsilon(1 + \ell) \ \text{ for all } \ (x^*, 0) \in \widehat{N}_\varepsilon\big((x, \varphi(x)); \operatorname{epi} \varphi\big), \quad x \in \bar{x} + \eta \mathbb{B},$$

$$\|x^*\| \leq \ell + \varepsilon(1 + \ell) \ \text{ for all } \ (x^*, -1) \in \widehat{N}_\varepsilon\big((x, \varphi(x)); \operatorname{epi} \varphi\big), \quad x \in \bar{x} + \eta \mathbb{B}.$$

Hint: Proceed by the definitions.

Exercise 1.64 (Smooth Variational Descriptions of Regular Subgradients in Infinite Dimensions). Let $\varphi\colon X \to \overline{\mathbb{R}}$ be finite at $\bar{x}$, and let $x^* \in \widehat{\partial} \varphi(\bar{x})$.

(i) Show that the first assertion of Theorem 1.27 holds in arbitrary Banach space X, while the second one requires that X admit a Fréchet smooth renorming. Furthermore, in the latter case, we have the *enhanced minimum condition*

$$\varphi(x) - \psi(x) - \|x - \bar{x}\|^2 \geq \varphi(\bar{x}) - \psi(\bar{x}) \text{ for all } x \in X. \tag{1.66}$$

(ii) Derive appropriate analogs of (1.66) in Banach spaces admitting $\mathcal{S}$-smooth bump functions of the classes listed in Exercise 1.51(iii).

Hint: Proceed similarly to the proof of Theorem 1.27 with taking into account the results of Exercise 1.51 and compare this with [522, Theorem 1.88].

Exercise 1.65 (Basic Subdifferential in Infinite Dimensions). Let X be Banach. Define the *basic subdifferential* of $\varphi\colon X \to \overline{\mathbb{R}}$ at $\bar{x} \in \operatorname{dom}\varphi$ geometrically

$$\partial\varphi(\bar{x}) := \big\{x^* \in X^* \big|\ (x^*, -1) \in N\big((\bar{x}, \varphi(\bar{x})); \operatorname{epi}\varphi\big)\big\} \tag{1.67}$$

via the basic/limiting normal cone (1.58) in Banach spaces.

(i) Show that $\partial\varphi(\bar{x})$ from (1.67) admits the following analytic representation

$$\partial\varphi(\bar{x}) := \operatorname*{Lim\,sup}_{\substack{x \xrightarrow{\varphi} \bar{x} \\ \varepsilon \downarrow 0}} \widehat{\partial}_\varepsilon \varphi(x) \tag{1.68}$$

via the sequential weak* outer limit (1.57) of ε-subgradients at points nearby. *Hint:* Deduce it from definition (1.58) and Theorem 1.26, which holds in an arbitrary Banach space without any change in the proof.

(ii) Let X be Asplund. Show that $\partial\varphi(\bar{x})$ admits the equivalent representation

$$\partial\varphi(\bar{x}) := \operatorname*{Lim\,sup}_{x \xrightarrow{\varphi} \bar{x}} \widehat{\partial}\varphi(x). \tag{1.69}$$

Hint: Employ the result from Exercise 1.43(i).

Exercise 1.66 (Subgradients of the Norm and Negative Norm Functions).

(i) Consider the norm function $\varphi(x) := \|x\|$ defined on an arbitrary Banach space X. Based on the definitions, show that

$$\widehat{\partial}\varphi(\bar{x}) = \partial\varphi(\bar{x}) = \begin{cases} \mathbb{B}^* & \text{if } \bar{x} = 0, \\ \big\{x^* \in X^* \big|\ \|x^*\| = 1,\ \langle x^*, \bar{x}\rangle = \|\bar{x}\|\big\} & \text{if } \bar{x} \neq 0. \end{cases}$$

(ii) Based on the above definitions, calculate $\widehat{\partial}\varphi(\bar{x})$ and $\partial\varphi(\bar{x})$ for $\varphi(x) := -\|x\|$ at $\bar{x} = 0$ and $\bar{x} \neq 0$ in **(a)** finite-dimensional Euclidean and non-Euclidean spaces, **(b)** Asplund spaces, and (c) Banach while not Asplund spaces.

Exercise 1.67 (Subgradients of Strictly Differentiable Functions). Let X be Banach, and let $\varphi\colon X \to \overline{\mathbb{R}}$ be strictly differentiable at $\bar{x}$.

(i) Show that $\widehat{\partial}\varphi(\bar{x}) = \partial\varphi(\bar{x}) = \big\{\nabla\varphi(\bar{x})\big\}$.

(ii) Is the strict differentiability of φ at $\bar{x}$ necessary for the validity of the second equality in (i) when $\varphi\colon \mathbb{R}^n \to \mathbb{R}$ is Fréchet differentiable at $\bar{x}$?

Exercise 1.68 (Singular Subdifferential in Infinite Dimensions). Let X be Banach. Define the *singular subdifferential* of $\varphi\colon X \to \overline{\mathbb{R}}$ at $\bar{x} \in \operatorname{dom}\varphi$ by

$$\partial^\infty\varphi(\bar{x}) := \operatorname*{Lim\,sup}_{\substack{x \xrightarrow{\varphi} \bar{x} \\ \lambda,\ \varepsilon \downarrow 0}} \lambda\widehat{\partial}_\varepsilon \varphi(x) \tag{1.70}$$

via the sequential weak* outer limit (1.57) in Banach spaces.

(i) Assume that X is Asplund and show that in this case

$$\partial^\infty\varphi(\bar{x}) := \mathop{\mathrm{Lim\,sup}}_{\substack{x\xrightarrow{\varphi}\bar{x}\\ \lambda\downarrow 0}} \lambda\widehat{\partial}\varphi(x), \tag{1.71}$$

i.e., $\varepsilon > 0$ can be equivalently dismissed in (1.70).

(ii) Verify that in Asplund spaces we have the geometric representation

$$\partial^\infty\varphi(\bar{x}) := \big\{x^* \in X^* \big|\ (x^*, 0) \in N\big((\bar{x}, \varphi(\bar{x})); \mathrm{epi}\,\varphi\big)\big\}. \tag{1.72}$$

Hint: Compare this with [522, Theorem 2.28] and simplify the proof in Hilbert spaces following the approach developed in [470].

(iii) Does representation (1.72) hold in general Banach spaces with the normal cone defined in (1.58) and the singular subdifferential defined in (1.70)?

Exercise 1.69 (Basic and Singular Subgradients of Lipschitzian Functions in Banach Spaces). Let X be a Banach space, and let $\varphi\colon X \to \overline{\mathbb{R}}$ be a locally Lipschitzian function around $\bar{x}$ with modulus $\ell \geq 0$.

(i) Prove the subgradient estimate (1.27). *Hint:* Use (1.68) and Exercise 1.63(i).

(ii) Prove that $\partial^\infty\varphi(\bar{x}) = \{0\}$. *Hint:* Use (1.70) and Exercise 1.63(i).

(iii) Give an example showing that the condition $\partial^\infty\varphi(\bar{x}) = \{0\}$ doesn't imply the local Lipschitz continuity of φ in infinite dimensions.

Exercise 1.70 (Scalarization of the Regular and Mixed Coderivatives in Banach Spaces). Let $f\colon X \to Y$ be a mapping between Banach spaces, which is assumed to be locally Lipschitzian around $\bar{x}$.

(i) Show that $\widehat{D}^* f(\bar{x})(y^*) = \widehat{\partial}\langle y^*, f\rangle(\bar{x})$ for all $y^* \in Y^*$.

(ii) Show that $D^*_M f(\bar{x})(\bar{x})(y^*) = \partial\langle y^*, f\rangle(\bar{x})$ for all $y^* \in Y^*$. *Hint:* Proceed as in the proof of Theorem 1.32 with using the ε-enlargements in (1.65) and (1.68) as well the norm convergence on Y^* in the construction of $D^*_M f(\bar{x}, \bar{y})(y^*)$.

(iii) Give an example showing that the scalarization formula in (i) is violated for the normal coderivative of Lipschitzian mappings with values in Hilbert spaces.

(iv) Does an analog of the scalarization formula hold for the coderivative generated by the convexified normal cone to graphs of locally Lipschitzian mappings between finite-dimensional spaces?

Exercise 1.71 (Scalarization of the Normal Coderivative for Strictly Lipschitzian Mappings). Let X, Y be Banach, and let $f\colon X \to Y$ be locally Lipschitzian around $\bar{x}$. It is w^*-*strictly Lipschitzian* at $\bar{x}$ if there is a neighborhood V of $0 \in X$ such that for any $u \in X$ and any sequences $x_k \to \bar{x}$, $t_k \downarrow 0$, and $y_k^* \xrightarrow{w^*} 0$, we have $\langle y_k^*, y_k\rangle \to 0$ as $k \to \infty$ with $y_k := t_k^{-1}[f(x_k + t_k u) - f(x_k)]$.

(i) Show that any mapping f strictly differentiable at $\bar{x}$ is w^*-strictly Lipschitzian at this point and find other conditions ensuring the validity of the w^*-strict Lipschitzian property of f at $\bar{x}$.

(ii) Show that the w^*-strict Lipschitzian property of f at $\bar{x}$ implies that for any sequences $\varepsilon_k \downarrow 0$, $x_k \to \bar{x}$, and $(y_k^*, x_k^*) \in \mathrm{gph}\,\widehat{D}^*_{\varepsilon_k} f(x_k)$, we have the implication

$$y_k^* \xrightarrow{w^*} 0 \Longrightarrow x^* \xrightarrow{w^*} 0 \ \text{ as } \ k \to \infty.$$

(iii) Assuming that X is Asplund and that f is w^*-strictly Lipschitzian at $\bar{x}$, justify the scalarization formula for the normal coderivative:

$$D^*_N f(\bar{x})(y^*) = \partial\langle y^*, f\rangle(\bar{x}) \ \text{ for all } \ y^* \in Y^*.$$

Hint: Use (ii) and compare it with the proof of [522, Theorem 3.28].

Exercise 1.72 (Subgradients of Compositions with Surjective Derivatives of Inner Mappings). Consider the composition $\varphi \circ g$ of a mapping $g\colon X \to Y$ between Banach spaces and a function $\varphi\colon Y \to \overline{\mathbb{R}}$. Assume that g is strictly differentiable at $\bar{x}$ with the surjective derivative $\nabla g(\bar{x})$ and that φ is finite at $\bar{y} := g(\bar{x})$. Verify the following subdifferential chain rules: $\widehat{\partial}(\varphi \circ g)(\bar{x}) = \nabla g(\bar{x})^* \widehat{\partial}\varphi(\bar{y})$,

$$\partial(\varphi \circ g)(\bar{x}) = \nabla g(\bar{x})^* \partial\varphi(\bar{y}), \quad \text{and } \partial^\infty(\varphi \circ g)(\bar{x}) = \nabla g(\bar{x})^* \partial^\infty\varphi(\bar{y}).$$

Hint: Deduce these equalities from the coderivative calculus in Exercise 1.61 by considering there the epigraphical multifunction $F = E_\varphi$ defined in (1.29).

Exercise 1.73 (Proximal Subgradients and Their Limits in Hilbert Spaces). Let $\varphi\colon X \to \overline{\mathbb{R}}$, where X is a Hilbert space. The *proximal subdifferential* of φ at $\bar{x} \in \operatorname{dom}\varphi$ is defined as the collection of proximal subgradients

$$\partial_P\varphi(\bar{x}) := \Big\{x^* \in X^* \Big|\ \liminf_{x \to \bar{x}} \frac{\varphi(x) - \varphi(\bar{x}) - \langle x^*, x - \bar{x}\rangle}{\|x - \bar{x}\|^2} > -\infty\Big\}.$$

(i) Show that $\partial_P\varphi(\bar{x}) \subset \widehat{\partial}\varphi(\bar{x})$ and that the proximal subgradient set $\partial_P\varphi(\bar{x})$ may not be closed in $\mathbb{R}^n$ in contrast to $\widehat{\partial}\varphi(\bar{x})$.

(ii) Give an example showing that the set $\partial_P\varphi(\bar{x})$ may be empty even for smooth functions on finite-dimensional spaces.

(iii) Show that for any $x^* \in \widehat{\partial}\varphi(\bar{x})$, there are sequences $x_k \overset{\varphi}{\to} \bar{x}$ and $x_k^* \in \partial_P\varphi(x_k)$ such that $\|x_k^* - x^*\| \to 0$ as $k \to \infty$. *Hint:* Compare this with the proof in [472, Theorem 5.5] and simplify it in the case of $X = \mathbb{R}^n$.

(iv) Based on (iii) and (1.69), derive the limiting subdifferential representation

$$\partial\varphi(\bar{x}) = \operatorname*{Lim\,sup}_{x \overset{\varphi}{\to} \bar{x}} \partial_P\varphi(x).$$

Exercise 1.74 (Subdifferential Regularity of Functions). Let X be a Banach space. A function $\varphi\colon X \to \overline{\mathbb{R}}$ is *subdifferentially* or *epigraphically regular* at $\bar{x} \in \operatorname{dom}\varphi$ if its epigraph is normally regular at $(\bar{x}, \varphi(\bar{x}))$.

(i) Show that the function φ is subdifferentially regular at $\bar{x}$ if and only if

$$\partial\varphi(\bar{x}) = \widehat{\partial}\varphi(\bar{x}) \ \text{ and } \ \partial^\infty\varphi(\bar{x}) = \Big\{v \in X^* \Big|\ (v, 0) \in \widehat{N}\big((\bar{x}, \varphi(\bar{x})); \operatorname{epi}\varphi\big)\Big\}, \tag{1.73}$$

where the first equality in (1.73) is known as the *lower regularity* of φ at $\bar{x}$.

(ii) Show that for locally Lipschitzian functions φ on arbitrary Banach spaces, the subdifferential regularity and lower regularity of φ at $\bar{x}$ are equivalent, while it is not the case in general even for $X = \mathbb{R}$.

(iii) It follows from Theorem 1.33 that the distance function d_Ω for $\Omega \subset \mathbb{R}^n$ is lower regular at $\bar{x} \in \Omega$ if and only if the set Ω is normally regular at this point, while d_Ω is lower regular at $\bar{x} \notin \Omega$ if and only if the Euclidean projector $\Pi(\bar{x}; \Omega)$ is a singleton. Do these facts hold in infinite dimensions?

Exercise 1.75 (Upper and Symmetric Subdifferentials). Given a function $\varphi\colon X \to [-\infty, \infty)$ finite at $\bar{x}$ on a Banach space X, define the *upper subdifferential* and *upper singular subdifferential* of φ at $\bar{x}$ by, respectively,

$$\partial^+\varphi(\bar{x}) := -\partial(-\varphi)(\bar{x}), \quad \partial^{\infty,+}\varphi(\bar{x}) := -\partial^\infty(-\varphi)(\bar{x}). \tag{1.74}$$

The *symmetric subdifferential* and the *symmetric singular subdifferential* of φ at $\bar{x}$ are defined by, respectively,

$$\partial^0\varphi(\bar{x}) := \partial\varphi(\bar{x}) \cup \partial^+\varphi(\bar{x}), \quad \partial^{\infty,0}\varphi(\bar{x}) := \partial^\infty\varphi(\bar{x}) \cup \partial^{\infty,+}\varphi(\bar{x}). \tag{1.75}$$

(i) Check the *plus-minus symmetry* properties of the constructions in (1.75):

$$\partial^0(-\varphi)(\bar{x}) = -\partial^0\varphi(\bar{x}), \quad \partial^{\infty,0}(-\varphi)(\bar{x}) = -\partial^{\infty,0}\varphi(\bar{x})$$

(ii) Let φ be locally Lipschitzian around $\bar{x}$ with modulus $\ell \geq 0$. Check that

$$\partial^{\infty,0}\varphi(\bar{x}) = \{0\} \text{ and } \|x^*\| \leq \ell \text{ for all } x^* \in \partial^0\varphi(\bar{x}).$$

Exercise 1.76 (Upper Regular Subgradients and Subdifferential Characterization of Fréchet Differentiability). Define the collection of *upper regular subgradients* of $\varphi\colon X \to [-\infty, \infty)$ finite at $\bar{x}$ by $\widehat{\partial}^+\varphi(\bar{x}) := -\widehat{\partial}(-\varphi)(\bar{x})$, i.e.,

$$\widehat{\partial}^+\varphi(\bar{x}) = \left\{x^* \in X^* \,\middle|\, \limsup_{x \to \bar{x}} \frac{\varphi(x) - \varphi(\bar{x}) - \langle x^*, x - \bar{x}\rangle}{\|x - \bar{x}\|} \leq 0\right\}. \tag{1.76}$$

(i) Give examples showing that the sets $\widehat{\partial}\varphi(\bar{x})$ and $\widehat{\partial}^+\varphi(\bar{x})$ may be empty simultaneously for a continuous function $\varphi\colon \mathbb{R} \to \mathbb{R}$ and that $\widehat{\partial}\varphi(\bar{x})$ may be a singleton when φ is not Fréchet differentiable at $\bar{x}$.

(ii) Show that φ is Fréchet differentiable at $\bar{x}$ if and only if we have simultaneously $\widehat{\partial}\varphi(\bar{x}) \neq \emptyset$ and $\widehat{\partial}^+\varphi(\bar{x}) \neq \emptyset$ in which case $\widehat{\partial}^+\varphi(\bar{x}) = \widehat{\partial}\varphi(\bar{x}) = \{\nabla\varphi(\bar{x})\}$.

Exercise 1.77 (Epigraphical Regularity and Symmetric Subgradients for Convex Functions). Let $\varphi\colon X \to \overline{\mathbb{R}}$ be convex on a Banach space X. Show that φ is epigraphically regular at every $\bar{x} \in \operatorname{dom}\varphi$, and we have

$$\partial^0\varphi(\bar{x}) = \partial\varphi(\bar{x}) = \left\{x^* \in X^* \,\middle|\, \langle x^*, x - \bar{x}\rangle \leq \varphi(x) - \varphi(\bar{x}) \text{ for all } x \in X\right\}.$$

Hint: All the claimed properties but the representation for $\partial^0\varphi(\bar{x})$ in Banach spaces are verified similarly to the proof of Proposition 1.25. To justify the latter representation, it remains to show that $\partial^+\varphi(\bar{x}) \subset \partial\varphi(\bar{x})$ for convex functions. The latter can be proved by applying (1.68) to $-\partial(-\varphi)(\bar{x})$ and observing that the condition $-\widehat{\partial}_\varepsilon(-\varphi)(x) \neq \emptyset$ for some x and $\varepsilon > 0$ ensures that φ is bounded from above around x and thus $\widehat{\partial}\varphi(x) = \partial\varphi(x) \neq \emptyset$ due the convexity of φ. Then apply Exercise 1.76(ii) and compare with [522, Theorem 1.93] for more details.

Exercise 1.78 (Characterizations of Two-Sided Regularity for Continuous Functions). A function $\varphi\colon X \to [-\infty, \infty)$ finite at $\bar{x}$ is *upper regular* at this point if $\partial^+\varphi(\bar{x}) = \widehat{\partial}^+\varphi(\bar{x})$, i.e., the function $-\varphi$ is lower regular at $\bar{x}$.

(i) Show that the graphically regular of φ at $\bar{x}$ (in both sense of Exercise 1.60 for $Y = \mathbb{R}$) implies that φ is simultaneously lower and upper regular at this point. The converse holds if φ is locally Lipschitzian around $\bar{x}$. *Hint:* Use the corresponding Banach space extension of (1.30) Theorem 1.23 and the result of Exercise 1.75(ii).

(ii) Check that the strict differentiability of φ at $\bar{x}$ ensures both lower and upper regularity of φ at this point. The converse holds if φ is locally Lipschitzian around $\bar{x}$ and $\dim X < \infty$. *Hint:* To verify the converse statement, use Exercise 1.60(iii).

Exercise 1.79 (Generalized Directional Derivative and Generalized Gradient). Let X be a Banach space.

(i) Assume that $\varphi\colon X \to \overline{\mathbb{R}}$ is locally *Lipschitzian* around $\bar{x}$. The (Clarke) *generalized directional derivative* of φ at $\bar{x}$ in the direction $w \in X$ is

$$\varphi^\circ(\bar{x}; w) := \limsup_{\substack{x \to \bar{x} \\ t \downarrow 0}} \frac{\varphi(x + tw) - \varphi(x)}{t} \tag{1.77}$$

and the corresponding *generalized gradient* of φ at $\bar{x}$ is

$$\overline{\partial}\varphi(\bar{x}) = \left\{x^* \in X^* \,\middle|\, \langle x^*, w\rangle \leq \varphi^\circ(\bar{x}; w) \text{ for all } w \in X\right\}. \tag{1.78}$$

Show that the function $w \mapsto \varphi^\circ(\bar{x}; w)$ is convex and satisfies the condition $\varphi^\circ(\bar{x}, -w) = -\varphi^\circ(\bar{x}; w)$, which implies the plus-minus symmetry

$$\overline{\partial}(-\varphi)(\bar{x}) = -\overline{\partial}\varphi(\bar{x}). \tag{1.79}$$

(ii) Verify that the convexified normal cone (1.61) admits the representation

$$\overline{N}(\bar{x}; \Omega) = \text{cl}^*\Big\{\bigcup_{\lambda \geq 0} \lambda \overline{\partial} d_\Omega(\bar{x})\Big\} \tag{1.80}$$

via the (topological) weak* closure of the cone spanned on the generalized gradient (1.78) of the Lipschitzian distance function, which induces the corresponding subdifferential of a general (l.s.c.) function $\varphi\colon X \to \overline{\mathbb{R}}$ at $\bar{x} \in \text{dom}\,\varphi$ by

$$\overline{\partial}\varphi(\bar{x}) := \big\{x^* \in X^* \big|\ (x^*, -1) \in \overline{N}\big((\bar{x}, \varphi(\bar{x})); \text{epi}\,\varphi\big)\big\}. \tag{1.81}$$

(iii) Show that for any $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ finite at $\bar{x}$, we have the representation

$$\overline{\partial}\varphi(\bar{x}) = \text{clco}\big[\partial\varphi(\bar{x}) + \partial^\infty\varphi(\bar{x})\big], \tag{1.82}$$

which leads us in the case of locally Lipschitzian functions to the following ones:

$$\overline{\partial}\varphi(\bar{x}) = \text{co}\,\partial\varphi(\bar{x}) = \text{co}\,\partial^+\varphi(\bar{x}) = \text{co}\,\partial^0\varphi(\bar{x}). \tag{1.83}$$

Hint: For (i) and (ii), consult [165]. To verify (iii), deduce (1.82) from (1.60) and then derive all the conditions in (1.83) from (1.82) by using Theorem 1.22 and the symmetry relationship (1.79) for locally Lipschitzian functions.

Exercise 1.80 (Generalized Jacobian of Lipschitzian Mappings and Subgradients of Scalarizations). Let $f\colon \mathbb{R}^n \to \mathbb{R}^m$ be locally Lipschitzian around $\bar{x}$. By the classical Rademacher theorem (see, e.g., [678, Theorem 9.60]), f is differentiable almost everywhere around $\bar{x}$. The (Clarke) *generalized Jacobian* $\overline{\partial} f(\bar{x})$ of f at $\bar{x}$ is a nonempty compact subset of $\mathbb{R}^{n\times m}$ defined as the convex hull of the set

$$\big\{\lim \nabla f(x_k) \big|\ x_k \to \bar{x},\ k \to \infty,\ f \text{ is differentiable at } x_k\big\}$$

via the limit of the Jacobian matrix $\nabla f(x_k)$ for f at x_k.

(i) Show that for $m = 1$, the generalized Jacobian of f at $\bar{x}$ reduces to the generalized gradient of f at this point. *Hint:* Proceed by the definitions with the usage of the classical Fubini theorem; compare it with [165, Theorem 2.5.1].

(ii) Show that for any $m \in \mathbb{N}$, we have the following relationships:

$$D^* f(\bar{x})(v) = \partial\langle v, f\rangle(\bar{x}) = \text{co}\big\{A^* v \big|\ A \in \overline{\partial} f(\bar{x})\big\} \text{ whenever } v \in \mathbb{R}^m.$$

Hint: Use (i), (1.83), and the coderivative scalarization from Theorem 1.32.

(iii) Establish appropriate infinite-dimensional versions of the relationships in (ii) for locally Lipschitzian mappings defined on Asplund spaces. *Hint:* Use the scalarization results from Exercises 1.70 and 1.71 together with Preiss' extension [647] of the Rademacher theorem to such mappings.

Exercise 1.81 (More Subgradient Calculations).

(i) Consider all the functions $\varphi\colon \mathbb{R} \to \mathbb{R}$ from Example 1.21(i–iv) and calculate for them the subgradient sets $\widehat{\partial}^+\varphi(0)$, $\partial^+\varphi(0)$, $\partial^0\varphi(0)$, $\overline{\partial}\varphi(0)$, $\partial^{\infty,+}\varphi(0)$, and $\partial^{\infty,0}\varphi(0)$. Draw the corresponding figures.

(ii) Consider the two Lipschitz functions $\varphi\colon \mathbb{R}^2 \to \mathbb{R}$ from Example 1.31, calculate for them the subgradient sets $\widehat{\partial}^+\varphi(0,0)$, $\partial^+\varphi(0,0)$, $\partial^0\varphi(0,0)$, and $\overline{\partial}\varphi(0,0)$, and then draw the illustrating figures.

(iii) Define the functions $\varphi\colon \mathbb{R}^2 \to \mathbb{R}$ by

$$\varphi(x_1, x_2) := |x_1|^\alpha - |x_2|, \quad \varphi(x_1, x_2) := |x_1| - |x_2|^\beta, \quad \varphi(x_1, x_2) := |x_1|^\alpha - |x_2|^\beta$$

for any $\alpha, \beta \in (0, 1)$. Calculate for these functions the sets $\widehat{\partial}\varphi(0,0)$, $\widehat{\partial}^+\varphi(0,0)$, $\partial\varphi(0,0)$, $\partial^+\varphi(0,0)$ $\partial^0\varphi(0,0)$, and $\overline{\partial}\varphi(0,0)$ as well as their singular counterparts $\partial^\infty\varphi(0,0)$, $\partial^{\infty,+}\varphi(0,0)$, $\partial^{\infty,0}\varphi(0,0)$ with the corresponding geometric illustrations.

Exercise 1.82 (Duality for Regular Subgradients and Contingent Derivatives in Finite and Infinite Dimensions).

(i) Given $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ and $\bar{x} \in \operatorname{dom}\varphi$, show that

$$\widehat{\partial}\varphi(\bar{x}) = \big\{ v \in \mathbb{R}^n \big|\ \langle v, w\rangle \le d\varphi(\bar{x}; w) \ \text{ for all } \ w \in \mathbb{R}^n \big\},$$

where $d\varphi(\bar{x}; w)$ is the contingent derivative from (1.41) and (1.42).

(ii) Does this representation hold in infinite dimensions?

Exercise 1.83 (Relationships Between Directional Derivatives). Let X be a Banach space, and let $\varphi\colon X \to \overline{\mathbb{R}}$ be finite $\bar{x}$.

(i) Assuming that the classical *directional derivative* of the function φ at the point $\bar{x}$ in the direction $w \in X$ given by

$$\varphi'(\bar{x}; w) := \lim_{t \downarrow 0} \frac{\varphi(\bar{x} + tw) - \varphi(\bar{x})}{t} \tag{1.84}$$

exists whenever $w \in X$, show that $d\varphi(\bar{x}; w) \le \varphi'(\bar{x}; w)$ for the contingent derivative (1.42), where the inequality may be strict for continuous functions on $\mathbb{R}$.

(ii) Assuming that φ is locally Lipschitzian around $\bar{x}$, show the relationship $d\varphi(\bar{x}; w) \le \varphi^\circ(\bar{x}; w)$ between the generalized directional derivatives (1.42) and (1.77) for all $w \in X$, where the inequality may be strict when $X = \mathbb{R}$.

(iii) Assuming that φ is locally Lipschitzian around $\bar{x}$ and that $\varphi'(\bar{x}; w)$ exists for all $w \in X$, show that the inequality $\varphi'(\bar{x}; w) \le \varphi^\circ(\bar{x}; w)$ may be strict even for $X = \mathbb{R}$. The case of equality therein is known as the *tangential, directional*, or *Clarke regularity* of φ at $\bar{x}$. Show that it always holds for convex function and that we have $d\varphi(\bar{x}) = \varphi^\circ(\bar{x})$ under this regularity.

Hint: See [124, 125] for detailed comparisons between the aforementioned and other regularity notions in variational analysis.

Exercise 1.84 (Calculus of Right-Sided Subgradients). Clarify which calculus properties are available for the right-sided subdifferential (1.53).

Exercise 1.85 (Subdifferentiation of the Distance Function in Infinite Dimensions). Let $\emptyset \neq \Omega \subset X$, where X is Banach.

(i) Derive counterparts of Theorem 1.33(i) and Lemma 1.34 for ε-normals and ε-subgradients at in-set and out-of-set points.

(ii) Prove the corresponding extensions of Theorem 1.33(i) and Theorem 1.38 in arbitrary Banach spaces X.

Hint: Use Ekeland's variational principle (see Chapter 2) and compare with the proofs in [522, Theorems 1.97, 1.99, 1.101].

Exercise 1.86 (Subgradients of the Distance Function via Projection Points).

(i) Show that in any infinite-dimensional Hilbert space X, there is a closed set Ω such that the formula for $\partial d_\Omega(\bar{x})$ in Theorem 1.33(ii) is violated. *Hint:* Construct Ω as an orthonormal basis of X and take $\bar{x} = 0 \notin \Omega$.

(ii) Let $\Omega \subset X$ be an nonempty subset of a Banach space X, and let $\bar{x} \notin \Omega$. We say that the best approximation problem is *well posed* for Ω at $\bar{x}$ if either **(a)** for every sequence of $x_k \to \bar{x}$ with $\widehat{\partial}_{\varepsilon_k} d_\Omega(x_k) \neq \emptyset$ as $\varepsilon_k \downarrow 0$ there is a sequence of $w_k \in \Pi(x_k; \Omega)$ that contains a convergent subsequence or **(b)** every sequence of $x_k \xrightarrow{\Omega} \bar{x}$ such that $\|x_k - \bar{x}\| \to d_\Omega(\bar{x})$ as $k \to \infty$ contains a convergent subsequence.

Show that the defined well-posedness property for Ω at $\bar{x}$ ensures the validity of

$$\partial d_\Omega(\bar{x}) \subset \bigcup_{w \in \Pi(\bar{x};\Omega)} \big[N(w;\Omega) \cap \mathbb{B}^*\big] \tag{1.85}$$

(iii) Let X be a reflexive Banach space with an equivalent *Kadec norm*, i.e., such that the strong and weak convergences agree on the boundary of the unit sphere. Verify that the best approximation problem is well posed on Ω and hence (1.85) holds if either Ω is weakly closed or Ω is closed and $\widehat{\partial} d_\Omega(\bar{x}) \neq \emptyset$.

Hint: Compare with the approach and results in [522, Theorem 1.105 and Corollary 1.106] for the proofs of the corresponding assertions in (ii) and (iii).

Exercise 1.87 (Fermat-Torricelli-Steiner Problems). Given an arbitrary number of closed subsets $\Omega_i \subset \mathbb{R}^n$ as $i = 1, \ldots, s$, consider the *generalized Fermat-Torricelli-Steiner problem* [536] defined by

$$\text{minimize } \sum_{i=1}^{s} d_{\Omega_i}(x) \text{ over all } x \in \mathbb{R}^n. \tag{1.86}$$

The classical Fermat-Torricelli problem corresponds to (1.86) with three singletons Ω_i in $\mathbb{R}^2$, while the Steiner problem deals with any finite number of points on the plane: see Section 1.5 for more discussions and references. Using Proposition 1.30(i) and Theorem 1.33(ii) as well as the classical subdifferential sum rule of convex analysis, find exact the solutions to problem (1.86) in the following two cases:

(i) The sets Ω_i, $i = 1, \ldots, s$, are disjoint interval $[a_i, b_i]$ on the real line with $a_1 \le b_1 < a_2 \le b_2 < \ldots < a_s \le b_s$.

(ii) The sets Ω_i are three pairwise disjoint balls on the plane.

Hint: See [537, Chapter 4] for formulations and solutions of various location problems involving the distance function and its extensions.

1.5 Commentaries to Chapter 1

Section 1.1. The central construction in the developed approach to variational analysis and generalized differentiation is that of the *normal cone* to a locally closed set from Definition 1.1. This construction and the corresponding subdifferential of extended-real valued functions were introduced by the author as a by-product of his method of *metric approximations* in the beginning of 1975 when he was not even familiar with Clarke's work on generalized gradients. It was first written and published in the author's paper [502] (initially rejected!), not in [528] as stated in [375]; there is a reference in [528] to [502] while not vice versa. Following this scheme of [502] for problems of time optimal control with nonsmooth constraints, the initial applications were given in the early papers by the author and Kruger [439, 503, 528] for various optimal control problems. The normal cone notion of [502], widely spread in variational analysis under the names of *basic/general, limiting*, or *Mordukhovich normal cone*, has been the key and striking departure from the conventional scheme of defining a normal cone to a set via duality (1.10) from a tangential approximation, which corresponds for functions to defining a subdifferential via duality from a directional derivative. As discussed above, the latter approach unavoidably leads one to *convex* sets of normals and subgradients, while our constructions are intrinsically *nonconvex*. Besides the inspiration from convex analysis, the underlying idea behind the construction of generalized normals via duality from tangential approximations relates to the well-accepted approach of deriving necessary conditions in constrained optimization by selecting convex subcones of certain tangent cones to sets associated with optimal solutions and then applying a *convex separation theorem*; see, e.g., Dubovitskii and Milyutin [234], Girsanov [296], and Neustadt [606]. A similar idea has been widely implemented in establishing the so-called marginal price equilibria in nonconvex models of welfare economics starting with Guesnerie [313]. The approach suggested in [502] is principally different from all the previous developments and leads us to the robust nonconvex normal cone (1.4) satisfying, together with the associated subdifferential and coderivative constructions for functions and multifunctions, comprehensive *calculus rules* at the points in question, without *any appeal* to tangential approximations.

Note that the convex closure of the basic normal cone

$$\overline{N}(\bar{x};\Omega) := \text{clco}\, N(\bar{x};\Omega), \quad \bar{x}\in\Omega,$$

as in (1.60) agrees with the *convexified/Clarke normal cone* to Ω at $\bar{x}$ defined in [163], based on Clarke's dissertation [162] under the direction of Rockafellar, by the duality scheme (1.10) via the (automatically convex) tangent cone introduced therein. The convexity of these tangent and normal cones provides the possibility to strongly use the machinery of convex analysis and to develop extensive calculus rules and various applications first for the corresponding generalized gradients of locally Lipschitzian functions by Clarke [165] and then for certain non-Lipschitzian cases of sets and functions by Rockafellar [671, 675]. At the same time, it has been realized after a while that the convexity of the normal cone $\overline{N}(\bar{x};\Omega)$ in (1.60) creates serious obstacles in deriving satisfactory necessary optimality conditions and adequate formalizations of marginal price equilibria in economic modeling; see Mordukhovich [507] and Khan [412]. Furthermore, it is proved by Rockafellar [676] that Clarke's normal cone to a *Lipschitzian manifold* of dimension d in $\mathbb{R}^n$ (i.e., a set locally homeomorphic around the point in question to the graph of a locally Lipschitzian mapping) inevitable has to be a *linear subspace* with dimension *greater* than $n-d$ unless the manifold was "strictly smooth" at this point. As an illustration, see the set $\Omega = \text{gph}\,|x|$ in Example 1.14(i), where $\overline{N}((0,0);\Omega) = \mathbb{R}^2$. It shows that for such *graphical* sets, the convexification operation in (1.60) may enlarge the normal cone dramatically to the extent of loosing any useful information about optimality and/or equilibria. Observe to this end that graphical sets always appear in the *coderivative* construction of Definition 1.11 and that, besides graphs of single-valued Lipschitz continuous mappings, Lipschitzian manifolds (or graphically Lipschitzian mappings) include graphs of set-valued *subgradient* mappings for *convex* and more general *prox-regular* functions $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ as well as *maximal monotone* operators, which play a crucial role in many aspects of variational analysis and optimization; see [676] and the books [522, 523, 678] for more details. Note also that the convexification operation in (1.60) may violate the *robustness* property of $\overline{N}(\bar{x};\Omega)$ as for the set $\Omega = \{(x_1,x_2,x_3)\in\mathbb{R}^3 |\ x_3 = |x_1x_2|\}$ at $\bar{x} = 0 \in \mathbb{R}^3$.

Both limiting representations of the normal cone in (1.7) were given in the papers by Kruger and Mordukhovich [440, 441] with the original proof (cf. [522, Theorem 1.6]) different from that presented above. Furthermore, it has been realized in [441, 440] that the *prenormal cone* $\widehat{N}(\bar{x};\Omega)$ in (1.5), known as the *regular* or *Fréchet normal cone*, occurs to be *dual* by Proposition 1.9 to the *contingent/tangent cone* from Definition 1.8 introduced simultaneously and independently by Bouligand [123] and Severi [687] (in fact, this notion goes back to the early work by Peano as well as a number of other notions related to differentiability, tangency, and set limits; see the historical investigation by Dolecki and Greco [218] with the references therein). Thus the combination of the first limiting representation of the normal cone in Theorem 1.6 with the result of Proposition 1.9 shows that the normal cone construction employed by Rockafellar and Wets [678] is equivalent to the original definition of the normal cone (1.4) from [502].

Recall that the original author's construction of the normal cone and its equivalent description in terms of limits of tangents are *finite-dimensional*. The *Banach space* extension of $N(\bar{x};\Omega)$ corresponding to the *second* representation in (1.7) has been suggested in [440, 441] and then further elaborated in Kruger's dissertation [426] conducted under the author's direction and fully reflected in [428, 430] as well as in the books [507, 522, 523] with carefully written commentaries therein. This extension defined the normal cone $N(\bar{x};\Omega)$ via the *weak* sequential* convergence of *ε-normals* in the space X^* dual to a *Fréchet smooth* space X. Symbolically it is represented in the sequential outer limit form

$$N(\bar{x};\Omega) := \mathop{\mathrm{Lim\,sup}}_{x \xrightarrow{\Omega} \bar{x},\ \varepsilon \downarrow 0} \widehat{N}_\varepsilon(x;\Omega)$$

as in (1.58). Note that, although the exact nature of the weak* closure was not explicitly specified in [440, 441], it was clear from the proofs given therein that the weak* limit was taken in the sequential sense due to the classical fact of functional analysis on the weak* sequential compactness of bounded sets in duals to Fréchet smooth spaces. Unfortunately, these well-known observations haven't been reflected in [375]. Note to this end that all the aforementioned publications by the author and Kruger (including the mimeographed papers [427, 428, 440] written in Russian) have been widely distributed from the very beginning among experts on nonsmooth analysis in the former Soviet Union and partly abroad and have been discussed in the seminar and conference meetings.

In the final line of developments in this direction, the author and Shao [580] justified the possibility to equivalently *drop* ε_k in (1.58), i.e., to get the representation

$$N(\bar{x};\Omega) = \mathop{\mathrm{Lim\,sup}}_{x \xrightarrow{\Omega} \bar{x}} \widehat{N}(x;\Omega)$$

from (1.59) as the *definition* of the normal cone in (1.7) via the *sequential* weak* outer limit (1.1) of regular/Fréchet normals (1.5) for closed sets in *Asplund* spaces. Since this class is essentially broader than the Fréchet smooth one considered in the aforementioned work by Kruger and Mordukhovich, the refined construction (1.59) allows us to improve the results obtained therein in the Fréchet smooth setting. Note that the possibility to pass from (1.58) to (1.59) is a highly nontrivial fact based, among other devices, on the *Borwein-Preiss variational principle* [108] and the method of *separable reduction* by Fabian [254]; see more discussions and references in [522]. Recall that Asplund spaces form a remarkable and beautiful subclass of Banach spaces that contains, in particular, every reflexive space (as Fréchet smooth), every space admitting a Fréchet smooth bump function, every space with a separable dual, etc.; see also Exercise 1.41. As shown in [522, 580], the Clarke normal cone (1.61), defined in Banach spaces via the tangential duality [165], reduces to the convexified normal (1.60) provided that the space X is Asplund and the closure operation in (1.60) is taken in the weak* topology of X^*.

In some infinite-dimensional situations, it is useful to consider a modified limiting normal cone construction, where the weak* convergence in (1.58) is replaced by the norm/strong convergence in dual spaces. This has been first done in [277] under the name of the "norm-limiting normal cone" and recently was nicely implemented in [494] under the name of the "strong limiting normal cone" to study optimization problems with complementarity constraints in Lebesgue spaces.

The variational description of regular normals in Theorem 1.10(i) holds in any Banach space as observed in [519], while the more delicate one in (ii) requires a Fréchet smooth renorming; see Fabian and Mordukhovich [257], where the reader can find other versions under some smooth bump geometric assumptions on the space in question. Another proof of the smooth variational description in Theorem 1.10 is given by Rockafellar and Wets [678] in finite dimensions but without the conclusion on convexity of the smooth support function ψ therein.

Section 1.2. The *coderivative* construction of Definition 1.11 was introduced by the author [504] motivated by deriving necessary optimality conditions in optimization problems with nonsmooth equality constraints and describing the adjoint system in the extended Euler-Lagrange conditions for optimal control of differential inclusions. Theorem 1.15 useful in optimal control can be found in [504]. As we see, the coderivative plays a role of a generalized *adjoint derivative* for nonsmooth and set-valued mappings. Note that, being nonconvex-valued, the coderivative $D^*F(\bar{x},\bar{y})$ of $F\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ *cannot* be obtained by duality via *any* tangentially generated derivative of F at $(\bar{x},\bar{y})$; in particular, by the *contingent/graphical derivative*

$$DF(\bar{x},\bar{y})(u) := \big\{v \in \mathbb{R}^m \,\big|\, (u,v) \in T\big((\bar{x},\bar{y});\operatorname{gph} F\big)\big\}, \quad u \in \mathbb{R}^n, \tag{1.87}$$

introduced by Aubin [34] motivated by different applications; see also [36]. Previous developments in this direction based on tangential approximations of graphs can be found in Pshenichnyi [648, 649] as "locally adjoint mappings" for convex-graph and convex-valued multifunctions. Serious disadvantages of the graphical derivative (1.87), which it shares with the contingent and regular normal cones, are nonrobustness and a poor calculus that create obstacles for applications. This is not the case of the basic coderivative D^* along with the basic normal cone from Definition 1.1 and its infinite-dimensional extensions. The principal importance of the coderivative construction in variational analysis is revealed in the subsequent material of this book; see also the prior monographs [507, 522, 523, 678] and the references therein.

For mappings $F\colon X \rightrightarrows Y$ between infinite-dimensional spaces, there are two distinct extensions of our basic coderivative (1.15) from the viewpoint of taking the limit in (1.17), where the *precoderivative* $\widehat{D}^*F$ (known also as the *regular/Fréchet coderivative*) is defined in (1.16) via the prenormal cone $\widehat{N}$ to the graph of F or via the ε-enlargements $\widehat{N}_\varepsilon$ in (1.6) . These enlargements are needed in the case of general Banach spaces X and Y while it suffices to use $\widehat{N}$ when both spaces are Asplund. The first extension, called the *normal coderivative* $D^*_N F(\bar{x}, \bar{y})$, is defined by the same formula (1.15) as in finite dimensions, while the normal cone $N(\cdot; \operatorname{gph} F)$ therein is given by (1.58) or by (1.59) in Asplund spaces, which corresponds to the *weak** convergence of *both* sequences $(x^*_k, -y^*_k)$ from the cone $\widehat{N}((x_k, y_k); \operatorname{gph} F) \subset X^* \times Y^*$ or its ε_k-enlargements; see (1.63). In the second extension, introduced by the author in [514] as the *mixed coderivative* $D^*_M F(\bar{x}, \bar{y})$, we take advantages of the product structure of $X^* \times Y^*$ and use the *strong* convergence of y^*_k in Y^* while still employing the *weak** convergence of x^*_k in X^*. This gives us (1.65), where we can put $\varepsilon = 0$ if both spaces X and Y are Asplund. It is obvious that $D^*_M F(\bar{x}, \bar{y})(y^*) \subset D^*_N F(\bar{x}, \bar{y})(y^*)$, that these coderivatives agree when $\dim Y < \infty$, and that they reduce to the basic coderivative from Definition 1.11 when $\dim X < \infty$ as well. In infinite dimensions the coderivatives D^*_N and D^*_M enjoy similar and rather comprehensive (in Asplund spaces) calculus rules while both being important in variational analysis and its applications; see, e.g., [522, 523] and the material presented below.

Finally, we observe here that the extremal property of Theorem 1.15, which holds in terms of the *normal* coderivative for convex-valued multifunctions between arbitrary Banach spaces, shows in particular that the coderivative Euler-Lagrange condition in optimal control of Lipschitzian differential inclusions obtained first by the author in [504] yields the Weierstrass-Pontryagin maximum condition in the convex-valued setting; see [507, 522, 523] for further discussions.

Section 1.3. Subdifferential theory of variational analysis has started from convex analysis with the fundamental developments by Fenchel, Moreau, and Rockafellar on generalized differentiation of convex functions; see the books [105, 352, 667, 678] for historical comments and references and also the recent one [537] for simplified proofs of the key results of convex subdifferential calculus based on a geometric variational approach. In particular, convex analysis offers a monumental idea of a *set-valued* extension of the classical gradient to nondifferentiable functions known now as the *subdifferential* or the *subgradient mapping*. Considering extended-real-valued functions, convex analysis strongly unifies analytic and geometric ideas revolving around convexity. Observe to this end that, although the classical definition of the convex subdifferential (1.35) is analytic, geometric considerations based on *convex separation/supporting hyperplane* theorems have been behind major results of subdifferential theory for convex functions including their subdifferentiation ($\partial\varphi(\bar{x}) \neq \emptyset$), on the relative interior of the domain, subdifferential calculus rules, etc. It has also been realized from the early days of convex analysis that subgradients of convex functions can be obtained geometrically via epigraphical normals as in (1.24). Nonconvex counterparts of these geometric ideas are underlying in subdifferential theory for general functions developed in this book following those in [507, 522].

On the other hand, it is well known that every convex function $\varphi\colon X \to \overline{\mathbb{R}}$ on a Banach space X (and in more general linear topological settings) admits at any $\bar{x} \in \operatorname{dom}\varphi$ and $w \in X$ the (one-sided) *directional derivative* $\varphi'(\bar{x}; w)$ as in (1.84), which is convex in w and generates the subdifferential (1.35) by the duality

$$\partial\varphi(\bar{x}) = \left\{ v \in X^* \,\middle|\, \langle v, w\rangle \le \varphi'(\bar{x}; w) \ \text{ for all } \ w \in X \right\}. \tag{1.88}$$

This duality scheme has become the dominating source of constructing various subdifferentials for nonconvex functions via appropriately defined generalized directional derivatives known under different names; see, e.g., (1.42) and the comprehensive comments on such constructions in the books [522, 678] with the references therein. The most successful attempt in this vein is generalized directional derivative $\varphi^\circ(\bar{x}; w)$ for locally *Lipschitzian* functions defined by Clarke [163] as in (1.77). The Lipschitz continuity of φ around $\bar{x}$ and the automatic convexity of (1.77) in directions are very essential for satisfactory properties of the corresponding *generalized gradient* $\overline{\partial}\varphi(\bar{x})$ obtained from (1.77) by the duality scheme (1.88) as given in (1.78). To recover the generalized gradient [163] for *l.s.c.* functions by scheme (1.88), Rockafellar [669] introduced a significantly more complicated directional derivative. It reduces to (1.77) for locally Lipschitzian functions φ while loosing some nice properties known for (1.78) in the Lipschitzian case. In particular, robustness and certain important calculus rules are generally lost for this construction even in finite dimensions.

It has been really surprising from the beginning that, despite its nonconvexity and no relation to any directional derivative, these and much better properties hold for the *basic/limiting subdifferential* (1.24) from Definition 1.18 that appeared first in Mordukhovich [502] and then employed in a number of publications (not so many though till 1988) summarized in the book [507]; see also the commentaries in [522, 678] for major developments and references during that period. ‘ Overall, it has been achieved by developing *variational/extremal principles* and *techniques* which are at the core of variational analysis; see more on it in the next Chapter 2.

Let us now comment on the main results presented in Section 1.3 and their infinite-dimensional extensions. The subdifferential description of locally Lipschitzian functions in Theorem 1.22 and the *singular subdifferential* construction (1.25) were given in Kruger and Mordukhovich [504] , while the singular subdifferential representation (1.38) in an equivalent limiting form via proximal subgradients was established by Rockafellar [672] together with the singular subdifferential characterization of local Lipschitz continuity given in Theorem 1.22; see also [678], where the first representation in (1.38) is taken for the definition of $\partial^\infty\varphi(\bar{x})$ as the collection of “horizon subgradients” of φ at $\bar{x}$, and [522] for infinite-dimensional extensions. Observe that the original proof of (1.38) in [672] (reproduced in [678, Theorem 8.9]) and those given in various infinite-dimensional settings [110, 370, 470, 522, 655] are heavily technically involved. Note also that the singular subdifferential characterization $\partial^\infty\varphi(\bar{x}) = \{0\}$ of local Lipschitz continuity for functions on Asplund spaces obtained in [522, Theorem 3.52] requires the additional “sequential normal epi-compactness” condition on φ at $\bar{x}$, which is automatic in finite dimensions.

Recall that, in contrast to classical analysis with its plus-minus symmetry for derivatives, convex analysis is “unilateral” (the expression of Moreau [593]). The negation of a convex function φ is not convex anymore (except of the linear case), and the generalized differential properties of $-\varphi$ are significantly different from those for φ. The subdifferential of a *concave* function $\varphi\colon X \to [-\infty, \infty)$ at $\bar{x}$ with $\varphi(\bar{x}) > -\infty$ is defined by Rockafellar [667] as $\partial\varphi(\bar{x}) := -\partial(-\varphi)(\bar{x})$ also being called the “superdifferential” of φ or—even better—the “upper subdifferential” of φ at this point. The situation is different for Clarke’s generalized gradient of locally Lipschitzian function, which possesses the classical *plus-minus symmetry* $\overline{\partial}(-\varphi)(\bar{x}) = -\overline{\partial}\varphi(\bar{x})$ and thus *doesn’t distinguish* between convex and concave functions as well as between maxima and minima. It seems to be rather unnatural for nonsmooth functions and doesn’t follow the line of convex analysis.

In the case of our basic subdifferential from Definition 1.18, we don’t have such a symmetry, and it makes sense to consider along with the (lower) subdifferential constructions (1.24) and (1.25) their *upper* counterparts $\partial^+\varphi(\bar{x})$ and $\partial^{\infty,+}\varphi(\bar{x})$ defined by (1.74), which may be significantly different from the lower ones as, e.g., for the simplest one-dimensional function $\varphi(x) := |x|$, where $\partial^+\varphi(0) = \{-1, 1\}$. Furthermore, the unions of the lower and upper constructions, defined as the *symmetric* basic and singular subdifferentials in (1.75), enjoy the plus-minus symmetry $\partial^0(-\varphi)(\bar{x}) = -\partial^0\varphi(\bar{x})$ and $\partial^{\infty,0}(-\varphi)(\bar{x}) = -\partial^{\infty,0}\varphi(\bar{x})$. Note that these symmetric constructions are generally nonconvex, that for convex functions φ, the set $\partial^0\varphi(\bar{x})$ reduces to the subdifferential of convex analysis, while for locally Lipschitzian functions, it may be essentially smaller than $\overline{\partial}\varphi(\bar{x})$. In fact, for functions $\varphi\colon \mathbb{R}^n \to \mathbb{R}$ Lipschitzian around $\bar{x}$, we have by (1.83) that

Clarke's generalized gradient $\overline{\partial}\varphi(\bar{x})$ is the convex hull of each of the sets $\partial\varphi(\bar{x})$, $\partial^+\varphi(\bar{x})$, and $\partial^0\varphi(\bar{x})$.

Let us illustrate these relationships for the functions $\varphi\colon \mathbb{R}^2 \to \mathbb{R}$ from Example 1.31. For $\varphi(x_1, x_2) = |x_1| - |x_2|$ from (i) therein, we have

$$\partial^+\varphi(0,0) = \big\{(-1, v) \in \mathbb{R}^2 \big| -1 \le v \le 1\big\} \cup \big\{(1, v) \in \mathbb{R}^2 \big| -1 \le v \le 1\big\},$$

which yields that $\partial^0\varphi(0,0)$ is the boundary of the unit square in $\mathbb{R}^2$, while $\overline{\partial}\varphi(0,0))$ is the whole unit square; see Fig. 1.13. For the function $\varphi(x_1, x_2) = \big|\,|x_1| + x_2\big|$ in Example 1.31(ii), we have the subdifferential calculations

$$\partial^+\varphi(0,0) = \big\{(v, -1) \in \mathbb{R}^2 \big| -1 \le v \le 1\big\} \cup \big\{(1, -1)\big\} \cup \big\{(-1, 1)\big\}$$

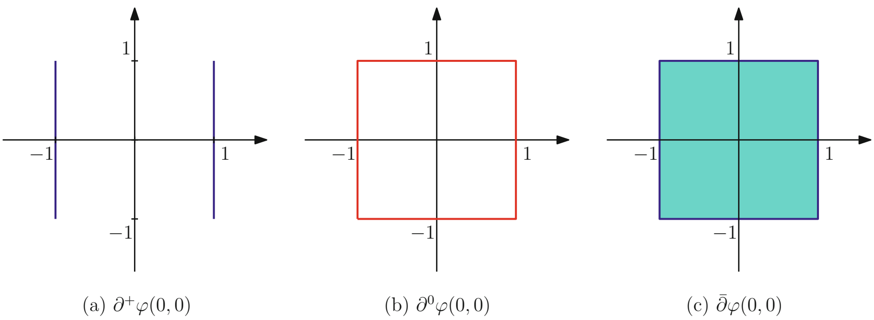


Fig. 1.13 Different subdifferentials of $\varphi(x_1, x_2) = |x_1| - |x_2|$

and thus $\partial^0\varphi(0,0) = \partial\varphi(0,0) \cup \{(v, -1) | -1 \le v \le 1\}$, where $\partial\varphi(0,0))$ is calculated in Example 1.31(ii), while $\overline{\partial}\varphi(0,0)$ is again the whole unit sphere in $\mathbb{R}^2$. Note that this function is taken from Warga [736], where his *derivate container* $\Lambda^0\varphi(0,0)$ is also depicted on this figure. It is proved in [507, Theorem 2.3] (for Lipschitzian functions in [440]) that for any function $\varphi\colon \mathbb{R}^n \to \mathbb{R}$ continuous around $\bar{x}$, we have $\partial^0\varphi(\bar{x}) \subset \Lambda^0\varphi(\bar{x})$; see also [522, Corollary 2.48] and the references therein for infinite-dimensional extensions including mappings between Banach spaces (Fig. 1.14).

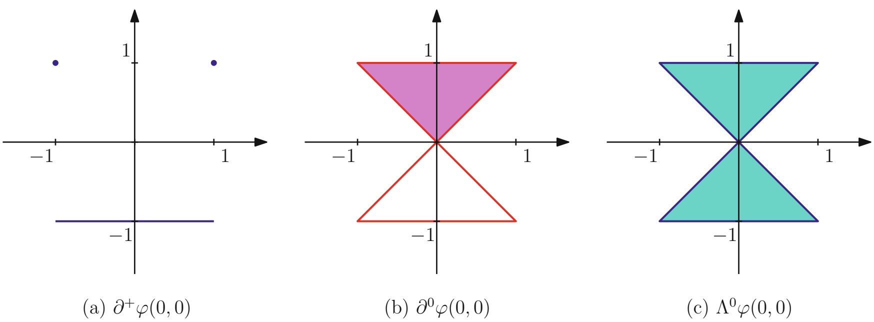


Fig. 1.14 Different subdifferentials of $\varphi(x_1, x_2) = \big|\,|x_1| + x_2\big|$

Theorem 1.23 appears here for the first time in the general case of l.s.c. functions, and its proof holds in any Banach space by taking into account definition (1.58) of the normal cone in this setting. When φ is continuous around $\bar{x}$, a somewhat different proof was given by the author in [522, Theorem 1.80].

To discuss next the limiting connections between the basic subdifferential and its presubdifferential/regular counterpart, note that the latter construction appeared first in Bazaraa, Goode, and Nashed [74] in finite-dimensional spaces under the name of "the set of $\geq$-gradients." Then it has been used in many publications under various names; in particular, as the "Fréchet subdifferential" by analogy with the classical Fréchet derivative (1.12); see, e.g., [114, 375, 522]. The term "regular subgradient" for any $v \in \widehat{\partial}\varphi(\bar{x})$ was suggested in Rockafellar and Wets [678] motivated probably by the property of *lower regularity* $\partial\varphi(\bar{x}) = \widehat{\partial}\varphi(\bar{x})$ holding for certain classes of nice functions such as smooth, convex, amenable ones, etc. Note that $\widehat{\partial}\varphi(\bar{x})$ is also known as the "subdifferential in the sense of viscosity solutions" (or the viscosity subdifferential as suggested by Borwein and Zhu [113, 114]) and has been widely used, starting with the paper by Crandall and Lions [183], in partial differential equations of the Hamilton-Jacobi type with a great many applications; see, e.g., [67, 136, 182] and the bibliographies therein. Finally, the "presubdifferential" (similarly "prenormal" and "precoderivative") terminology comes from the abstract presubdifferential theory by Thibault and Zagrodny [710], where the regular/Fréchet-like constructions take a prominent role.

Along with the set $\widehat{\partial}\varphi(\bar{x})$ of lower subgradients, its *upper* counterpart $\widehat{\partial}^+\varphi(\bar{x}) := -\widehat{\partial}(-\varphi)(\bar{x})$ was also introduced in [74] under the name of "the set of $\leq$ gradients" and then was called in [183] the *superdifferential* in the sense of viscosity solutions. It is easy to see that the sets $\widehat{\partial}\varphi(\bar{x})$ and $\widehat{\partial}^+\varphi(\bar{x})$ are nonempty simultaneously if and only if φ is *Fréchet differentiable* at $\bar{x}$. Thus, contrary to (1.75), the corresponding "symmetric" set $\widehat{\partial}^0\varphi(\bar{x}) := \widehat{\partial}\varphi(\bar{x}) \cup \widehat{\partial}^+\varphi(\bar{x})$ doesn't play any independent role, since it always reduces to either $\widehat{\partial}\varphi(\bar{x})$ or $\widehat{\partial}^+\varphi(\bar{x})$.

Considering the ε-subgradient sets (1.34) goes back to Kruger and Mordukhovich [441, 440] motivated by seeking a convenient description of basic subgradients in Banach spaces corresponding to their second representation in (1.37) of Theorem 1.28. Note that the Fréchet-type ε-subgradients $\widehat{\partial}_\varepsilon\varphi(\bar{x})$ for convex functions are different from the approximate ε-subgradients $\partial_\varepsilon\varphi(\bar{x})$ in the sense of convex analysis; see Proposition 1.25 for the representation of $\widehat{\partial}_\varepsilon\varphi(\bar{x})$ in the convex case while

$$\partial_\varepsilon\varphi(\bar{x}) := \big\{x^* \in X^* \big|\ \varphi(x) - \varphi(\bar{x}) \leq \langle x^*, x - \bar{x}\rangle + \varepsilon \ \text{ whenever } \ x \in X\big\}$$

for the approximate ε-subdifferential of convex analysis. The exact formulations and the presented proof of the relationships with ε-normals in Theorem 1.26 are due to Kruger [427, 430]. The first representation of basic subgradients in (1.37) in finite dimensions follows directly from properties of the Euclidean norm exploited in Theorem 1.6 and thus shows that general subgradients in Rockafellar and Wets [678] are the same (1.24) as introduced by the author [502]. However, the validity of this representation (without $\varepsilon > 0$ involved) in an arbitrary *Asplund* space is a deep variational fact revealed by the author and Shao [580] based on the previous developments; cf. the normal cone commentaries above and the book [522] for more details and references. Note that the first representation in (1.37) for any l.s.c. function $\varphi\colon X \to \overline{\mathbb{R}}$ is actually a *characterization* of Asplund spaces as shown by Fabian and Mordukhovich [257]; see also [522].

The smooth variational description of regular subgradients from Theorem 1.27 was established in [257] in Fréchet smooth spaces, where it was shown that this smooth renorming assumption is also *necessary* for the *concavity* of the smooth support function ψ in Theorem 1.27; see [257] for other smooth variational descriptions in infinite-dimensional spaces. A weaker version of this result without the convexity property of ψ in finite dimensions was given in [678, Proposition 8.5] based on the reduction to the corresponding description of regular normals.

The dual representation (1.40) of regular subgradients via the contingent directional derivative introduced by Penot [634] in form (1.42) follows directly from the definitions, while this fact is essentially *finite-dimensional*; see [522] and the references therein for some analogs of (1.9) and (1.40) via the *weak* contingent cone and the *weak* contingent derivative in *reflexive* spaces.

The other line of extensions of the author's generalized differential constructions to objects in infinite-dimensional spaces has been developed by Ioffe in the series of publications starting from 1981 under different names (M-subdifferential, analytic and geometric "approximate" subdifferentials, their nuclei, etc.), while all of them reduce to [502] in finite dimensions. He was well familiar with and fully acknowledged the previous aforementioned developments by the author in finite dimensions and then with the joint work by the author and Kruger [440, 441] and by Kruger alone [427, 428] in Fréchet smooth spaces. It is written, e.g., in the first part [364] of Ioffe's original work [364, 365] on "approximate subdifferentials" containing the core of his subsequent developments in this direction: "It all essentially arises from thinking over Mordukhovich's approximate approach to necessary conditions for an extremum [502]." This is reflected by the "approximate" term for such subgradients that doesn't correspond to the conventional approximate subgradients used in convex analysis; see, e.g., the book [352].

We are not going to discuss here the essence of the "approximate subdifferentials" and their comparison with the our basic subdifferential constructions in infinite dimensions while referring the reader to [522, Subsections 2.6.9 and 3.2.3] and the commentaries therein for a full account. Note that the best of his constructions, called "nuclei of the geometric subdifferential and the geometric normal cone" [369], satisfy strong calculus rules in general Banach spaces, being however significantly more complicated and always larger than our basic sequential constructions discussed above. Observe to this end that the claims made in [369, Proposition 8.2] and [370, Theorem 1] about the relationships between the "approximate" and our constructions in infinite dimensions are incorrect; in fact, the *opposite* inclusions hold *strictly* even for Lipschitz continuous functions in $\mathcal{C}^\infty$-smooth spaces as shown by Borwein and Fitzpatrick [101]; see also [522, Example 3.61]. The mistakes in the proofs of [369, 370] came from the confusion between the sequential and topological weak* closures. Comprehensive relationships between the sequential limiting and "approximate" subdifferentials of *integral functionals* in the $L^1(T;\mathbb{R}^n)$ (non-Asplund) space have been recently established by Jourani and Thibault [403].

Let us proceed with commentaries on other topics and results presented in Sections 1.3 and 1.4. The *scalarization formula* of Theorem 1.32 was first obtained by Kruger [426, 428] for locally Lipschitzian mappings $f\colon X \to \mathbb{R}^m$ on Banach spaces; cf. also [368] when X is finite-dimensional. The extension of this result to the *mixed* coderivative (1.65) of Lipschitzian mappings between arbitrary Banach spaces was given by the author and Shao [584]; see also [522, Theorem 1.90]. The *normal* coderivative counterpart of the scalarization is significantly more involved; see [580] for mappings $f\colon X \to Y$ from Asplund to general Banach spaces that are *strictly Lipschitzian* at $\bar{x}$; as shown in [709], this notion goes back to the basic version of "compactly Lipschitzian" behavior introduced and studied by Thibault [704] in connection with subdifferential calculus for vector mappings . An improved version of the normal coderivative scalarization result was derived by the author and Wang [590] for the weaker "w^*-strictly Lipschitzian" mappings and was also presented in [522, Theorem 3.28] with more discussions therein.

The classical *distance function* $d_\Omega(x)$ is intrinsically nondifferentiable while Lipschitz continuous, and its generalized differentiation has played a significant role in nonsmooth analysis from the very beginning. It has been well recognized the importance of the distance function in implementing *variational techniques* involving, e.g., penalization in constrained optimization and via the powerful *Ekeland's variational principle* [249, 250]. Theorem 1.33 in finite-dimensional Euclidean spaces goes back to [507, Proposition 2.7] and [678, Example 8.53] (with a new and complete proof given here), while its infinite-dimensional versions are significantly more involved; see the book [522] and the more recent paper [535] for a comprehensive account.

To the best of our knowledge, ε-subgradients of the distance function d_Ω for closed subsets of Banach spaces were first calculated by Kruger [427] for any $\varepsilon \geq 0$ at both in-set and out-of-set points. However, his proof in the out-of-set case via ε-normals to the ρ-enlargement

$$\Omega(\rho) := \big\{x \in X\big|\ d_\Omega(x) \leq \rho\big\} \ \text{ with } \ \rho = d_\Omega(\bar{x})$$

was incomplete, and then it was further clarified by Bounkhel and Thibault [125]. The in-set case for $\widehat{\partial} d_\Omega(\bar{x})$ in Theorem 1.33(i) was also considered by Ioffe [370], while the result for the basic subdifferential $\partial d_\Omega(\bar{x})$ at $\bar{x} \in \Omega$ in general Banach spaces was first derived by Thibault [706] by using Ekeland's variational principle.

Observe that the out-of-set point results in Theorem 1.33(ii) are essentially *finite-dimensional* and depend on the *Euclidean* norm on $\mathbb{R}^n$. Their various infinite-dimensional counterparts for regular and basic subgradients of d_Ω via the corresponding normals to the projection Π_Ω as well as to the enlargement $\Omega(\rho)$ were obtained by the author and Nam [530, 531]. In particular, it was revealed there the *failure*—even in finite dimensions—of an expected counterpart of the relationship between $\partial d_\Omega(\bar{x})$ at $\bar{x} \notin \Omega$ and the normal cone to the enlargement $\Omega(\rho)$ similar to that in Theorem 1.33(i) for $\bar{x} \in \Omega$. To get an appropriate version of this result, the *right-sided subdifferential* of $\varphi\colon X \to \overline{\mathbb{R}}$ at $\bar{x} \in \operatorname{dom}\varphi$ was introduced in [530] by

$$\partial_{\geq}\varphi(\bar{x}) := \operatorname*{Lim\,sup}_{x \overset{\varphi+}{\to} \bar{x}} \widehat{\partial}\varphi(x), \tag{1.89}$$

where the symbol $x \overset{\varphi+}{\to} \bar{x}$ indicated that $x \to \bar{x}$ with $\varphi(x) \to \varphi(\bar{x})$ and $\varphi(x) \geq \varphi(\bar{x})$. Then it was shown therein (see also [522, Theorem 1.101]) that

$$N\big(\bar{x}; \Omega(\rho)\big) = \bigcup_{\lambda \geq 0} \lambda \partial_{\geq} d_\Omega(\bar{x}) \ \text{ with } \ \rho = d_\Omega(\bar{x})$$

for closed subsets of arbitrary Banach spaces. Some extended and axiomatically defined versions of the right-sided subdifferential for the distance function $\partial d_\Omega(\bar{x})$ at the out-of-set point $\bar{x}$ were introduced by the author and Mou [529] under the names of the sequential and topological *outer regular subdifferentials*. These constructions were efficiently used in [529] to derive necessary optimality conditions for optimization problems on metric spaces with inclusion constraints given in arbitrary Banach spaces via *approximately convex sets* in the sense of Ngai, Luc, and Théra [609]. Yet another enhanced version of the right-sided subdifferential (1.89) with replacing "$\geq$" by the strict inequality "$>$" was defined by Ioffe and Outrata [376] in finite-dimensional spaces under the name of the *outer subdifferential* similar to [529] while the constructions are essentially different. This outer subdifferential of [376] and the corresponding notion of *outer coderivative* were utilized in [376] and then in [138, 465, 637] for various applications to optimization and related topics. We refer the reader to the recent paper by Ivanov and Thibault [381] for the impressive usage of the right-sided subdifferential (1.89) in the study of minimum time functions.

Another interesting research topic is subdifferentiation of nonsmooth *integral functionals* (generalized Leibniz rules), which has received a growing interest over the recent years from both viewpoints of variational theory and applications; see, e.g., [1, 149, 169, 295, 330, 572, 636] and the references therein. In particular, the papers by Ackooij and Henrion [1] and by Hantoute, Henrion, and Pérez-Aros [330] contain impressive results via basic subgradients and generalized gradients in the framework of probability functions for parameter-dependent random inequality systems under the Gaussian distribution. The results by Mordukhovich and Sagara [572] concern nonsmooth versions of the Leibniz rule, in terms of the aforementioned subdifferential constructions, for Gelfand integral functionals on general measure spaces as well as on those with *saturated* measures, where the rather involved *weak***-closure* operation for integral values can be avoided. The applications of there results are given in [572] to stochastic dynamic programming and economic modeling. Economically motivated deterministic applications of these subgradients to cooperative games suggested by Sagara [683] have been recently provided by Adam and Kroupa [4].

The *directional limiting subdifferential* together with the corresponding normals and coderivatives was introduced and investigated in the joint papers by the author and Ginchev [293, 294] with involving *tangential directions* in the limiting process. These constructions were used in [293, 294] for deriving more selective necessary conditions in constrained optimization. Strong results in this vein for directional metric regularity and subregularity were established by Gfrerer [281, 282] with various applications to optimization; see also Thinh and Chuong [711] for further developments and applications to multiobjective problems. We specially emphasize the recent papers by Gfrerer and Outrata [287, 289] who obtained, by developing a primal-dual directional variational approach, efficient conditions for Lipschitzian stability of solution maps to parametric generalized equations and applied them to broad spectrum of problems in mathematical programming with equilibrium constraints, conic programming, etc. Note to this end that the "directional" terminology proposed by Penot [637] concerns constructions of the Dini-Hadamard type and their limits, which are completely different from those discussed above. The reader can find more information about other subdifferential (in particular, moderate/Michel-Penot and linear/Treiman) as well as subderivative constructions for extended-real-valued functions, used in variational analysis and not considered here, in [496, 522, 637, 685, 678, 715, 716] and the references therein.

Yet another topic of recent developments concerns subdifferential properties and recent applications of the so-called *minimum time functions* defined by

$$\tau_F(x;\Omega) := \inf_{z\in\Omega} p_F(z-x), \quad x \in X, \tag{1.90}$$

where $F \subset X$ is a closed, convex, and bounded with $0 \in \operatorname{int} F$, where

$$p_F(u) := \inf\{t > 0 \big|\ t^{-1}u \in F\big\}$$

is its *Minkowski gauge*, and where $\Omega \subset X$ is a closed while generally nonconvex *target* set. When F is the closed unit ball $\mathbb{B}$ of the space X, we have $p_F(u) = \|u\|$, and (1.90) reduces to the distance function d_Ω. It has been well recognized that the minimum time functions generated by various sets F and Ω play an important role in many aspects of variational analysis, optimization, control theory, partial differential equations, approximation theory, etc.; see, e.g., [66, 136, 171, 176, 334, 380, 381, 534, 535, 601] and the references therein, where some subgradient properties of (1.90) and their applications can be found. We mention, in particular, the papers [171, 176, 334, 380, 381, 534, 535, 601] for various results in this direction involving the aforementioned subdifferential constructions. Furthermore, recently some of these subdifferential results have been successfully applied in [122, 535, 536, 537, 542, 543, 544, 602, 604] and other publications to solving a number of *facility location problems* whose original versions go back to Fermat, Torricelli, Sylvester, Steiner, and Weber. Strong interest has been revived to investigate problems of this type due to their importance in location science, optimal networks, wireless communications, etc.; see [13, 488, 615, 616, 682] and the references therein.

Recent years have witnessed a rapidly growing interest to *algorithmic* aspects of optimization involving basic subgradients and their applications to *numerical analysis*; see, e.g., [31, 32, 33, 46, 47, 48, 71, 72, 73, 118, 78, 81, 90, 91, 92, 134, 148, 190, 229, 233, 231, 306, 309, 310, 315, 344, 346, 333, 413, 422, 452, 457, 460, 458, 465, 466, 467, 480, 566, 640, 639, 679, 680, 761, 762] among other publications. In particular, a largely unexplored algorithmic area concerns the usage of basic subgradients in *automatic/algorithmic differentiation* [308]; see [68, 307, 309, 410] for related results and discussions in some special settings highly important in applications. Note that the papers [68, 410] impressively demonstrate algorithmic advantages of Nesterov's *lexicographical differentiation* [605] for these classes of nonsmooth functions.

Section 1.4. This section collects some additional material related to the basic content of Section 1.1–1.3 and infinite-dimensional extensions of the results presented therein. Along with rather simple exercises that require just the clear understanding of the basic material and performing calculations, the reader can find in Section 1.4 more involved results with the hints to solving the problems and the references to the corresponding publications. We specially emphasize the *unsolved issues* concerning the development of adequate calculus rules for the right-sided subdifferential (1.53), which are largely open in both finite and infinite dimensions; see Exercise 1.84. The same can be said about its ">" (outer) counterpart discussed in the commentaries above and the corresponding outer coderivative from [376]. Resolving these issues would be of great importance for various applications.

Chapter 2
Fundamental Principles of Variational Analysis

This chapter is devoted to the exposition and developments of the fundamental principles of variational analysis, which play a crucial role in resolving many issues of variational theory and applications by employing optimization ideas and techniques. In our geometric dual-space approach to variational analysis, the major result in this direction is the *extremal principle* for closed sets, which can be treated as a *variational nonconvex* counterpart of the powerful convex separation principle with no presence of convexity. We derive the basic version of the extremal principle for finitely many sets and then continue with new developments for countable set systems. Related variational principles for extended-real-valued functions are also discussed in the main exposition here as well as in the exercise and commentary parts. As a direct consequence of the extremal principle, we establish in this chapter the normal cone intersection rule, which is the *key result* of the nonconvex generalized differential calculus allowing us to derive via a geometric approach comprehensive calculus rules for the robust generalized differential constructions of Chapter 1. Roughly speaking, the extremal principle and related variational ideas are solely responsible for the validity of comprehensive calculus rules for the nonconvex limiting constructions under consideration that occur to be essentially better in comparison with their convex counterparts. This partly justifies the name of "variational analysis" for our discipline.

2.1 Extremal Principle for Finite Systems of Sets

In this section we define and study the notion of local extremality of a given point relative to a system of finitely many sets.

B. S. Mordukhovich, *Variational Analysis and Applications*, Springer Monographs in Mathematics, https://doi.org/10.1007/978-3-319-92775-6_2

2.1.1 The Concept and Examples of Set Extremality

We begin with the definition of the extremal system of finitely many sets. Although it is not used in the definition, suppose unless otherwise stated that all the sets are *locally closed* around the point in question. This is our standing assumption, which is indeed needed in the proofs of the basic extremal principle and the related results presented below.

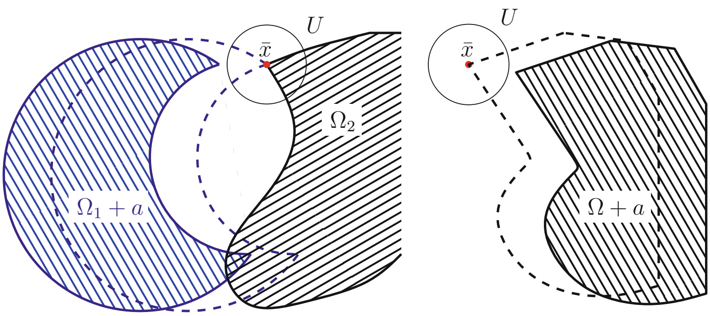
Definition 2.1 (Local Extremality of Finitely Many Sets). *Let $\Omega_1, \ldots, \Omega_s$ with $s \geq 2$ be nonempty subsets of $\mathbb{R}^n$, and let $\bar{x}$ be their common point. We say that $\bar{x}$ is a* LOCALLY EXTREMAL POINT *of the set system $\{\Omega_1, \ldots, \Omega_s\}$ if there are sequences $\{a_{ik}\} \subset \mathbb{R}^n$ for $i = 1, \ldots, s$ and a neighborhood U of $\bar{x}$ such that $a_{ik} \to 0$ as $k \to \infty$ and*

$$\bigcap_{i=1}^{s} \Big(\Omega_i - a_{ik}\Big) \cap U = \emptyset \ \text{ for all large } \ k \in \mathbb{N}. \tag{2.1}$$

In this case we say that $\{\Omega_1, \ldots, \Omega_s, \bar{x}\}$ is an EXTREMAL SYSTEM *in $\mathbb{R}^n$.*

In the sequel we'll drop the word "locally" for $\bar{x}$ in Definition 2.1 if $U = \mathbb{R}^n$ in (2.1). In fact, it is possible to assume without loss of generality that $U = \mathbb{R}^n$ in all the (local) results below concerning locally extremal points.

Geometrically the local extremality of sets at a common point means that they can be locally "pushed apart" by a small perturbation/translation of at least one of them. For $s = 2$, the local extremality of $\{\Omega_1, \Omega_2, \bar{x}\}$ can be equivalently described as follows: there exists a neighborhood U of $\bar{x}$ such that for any $\varepsilon > 0$, there is $a \in \varepsilon\mathbb{B}$ with $(\Omega_1 + a) \cap \Omega_2 \cap U = \emptyset$; see Fig. 2.1(a). Obviously, the condition $\Omega_1 \cap \Omega_2 = \{\bar{x}\}$ doesn't necessarily imply that $\bar{x}$ is a locally extremal point of $\{\Omega_1, \Omega_2\}$. A simple example is given by the two sets on the plane $\Omega_1 := \{(v, v) |\ v \in \mathbb{R}\}$ and $\Omega_2 := \{(v, -v) |\ v \in \mathbb{R}\}$.



(a) $(\Omega_1 + a) \cap \Omega_2 \cap U = \emptyset$ (b) $\bar{x} \in \text{bd}\, \Omega$

Fig. 2.1 Extremal systems of sets

It is easy to see that any *boundary point* $\bar{x}$ of a closed set Ω is a locally extremal point of the pair $\{\Omega_1, \Omega_2\}$ with $\Omega_1 := \Omega$ and $\Omega_2 := \{\bar{x}\}$; see Fig. 2.1(b). Furthermore, the geometric notion of local extremality for set systems can be treated as a direct extension of *local optimality* of feasible solutions to optimization problems. Indeed, consider the general problem of *constrained optimization* with the scalar objective given by

$$\text{minimize } \varphi(x) \text{ subject to } x \in \Omega \subset \mathbb{R}^n,$$

where the constraint set Ω is closed and the cost/objective function φ is l.s.c. around $\bar{x}$. It follows directly from the definitions that any locally optimal solution $\bar{x} \in \Omega$ to this problem generates the locally extremal point $(\bar{x}, \varphi(\bar{x}))$ of the system of locally closed sets $\{\Omega_1, \Omega_2\}$ in $\mathbb{R}^{n+1}$ defined by

$$\Omega_1 := \operatorname{epi}\varphi \text{ and } \Omega_2 = \Omega \times \big\{\varphi(\bar{x})\big\}.$$

To verify the extremality condition (2.1) in Definition 2.1, take the sequences $a_{1k} := (0, \nu_k) \subset \mathbb{R}^n \times \mathbb{R}$, $a_{2k} := 0$ and the neighborhood $U = O \times \mathbb{R}$ therein, where $\nu_k \uparrow 0$ and where O is a neighborhood of the local minimizer $\bar{x}$. In the subsequent sections of this chapter and in other chapters of the book, the reader can find many examples of extremal systems in optimization-related (including those of vector and set optimization) and equilibrium problems, variational principles, generalized differential calculus, economic modeling, etc.

Let us now compare the introduced notion of set extremality with the conventional separation property for finitely many sets, not necessarily convex, which have a common point. Recall that such sets $\Omega_i \subset \mathbb{R}^n$ as $i = 1, \ldots, s$ are said to be *separated* if there exist vectors $v_i \in \mathbb{R}^n$, not equal to zero simultaneously, and numbers $\alpha_i \in \mathbb{R}$ for which

$$\langle v_i, x\rangle \le \alpha_i \text{ whenever } x \in \Omega_i, \quad i = 1, \ldots, s, \tag{2.2}$$

$$v_1 + \ldots + v_s = 0, \text{ and } \alpha_1 + \ldots + \alpha_s = 0. \tag{2.3}$$

A crucial issue of this definition is the *existence* of vectors v_i and numbers α_i satisfying (2.2) and (2.3). Although the notion of separation is defined in the general setting, we are able to justify its applicability only in the convex case in connection with set extremality. This is done in the next proposition.

Proposition 2.2 (Extremality and Separation). *Let $\Omega_1, \ldots, \Omega_s$ for $s \ge 2$ be subsets of $\mathbb{R}^n$ having at least one common point. The following hold:*

(i) *If these sets are separated, then the system $\{\Omega_1, \ldots, \Omega_s, \bar{x}\}$ is extremal with $U = \mathbb{R}^n$ for every common point $\bar{x}$ of these sets.*

(ii) *The converse is true if all the sets Ω_i are convex.*

(iii) *Thus the convex sets $\Omega_1, \ldots, \Omega_s$ are separated in $\mathbb{R}^n$ if and only if each of their common point is extremal.*

Proof. Suppose that Ω_i are separated with $v_s \neq 0$ for definiteness. Pick any $a \in \mathbb{R}^n$ with $\langle v_s, a\rangle > 0$ and put $a_k := a/k$ for all $k \in \mathbb{N}$. Let us verify that

$$\Omega_1 \cap \ldots \cap \Omega_{s-1} \cap (\Omega_s - a_k) = \emptyset, \quad k \in \mathbb{N},$$

which obviously implies the extremality of $\{\Omega_1, \ldots, \Omega_s, \bar{x}\}$ for every common point $\bar{x}$. Assuming the contrary and taking any x from the above intersection, we get by the separation property that

$$\langle v_i, x\rangle \leq \alpha_i, \quad i = 1, \ldots, s-1, \quad \text{and} \quad \langle v_s, x + a_k\rangle \leq \alpha_s, \quad k \in \mathbb{N}.$$

Summing this up gives us $\alpha_1 + \ldots + \alpha_s \geq \frac{1}{k}\langle v_s, a\rangle > 0$, which is a contradiction that justifies (i). The converse assertion in (ii) follows from the extremal principle of Theorem 2.3 and the normal cone expression (1.9) for convex sets. Assertion (iii) is a direct consequence of (i) and (ii). △

2.1.2 Basic Extremal Principle and Some Consequences

The next result establishes the underlying extremal principle for systems of finitely many sets in finite-dimensional spaces. It shows, in particular, that the set extremality, but not relationships (2.2) and (2.3), is a natural *variational* counterpart of separation for nonconvex sets, and that the extremal principle is an appropriate *variational* counterpart of the separation theorem in nonconvex settings. The proof of the extremal principle is based on the *method of metric approximations*, which provides a constructive approximation of the extremal set system under consideration by families of *unconstrained optimization* problems with cost functions *smooth* around the points of interest.

Theorem 2.3 (Basic Extremal Principle). *Let $\{\Omega_1, \ldots, \Omega_s, \bar{x}\}$ with $s \geq 2$ be an extremal system in $\mathbb{R}^n$. Then there are basic normals*

$$v_i \in N(\bar{x}; \Omega_i), \quad i = 1, \ldots, s, \tag{2.4}$$

to the sets Ω_i at the locally extremal point $\bar{x}$ such that

$$v_1 + \ldots + v_s = 0 \;\; \textit{and} \;\; \|v_1\|^2 + \ldots + \|v_s\|^2 = 1. \tag{2.5}$$

Proof. Without loss of generality, suppose that $U = \mathbb{R}^n$ in the definition of the extremal point $\bar{x} \in \Omega_1 \cap \ldots \cap \Omega_s$. Take the sequences $\{a_{ik}\}$ from Definition 2.1, and for each $k = 1, 2, \ldots$, consider the following problem of *unconstrained minimization* overall $x \in \mathbb{R}^n$:

$$\text{minimize} \;\; d_k(x) := \Big[\sum_{i=1}^{s} \text{dist}^2(x + a_{ik}; \Omega_i)\Big]^{1/2} + \|x - \bar{x}\|^2. \tag{2.6}$$

Since the function d_k is continuous and its level sets are bounded, there is an optimal solution x_k to (2.6) by the classical Weierstrass theorem. Due to the extremality of $\bar{x}$ in (2.1), we readily have that

$$\alpha_k := \Big[\sum_{i=1}^{s} \text{dist}^2(x_k + a_{ik}; \Omega_i)\Big]^{1/2} > 0.$$

Furthermore, the optimality of x_k in (2.6) ensures that

$$d_k(x_k) = \alpha_k + \|x_k - \bar{x}\|^2 \le \Big[\sum_{i=1}^{s} \|a_{ik}\|^2\Big]^{1/2} \downarrow 0,$$

which implies that $x_k \to \bar{x}$ and $\alpha_k \downarrow 0$ as $k \to \infty$. Now for each $i = 1, \ldots, s$ we pick an arbitrary Euclidean projection $w_{ik} \in \Pi(x_k + a_{ik}; \Omega_i)$ and consider yet another unconstrained optimization problem over $x \in \mathbb{R}^n$:

$$\text{minimize} \;\; \rho_k(x) := \Big[\sum_{i=1}^{s} \|x + a_{ik} - w_{ik}\|^2\Big]^{1/2} + \|x - \bar{x}\|^2, \tag{2.7}$$

which obviously has the same optimal solution x_k as (2.6). Since $\alpha_k > 0$ and the Euclidean norm $\|\cdot\|$ is smooth on $\mathbb{R}^n \setminus \{0\}$, the function $\rho_k(x)$ is continuously differentiable around x_k, and so (2.7) is a *smooth* problem of unconstrained minimization. Thus the classical Fermat stationary rule yields

$$\nabla \rho_k(x_k) = \sum_{i=1}^{s} v_{ik} + 2(x_k - \bar{x}) = 0, \tag{2.8}$$

where $v_{ik} = (x_k + a_{ik} - w_{ik})/\alpha_k$, $i = 1, \ldots, s$, with

$$\|v_{1k}\|^2 + \ldots + \|v_{sk}\|^2 = 1 \;\; \text{for all} \;\; k \in \mathbb{N}.$$

By the compactness of the unit sphere in $\mathbb{R}^n$, we find vectors $v_i \in \mathbb{R}^n$, $i = 1, \ldots, s$, satisfying the nontriviality condition in (2.5) and such that $v_{ik} \to v_i$ as $k \to \infty$. Passing to the limit in (2.8) gives us also the first equation in (2.5). Finally, it follows directly from the definition of basic normals in (1.4) that each v_i satisfies (2.4), which thus completes the proof of the theorem. △

For the case of two sets Ω_1, Ω_2 in the extremal system, the relationships of the extremal principle in Theorem 2.3 reduce to

$$0 \ne v \in N(\bar{x}; \Omega_1) \cap \big(-N(\bar{x}; \Omega_2)\big). \tag{2.9}$$

When both Ω_1 and Ω_2 are *convex*, we have from (2.9) by the normal cone representation for convex sets (1.9) that

$$\langle v, x_1 \rangle \le \langle v, x_2 \rangle \;\; \text{for all} \;\; x_1 \in \Omega_1 \;\; \text{and} \;\; x_2 \in \Omega_2 \;\; \text{with} \;\; v \ne 0,$$

which is the contents of the classical *separation theorem* for two convex sets. This allows us to get a *full characterization* of extremal points of finitely many convex sets via their *relative interiors* ri Ω_i, i.e., the interior of each convex set Ω_i in Theorem 2.3 with respect to its affine hull ; see, e.g., [667].

Corollary 2.4 (Relative Interiority Condition for Extremality of Convex Sets). *A system of convex sets* $\{\Omega_1, \ldots, \Omega_s, \bar{x}\}$ *with* $s \geq 2$ *is extremal at each of their common point* $\bar{x}$ *if we have the condition*

$$\text{ri}\,\Omega_1 \cap \ldots \cap \text{ri}\,\Omega_s = \emptyset. \tag{2.10}$$

Proof. The separation result from [667, Theorem 11.3]) tells us that the condition $\text{ri}\,\Omega_1 \cap \text{ri}\,\Omega_2 = \emptyset$ is necessary and sufficient for the so-called *proper separation* of two convex sets in $\mathbb{R}^n$; hence it yields the usual separation property for two sets. This allows us to conclude by induction that (2.10) ensures the separation of many convex sets in the sense discussed above. Since extremality and separation are equivalent for convex sets by Proposition 2.2, we get (2.10) as a sufficient condition for set extremality. △

Note that the *convexity* of Ω_i is essential for the validity of Corollary 2.4. Indeed, let Ω_1 be the union of the first and third quadrants and Ω_2 be the union of the second and fourth quadrants of the plane with the common point $(0, 0)$, which is not extremal, while condition (2.10) holds; see Fig. 2.2.

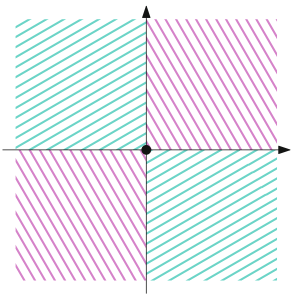


Fig. 2.2 Extremality and relative interior

When $\bar{x}$ is a boundary point of the closed (not necessarily convex) set Ω, applying Theorem 2.3 to the extremal system $\{\Omega, \{\bar{x}\}, \bar{x}\}$ gives us that $N(\bar{x}; \Omega) \neq 0$, i.e., we recover the result of Proposition 1.2.

Observe that the basic extremal principle of Theorem 2.3 is given in the *exact/pointbased* form involving only the locally extremal point $\bar{x}$ in question. The next consequence of Theorem 2.3 in the finite-dimensional setting considered here is the following *approximate extremal principle*, which plays an independent role in infinite dimensions; see Sections 2.5 and 2.6.

Corollary 2.5 (Approximate Extremal Principle). *Let $\{\Omega_1, \ldots, \Omega_s, \bar{x}\}$ with $s \geq 2$ be an extremal system in $\mathbb{R}^n$. Then for any number $\varepsilon > 0$, there are points $x_i \in \Omega_i \cap (\bar{x} + \varepsilon\mathbb{B})$ and approximate normals*

$$v_i \in \widehat{N}(x_i; \Omega_i) + \varepsilon\mathbb{B}, \quad i = 1, \ldots, s, \tag{2.11}$$

such that both relationships in (2.5) are satisfied.

Proof. It follows directly from the extremal principle of Theorem 2.3 and the first limiting representation of basic normals in Theorem 1.6. △

It is easy to see that the result of Corollary 2.5 is in fact *equivalent* to the basic extremal principle of Theorem 2.3 in the finite-dimensional setting under consideration since we can get (2.4) by passing to the limit from (2.11) due to (2.5) and the compactness of the unit sphere in $\mathbb{R}^n$.

2.2 Extremal Principles for Countable Systems of Sets

Next we consider appropriate versions of set extremality and extremal principles for collections of *infinite/countable* systems of sets. This issue is significantly more involved in comparison with finite systems of sets, even in the presence of convexity. The study of extremality of infinite set systems is important for many aspects of variational analysis and optimization, in particular for problems of *semi-infinite programming* considered later in Chapter 8.

2.2.1 Versions of Extremality for Countable Set Systems

In contrast to the constructions and results above concerning finite systems of sets, the following notions of the *conic* and *tangential/contingent extremality* play a crucial role in the study of infinite set systems.

Definition 2.6 (Conic and Contingent Extremal Systems). *We say that:*

(a) *A countable system of cones $\{\Lambda_i\}_{i\in\mathbb{N}} \subset \mathbb{R}^n$ is* EXTREMAL AT THE ORIGIN, *or simply $\{\Lambda_i\}_{i\in\mathbb{N}}$ is an* EXTREMAL SYSTEM OF CONES, *if there is a bounded sequence $\{a_i\}_{i\in\mathbb{N}} \subset \mathbb{R}^n$ such that*

$$\bigcap_{i=1}^{\infty} \big(\Lambda_i - a_i\big) = \emptyset. \tag{2.12}$$

(b) *Let $\{\Omega_i\}_{i\in\mathbb{N}} \subset \mathbb{R}^n$ be a countable system of sets with $\bar{x} \in \cap_{i=1}^{\infty}\Omega_i$, and let $T(\bar{x}; \Omega_i)$ be the contingent cone (1.11) to Ω_i at $\bar{x}$. Then $\{\Omega_i, \bar{x}\}_{i\in\mathbb{N}}$ is a* CONTINGENT EXTREMAL SYSTEM *with the* CONTINGENT LOCALLY EXTREMAL POINT $\bar{x}$ *if the conic system $\{T(\bar{x}; \Omega_i)\}_{i\in\mathbb{N}}$ is extremal at the origin.*

Note that in this way, we can naturally define other types of *tangential extremal systems* by replacing $T(\bar{x}; \Omega_i)$ in Definition 2.6(b) with other tangent cones to Ω_i at $\bar{x}$, but the main tangential extremal principle presented below in this section essentially uses specific properties of the contingent cone.

Observe also that the extremality notions in Definition 2.6 obviously apply to the case of systems containing *finitely* many sets; indeed, in such a case, the other sets reduce to the whole space $\mathbb{R}^n$. It is easy to check that any finite system of cones $\{\Lambda_1, \ldots, \Lambda_s\}$ is extremal at the origin *if and only if* $\bar{x} = 0$ is a locally extremal point of $\{\Lambda_1, \ldots, \Lambda_s\}$ in the sense of Definition 2.1. However, in general the local extremality (2.1) and the contingent extremality from Definition 2.6 are *independent* notions even in the case of two sets in $\mathbb{R}^2$.

Example 2.7 (Contingent Extremality vs. Local Extremality).

(i) Consider the function $\varphi(x) := x\sin(1/x)$ for $x \neq 0$ with $\varphi(0) = 0$, and construct the closed sets in $\mathbb{R}^2$ by

$$\Omega_1 := \operatorname{epi}\varphi \ \text{ and } \ \Omega_2 := (\mathbb{R} \times \mathbb{R}_-) \setminus \operatorname{int}\Omega_1.$$

Take $\bar{x} = (0, 0) \in \Omega_1 \cap \Omega_2$, and observe that the contingent cones to Ω_1 and Ω_2 at $\bar{x}$ are calculated, respectively, by

$$T(\bar{x}; \Omega_1) = \operatorname{epi}\big(-|\cdot|\big) \ \text{ and } \ T(\bar{x}; \Omega_2) = \mathbb{R} \times \mathbb{R}_-.$$

It is easy to conclude that $\bar{x}$ is a locally extremal point of $\{\Omega_1, \Omega_2\}$ but not a contingent locally extremal point of this set system; see Fig. 2.3.

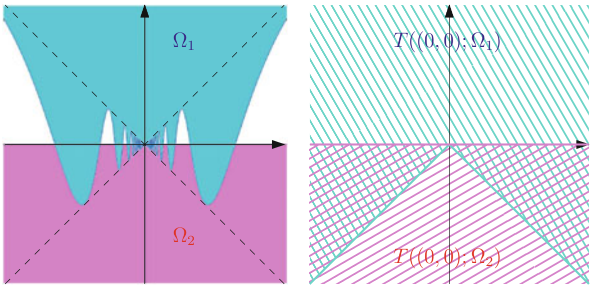


Fig. 2.3 Local extremality but not contingent local extremality

(ii) Define the two closed subsets of $\mathbb{R}^2$ by

$$\Omega_1 := \big\{(x, y) \in \mathbb{R}^2 \big|\ y \geq -x^2\big\} \ \text{ and } \ \Omega_2 := \mathbb{R} \times \mathbb{R}_-.$$

The contingent cones to Ω_1 and Ω_2 at $\bar{x} = (0, 0)$ are $T(\bar{x}; \Omega_1) = \mathbb{R} \times \mathbb{R}_+$ and $T(\bar{x}; \Omega_2) = \mathbb{R} \times \mathbb{R}_-$. It shows that $\{\Omega_1, \Omega_2, \bar{x}\}$ is a contingent extremal system but not an extremal system from Definition 2.1; see Fig. 2.4.

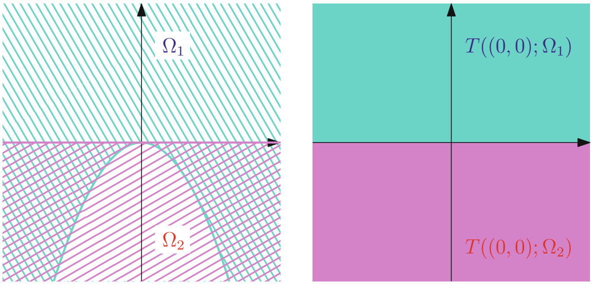


Fig. 2.4 Contingent local extremality but not local extremality

2.2.2 Conic and Contingent Extremal Principles

Our goal now is to derive meaningful *extremality conditions* for *countable* conic and contingent systems from Definition 2.6 via the basic normal cone (1.4) to the sets involved. Let us first formulate and discuss such conditions and then justify them under appropriate assumptions.

Definition 2.8 (Extremality Conditions for Countable Systems of Sets). *Considering countable set systems from Definition 2.6, we say that:*

(a) *The system of cones $\{\Lambda_i\}_{i\in\mathbb{N}}$ in $\mathbb{R}^n$ satisfies the* CONIC EXTREMALITY CONDITIONS *at the origin if there are normals $v_i \in N(0;\Lambda_i)$ for $i \in \mathbb{N}$ with*

$$\sum_{i=1}^{\infty}\frac{1}{2^i}v_i = 0 \quad \textit{and} \quad \sum_{i=1}^{\infty}\frac{1}{2^i}\|v_i\|^2 = 1. \tag{2.13}$$

(b) *The system of sets $\{\Omega_i\}_{i\in\mathbb{N}}$ in $\mathbb{R}^n$ satisfies the* CONTINGENT EXTREMALITY CONDITIONS *at $\bar{x} \in \bigcap_{i=1}^{\infty}\Omega_i$ if the systems of their contingent cones $\{T(\bar{x};\Omega_i\}_{i\in\mathbb{N}}$ satisfies the conic extremality conditions from* (a).

(c) *The system of sets $\{\Omega_i\}_{i\in\mathbb{N}}$ in $\mathbb{R}^n$ satisfies the* NORMAL EXTREMALITY CONDITIONS *at $\bar{x} \in \bigcap_{i=1}^{\infty}\Omega_i$ if there are basic normals $v_i \in N(\bar{x};\Omega_i)$ for $i \in \mathbb{N}$ satisfying the relationships in* (2.13).

It is easy to see that the introduced contingent and normal extremality conditions are *equivalent* if all the sets Ω_i are either *cones* with $\bar{x} = 0$ or *convex* near $\bar{x}$. We will prove below that the *contingent* extremality conditions always *imply* the *normal* ones. However, the opposite implication *doesn't hold* even for systems of two sets in $\mathbb{R}^2$. Indeed, consider the two sets from Example 2.7(i) for which $\bar{x} = (0,0)$ is a locally extremal point in the sense of Definition 2.1. Thus the normal extremality conditions, which reduce in this case to (2.4) and (2.5), hold by the basic extremal principle of Theorem 2.3. On the other hand, we can directly check by the calculation of Example 2.7(i) that the contingent extremality conditions are violated for these sets.

The following *conic extremal principle* (CEP) justifies the validity of the conic extremality conditions from Definition 2.8(a) for any countable extremal systems of nonoverlapping cones. Its proof is based on a countable extension of the *method of metric approximations* used in the proof of Theorem 2.3. The countability of the system requires additional arguments, which take into account the conic structure of the sets involved.

Theorem 2.9 (Conic Extremal Principle). *Let $\{\Lambda_i\}_{i\in\mathbb{N}}$ be an extremal system of cones in $\mathbb{R}^n$ with the* NONOVERLAPPING PROPERTY

$$\bigcap_{i=1}^{\infty}\Lambda_i=\{0\}. \tag{2.14}$$

Then the conic extremality conditions from Definition 2.8(a) *hold. Furthermore, for each $i\in\mathbb{N}$, there is $x_i\in\Lambda_i$ such that $v_i\in\widehat{N}(x_i;\Lambda_i)$ for the corresponding basic normal $v_i\in N(0;\Lambda_i)$ satisfying* (2.13).

Proof. By Definition 2.6(a) of conic extremal systems, find a bounded sequence $\{a_i\}_{i\in\mathbb{N}}\subset\mathbb{R}^n$ with property (2.12), and consider the problem:

$$\text{minimize } \varphi(x):=\left[\sum_{i=1}^{\infty}\frac{1}{2^i}\text{dist}^2(x+a_i;\Lambda_i)\right]^{\frac{1}{2}}\quad\text{over}\;\; x\in\mathbb{R}^n. \tag{2.15}$$

Step 1: *Problem* (2.15) *admits an optimal solution*.

Indeed, since the function φ in (2.15) is continuous on $\mathbb{R}^n$ due the uniform convergence of the series therein, it suffices to show that there is $\alpha>0$ for which the level set $\{x\in\mathbb{R}^n|\ \varphi(x)\le\inf\varphi+\alpha\}$ is bounded and then to apply the classical Weierstrass theorem. Suppose by the contrary that the level sets are unbounded whenever $\alpha>0$ and, for any $k\in\mathbb{N}$, find $x_k\in\mathbb{R}^n$ satisfying

$$\|x_k\|>k\;\text{ and }\;\varphi(x_k)\le\inf\varphi+\frac{1}{k}.$$

Setting $u_k:=x_k/\|x_k\|$ and taking into account that all Λ_i are cones give us

$$\frac{1}{\|x_k\|}\varphi(x_k)=\left[\sum_{i=1}^{\infty}\frac{1}{2^i}\text{dist}^2\Big(u_k+\frac{a_i}{\|x_k\|};\Lambda_i\Big)\right]^{\frac{1}{2}}\le\frac{1}{\|x_k\|}\Big(\inf\varphi+\frac{1}{k}\Big)\to 0$$

as $k\to\infty$. Furthermore, there is $M>0$ such that for large $k\in\mathbb{N}$, we get

$$\text{dist}\Big(u_k+\frac{a_i}{\|x_k\|};\Lambda_i\Big)\le\Big\|u_k+\frac{a_i}{\|x_k\|}\Big\|\le M.$$

Suppose without relabeling that $u_k\to u$ as $k\to\infty$ for some $u\in\mathbb{R}^n$. Passing now to the limit above and employing the uniform convergence of the series therein together with the fact that $a_i/\|x_k\|\to 0$ uniformly in $i\in\mathbb{N}$ due the boundedness of $\{a_i\}_{i\in\mathbb{N}}$, we have

$$\left[\sum_{i=1}^{\infty}\frac{1}{2^i}\text{dist}^2(u;\Lambda_i)\right]^{\frac{1}{2}}=0.$$

This implies by the nonoverlapping condition (2.14) that $u\in\bigcap_{i=1}^{\infty}\Lambda_i=\{0\}$. The latter is impossible due to $\|u\|=1$, which contradicts our intermediate assumption on the unboundedness of the level sets for φ and thus justifies the existence of an optimal solution $\widetilde{x}$ to problem (2.15).

Step 2: *Reduction to smooth unconstrained optimization.*

Observe first that for any closed cone $\Lambda\subset\mathbb{R}^n$ and any $w\in\Lambda$, we have

$$\widehat{N}(w,\Lambda)\subset N(0;\Lambda). \tag{2.16}$$

Indeed, pick any $v\in\widehat{N}(w;\Lambda)$, and get by definition (1.5) that

$$\limsup_{x\xrightarrow{\Lambda}w}\frac{\langle v,x-w\rangle}{\|x-w\|}\le 0.$$

Fix $x\in\Lambda$, $t>0$, and let $u:=x/t$. Then $x/t\in\Lambda$, $tw\in\Lambda$, and

$$\limsup_{x\xrightarrow{\Lambda}tw}\frac{\langle v,x-tw\rangle}{\|x-tw\|}=\limsup_{x\xrightarrow{\Lambda}tw}\frac{t\langle v,(x/t)-w\rangle}{t\|(x/t)-w\|}=\limsup_{u\xrightarrow{\Lambda}w}\frac{\langle v,u-w\rangle}{\|u-w\|}\le 0,$$

which gives us $v\in\widehat{N}(tw;\Lambda)$. Letting $t\to 0$ yields $v\in N(0;\Lambda)$ and so (2.16).

To proceed further, deduce from the cone extremality of $\{\Lambda_i\}_{i\in\mathbb{N}}$ and the construction of φ in (2.15) that $\varphi(\widetilde{x})>0$. Pick any $w_i\in\Pi(\widetilde{x}+a_i;\Lambda_i)$ as $i\in\mathbb{N}$, and get from (2.16) and the proof of Theorem 1.6 that

$$\widetilde{x}+a_i-w_i\in\Pi^{-1}(w_i;\Lambda_i)-w_i\subset\widehat{N}(w_i;\Lambda_i)\subset N(0;\Lambda_i). \tag{2.17}$$

Moreover, the sequence $\{a_i-w_i\}_{i\in\mathbb{N}}$ is bounded in $\mathbb{R}^n$ due to $\|x+a_i-w_i\|=\text{dist}\,(x+a_i;\Lambda_i)\le\|x+a_i\|$. Considering now the unconstrained problem

$$\text{minimize } \psi(x):=\left[\sum_{i=1}^{\infty}\frac{1}{2^i}\|x+a_i-w_i\|^2\right]^{\frac{1}{2}}\quad\text{over}\;\; x\in\mathbb{R}^n, \tag{2.18}$$

observe from $\psi(x)\ge\varphi(x)\ge\varphi(\widetilde{x})=\psi(\widetilde{x})$ that its optimal solution is the same $\widetilde{x}$ as for (2.15). To verify the smoothness of ψ around $\widetilde{x}$, define the function

$$\vartheta(x):=\sum_{i=1}^{\infty}\frac{1}{2^i}\|x-z_i\|^2,\quad x\in\mathbb{R}^n,$$

and show that it is continuously differentiable on $\mathbb{R}^n$ with the derivative

$$\nabla\vartheta(x) = \sum_{i=1}^{\infty} \frac{1}{2^{i-1}}\big(x - z_i\big), \quad x, z_i \in \mathbb{R}^n,$$

Indeed, it is easy to see that both series above converge for every $x \in \mathbb{R}^n$. Taking now any $u, \xi \in \mathbb{R}^n$ with the norm $\|\xi\|$ sufficiently small, we have

$$\|u+\xi\|^2-\|u\|^2-2\langle u, \xi\rangle = \|u\|^2+2\langle u, \xi\rangle+\|\xi\|^2-\|u\|^2-2\langle u, \xi\rangle = \|\xi\|^2 = o(\|\xi\|).$$

Thus it follows for any $x \in \mathbb{R}^n$ and y close to x that

$$\begin{aligned}&\vartheta(y) - \vartheta(x) - \Big\langle \nabla\vartheta(x), y - x\Big\rangle\\ &= \sum_{i=1}^{\infty} \frac{1}{2^i}\Big[\|y - z_i\|^2 - \|x - z_i\|^2 - 2\big\langle x - z_i, y - x\big\rangle\Big]\\ &= \sum_{i=1}^{\infty} \frac{1}{2^i}\|y - x\|^2 = o(\|y - x\|).\end{aligned}$$

This justifies that $\nabla\vartheta(x)$ is the derivative of ϑ at x, which is obviously continuous on $\mathbb{R}^n$. Then the claim follows from the smoothness of the function $\sqrt{t}$ around nonzero points and the fact that $\psi(\widetilde{x}) \neq 0$ due to the cone extremality.

Step 3: *Applying the Fermat stationary rule.*

The above derivative calculation gives us by the stationary principle that

$$\nabla\psi(\widetilde{x}) = \sum_{i=1}^{\infty} \frac{1}{2^i} v_i = 0 \ \text{ with } \ v_i := \frac{1}{\psi(\widetilde{x})}\Big(\widetilde{x} + a_i - w_i\Big), \quad i \in \mathbb{N}.$$

This implies by (2.17) that $v_i \in \widehat{N}(w_i; \Lambda_i) \subset N(0; \Lambda_i)$ for all $i \in \mathbb{N}$. Furthermore, it follows from the constructions of v_i and ψ that

$$\sum_{i=1}^{\infty} \frac{1}{2^i}\|v_i\|^2 = 1,$$

which thus completes the proof of the theorem. △

The following example demonstrates that the setting of Theorem 2.9 is essential for the validity of the extremality conditions therein.

Example 2.10 (Nonoverlapping Property and Conic Structure of Sets Are Essential for the Validity of CEP).

(i) Let us first show that the conclusion of Theorem 2.9 may fail for countable extremal systems of *convex* cones in $\mathbb{R}^2$ if the *nonoverlapping property* (2.14) is *violated*. Define the convex cones $\Lambda_i \subset \mathbb{R}^2$ as $i \in \mathbb{N}$ by

$$\Lambda_1 := \mathbb{R} \times \mathbb{R}_+ \text{ and } \Lambda_i := \left\{ (x, y) \in \mathbb{R}^2 \middle|\, y \le \frac{x}{i} \right\} \text{ for } i = 2, 3, \ldots$$

as depicted in Fig. 2.5. Observe that for any number $\nu > 0$, we have

$$\Big(\Lambda_1 + (0, \nu)\Big) \bigcap \Big(\bigcap_{i=2}^{\infty} \Lambda_i \Big) = \emptyset,$$

showing that the system $\{\Lambda_i\}_{i \in \mathbb{N}}$ is *extremal at the origin*. On the other hand,

$$\bigcap_{i=1}^{\infty} \Lambda_i = \mathbb{R} \times \{0\},$$

i.e., the nonoverlapping property (2.14) is violated. Furthermore, we can easily calculate the corresponding normal cones by

$$N(0; \Lambda_1) = \big\{ \lambda(0, -1) \big|\, \lambda \ge 0 \big\} \text{ and } N(0; \Lambda_i) = \big\{ \lambda(-1, i) \big|\, \lambda \ge 0 \big\}, \quad i = 2, 3, \ldots.$$

Taking now any $v_i \in N(0; \Lambda_i)$ as $i \in \mathbb{N}$, observe the equivalence

$$\Big[\sum_{i=1}^{\infty} \frac{1}{2^i} v_i = 0 \Big] \Longleftrightarrow \Big[\frac{\lambda_1}{2} \Big(0, -1\Big) + \sum_{i=2}^{\infty} \frac{\lambda_i}{2^i} \Big(-1, i\Big) = 0 \text{ with } \lambda_i \ge 0 \text{ as } i \in \mathbb{N} \Big].$$

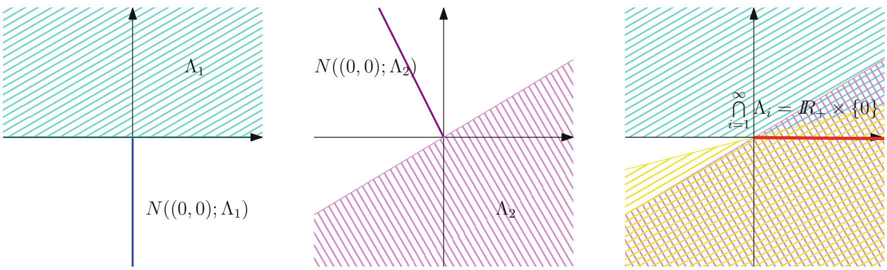


Fig. 2.5 Failure of CEP without nonoverlapping

This implies that $\lambda_i = 0$ and hence $v_i = 0$ for all $i \in \mathbb{N}$. Thus the nontriviality condition in (2.13) is not satisfied, which shows that the conic extremal principle fails for the extremal countable system of cones.

(ii) Next we demonstrate that the extremality conditions of Theorem 2.9 are violated if the sets $\Lambda_i \subset \mathbb{R}^2$ are *convex* with the nonoverlapping property, while some of them are *not cones*. Indeed, consider a countable system of closed and convex sets in $\mathbb{R}^2$ defined by

$$\Lambda_1 := \big\{ (x, y) \in \mathbb{R}^2 \big|\, y \ge x^2 \big\} \text{ and } \Lambda_i := \left\{ (x, y) \in \mathbb{R}^2 \middle|\, y \le \frac{x}{i} \right\} \text{ for } i = 2, 3, \ldots$$

as shown in Fig. 2.6, and observe that only the set Λ_1 is not a cone and that the nonoverlapping property (2.14) is satisfied. Furthermore, the system $\{\Lambda_i\}_{i\in\mathbb{N}}$ is *extremal at the origin* in the sense that (2.12) holds. However, the arguments similar to part (i) of this example show that the extremality conditions (2.13) with $v_i \in N(0;\Omega_i)$ as $i \in \mathbb{N}$ fail to fulfill.

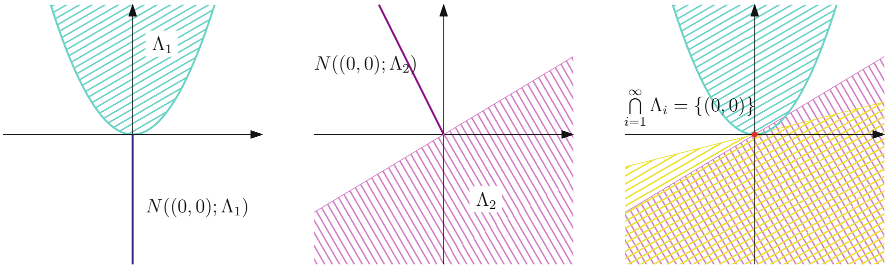


Fig. 2.6 Failure of CEP without conic structure

Our next result is the following *contingent extremal principle* for contingent extremal systems of sets from Definition 2.6(b) justifying the validity of both *contingent* and *normal extremality conditions* from Definition 2.8(b,c) for contingent locally extremal points of such systems.

Theorem 2.11 (Contingent Extremal Principle for Countable Systems of Sets). *Let $\bar{x} \in \bigcap_{i=1}^{\infty} \Omega_i$ be a contingent locally extremal point of a countable system of sets $\{\Omega_i\}_{i\in\mathbb{N}}$ in $\mathbb{R}^n$. Assume that the contingent cones $T(\bar{x};\Omega_i)$ to the sets Ω_i at $\bar{x}$ don't overlap*

$$\bigcap_{i=1}^{\infty} \Big\{T(\bar{x};\Omega_i)\Big\} = \{0\}.$$

Then there are vectors $v_i \in \mathbb{R}^n$ for $i \in \mathbb{N}$ satisfying simultaneously the contingent extremality conditions from Definition 2.8(b) *and the normal extremality conditions from Definition* 2.8(c).

Proof. The existence of $v_i \in N(0;\Lambda_i)$ with $\Lambda_i = T(\bar{x};\Omega_i)$, $i \in \mathbb{N}$, satisfying the extremality conditions (2.13) under the assumed nonoverlapping property of $\{T(\bar{x};\Omega_i)\}_{i\in\mathbb{N}}$ follows directly from Definition 2.6(b) of contingent locally extremal points and the conic extremal principle of Theorem 2.9. To derive from this the claimed normal extremality condition, it suffices to show that for any set $\Omega \subset \mathbb{R}^n$ locally closed around $\bar{x} \in \Omega$, we have the inclusion

$$N(0;\Lambda) \subset N(\bar{x};\Omega) \ \text{ with } \ \Lambda := T(\bar{x};\Omega). \tag{2.19}$$

To verify this inclusion, pick any $v \in N(0;\Lambda)$, and by Definition 1.8 of the contingent cone $T(\bar{x};\Omega)$, find sequences $t_k \downarrow 0$ and $v_k \in N(w_k;T_k)$ with $T_k := (\Omega-\bar{x})/t_k$ such that $w_k \to 0$ and $v_k \to v$ as $k \to \infty$. We have $N(w_k;T_k) = N(x_k;\Omega)$ for

$x_k := \bar{x} + t_k w_k \to \bar{x}$ and conclude therefore by the robustness property from Proposition 1.3 that $v \in N(\bar{x}; \Omega)$. This justifies (2.19) and thus completes the proof of the theorem. $\triangle$

2.3 Variational Principles for Functions

In this short section, we discuss some results known as variational principles for lower semicontinuous functions. They play a crucial role in infinite-dimensional variational analysis, where they are strongly connected with appropriate versions of the extremal principle for systems of two sets; see [522, Chapter 2]. In finite-dimensional spaces, variational principles are rather elementary (they are in fact consequences of the classical Weierstrass existence theorem for l.s.c. functions and finite-dimensional geometry), while even in this case, they provide useful conclusions convenient for applications.

Following the conventional terminology of variational analysis, by *variational principles*, we understand a group of results stating that for any l.s.c. and bounded from below function $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ and any given point x_0 close enough to its minimum, there is an arbitrarily small perturbation $\theta(\cdot)$ such that the resulting function $\varphi + \theta$ achieves its minimum at some point $\bar{x}$ near x_0. In the rest of this section, we assume unless otherwise stated that $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ is a proper *l.s.c.* and *bounded from below* extended-real-valued function while postponing infinite-dimensional discussions till Sections 2.5 and 2.6.

2.3.1 General Variational Principle

The following result presents a general variational principle in the finite-dimensional setting under consideration.

Theorem 2.12 (General Variational Principle in Finite Dimensions). *Let $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ be under the standing assumptions made, and let $\theta\colon \mathbb{R}^n \to \mathbb{R}_+$ be l.s.c. satisfying the growth condition $\theta(x) \to \infty$ as $\|x\| \to \infty$. Then for any $\varepsilon, \lambda > 0$ and any $x_0 \in \mathbb{R}^n$ with $\varphi(x_0) \le \inf \varphi + \varepsilon$, there is $\bar{x} \in \mathbb{R}^n$ such that*

$$\varphi(\bar{x}) \le \varphi(x) + \big(\varepsilon/\lambda\big)\big[\theta(x - x_0) - \theta(\bar{x} - x_0)\big] \text{ for all } x \in \mathbb{R}^n. \tag{2.20}$$

Furthermore, in the case of $\theta(0) = 0$, we have the estimates

$$\varphi(\bar{x}) \le \varphi(x_0) \text{ and } \theta(\bar{x} - x_0) \le \lambda.$$

If in addition the function θ is subadditive on $\mathbb{R}^n$, i.e., $\theta(x + z) \le \theta(x) + \theta(z)$ for all $x, z \in \mathbb{R}^n$, then it follows from (2.20) that

$$\varphi(\bar{x}) \le \varphi(x) + \big(\varepsilon/\lambda\big)\theta(x - \bar{x}) \text{ for all } x \in \mathbb{R}^n, \tag{2.21}$$

where the inequality is strict for all $x \ne \bar{x}$ if $x = 0$ is the only root of $\theta(x) = 0$.

Proof. Consider the unconstrained optimization problem:

$$\text{minimize } \vartheta(x) := \varphi(x) + \big(\varepsilon/\lambda\big)\theta(x - x_0) \text{ over } x \in \mathbb{R}^n. \tag{2.22}$$

Since φ is bounded from below and θ satisfies the imposed growth condition , the level sets $\{x \in \mathbb{R}^n |\ \vartheta(x) \le \gamma\}$ of ϑ are bounded and thus compact in $\mathbb{R}^n$ due to the lower semicontinuity of the function ϑ in (2.22). Then the classical Weierstrass theorem ensures the existence of an optimal solution $\bar{x}$ to (2.22), which verifies (2.20). When $\theta(0) = 0$, we directly deduce from (2.20) that $\varphi(\bar{x}) \le \varphi(x_0)$ by putting $x = x_0$ therein and that $\theta(\bar{x} - x_0) \le \lambda$. If furthermore θ is subadditive, then (2.20) yields (2.21) by applying the former to the representation $x - x_0 = (x - \bar{x}) + (\bar{x} - x_0)$. The last statement of the theorem obviously follows from (2.21). △

Loosely speaking, the result of Theorem 2.12 tells us that for any ε optimal (or *suboptimal*) starting point x_0 in the problem of minimizing the function φ, there exists another ε optimal vector $\bar{x}$ arbitrarily close to x_0 by modulus θ such that $\bar{x}$ is an *exact* solution for the *perturbed* optimization problem in (2.20). Specifying θ in Theorem 2.12 gives us various versions of the variational conditions therein. In particular, for $\theta(x) := \|x\|$, we arrive at the following conditions of *Ekeland's variational principle*, which has a great many consequences and applications in both finite and infinite dimensions; see below.

Corollary 2.13 (Ekeland's Variational Principle). *Let φ, x_0, and ε be given as in Theorem 2.12. Then for every $\lambda > 0$, there is $\bar{x} \in \mathbb{R}^n$ such that $\|\bar{x} - x_0\| \le \lambda$, $\varphi(\bar{x}) \le \varphi(x_0)$, and*

$$\varphi(\bar{x}) < \varphi(x) + \big(\varepsilon/\lambda\big)\|x - \bar{x}\| \text{ whenever } x \ne \bar{x}. \tag{2.23}$$

Observe that the suboptimal solution $\bar{x}$ in Corollary 2.13 satisfies the following *almost stationary condition*

$$\|\nabla\varphi(\bar{x})\| \le \varepsilon/\lambda \tag{2.24}$$

provided that φ is differentiable at $\bar{x}$. Indeed, it follows by applying the elementary sum rule from Proposition 1.30(ii) to the inclusion $0 \in \widehat{\partial}(\varphi + \theta)(\bar{x})$ from Proposition 1.30(i) with $\theta(x) := (\varepsilon/\lambda)\|x - \bar{x}\|$ due to the optimality of $\bar{x}$ for this sum and the fact that $\partial(\|\cdot - \bar{x}\|)(\bar{x}) = \mathbb{B}$ in convex analysis. We'll see below that the flexibility of choosing an auxiliary function θ in Theorem 2.12, not just as the norm $\|\cdot\|$, allows us to gain more information for applications.

2.3.2 Applications to Suboptimality and Fixed Points

Note that the almost stationary condition (2.24) and its verification based on Corollary 2.13 unavoidably require the differentiability of φ via the application of Proposition 1.30(ii). However, we can derive in this way some extended conditions for suboptimal points of smooth and nonsmooth functions φ by an appropriate choice of perturbations θ in the general variational principle of Theorem 2.12. The next theorem contains two independent versions of *subdifferential* almost stationary conditions obtained in this way for suboptimal solutions. The first one is expressed in terms of (*lower*) regular subgradients from $\widehat{\partial}\varphi(\bar{x})$, while the other condition is given in a new enhanced form via the *entire set* of *upper* regular subgradients from $\widehat{\partial}^+\varphi(\bar{x}) := -\widehat{\partial}(-\varphi)(\bar{x})$ provided that this set is nonempty. The proof of the latter result invokes the smooth variational description of regular subgradients from Theorem 1.27.

Theorem 2.14 (Subdifferential Almost Stationary Conditions for Suboptimal Solutions). *Let φ, ε, λ, and x_0 be as in Theorem 2.12. Then there exist a suboptimal solution $\bar{x} \in \mathbb{R}^n$ and a regular subgradient $v \in \widehat{\partial}\varphi(\bar{x})$ such that $\|\bar{x} - x_0\| \le \lambda$, $\varphi(\bar{x}) \le \varphi(x_0)$, and $\|v\| \le \varepsilon/\lambda$. If furthermore $\widehat{\partial}^+\varphi(\bar{x}) \neq \emptyset$, then in addition the latter estimate holds for any $v \in \widehat{\partial}^+\varphi(\bar{x})$.*

Proof. Taking ε, λ, and x_0 from Theorem 2.12, we select

$$\theta(x) := \frac{1}{4\lambda}\|x\|^2, \quad x \in \mathbb{R}^n,$$

and find a vector $\bar{x} \in B_{2\lambda}(x_0)$ with $\varphi(\bar{x}) \le \varphi(x_0)$, minimizing the function

$$\vartheta(x) := \varphi(x) + \frac{\varepsilon}{4\lambda^2}\|x - x_0\|^2 \text{ over } x \in \mathbb{R}^n.$$

Applying now both parts of Proposition 1.30 to the sum $\vartheta(\cdot)$ shows that

$$0 \in \widehat{\partial}\varphi(\bar{x}) + \frac{\varepsilon}{2\lambda^2}\Big(\bar{x} - x_0\Big),$$

which justifies, by taking into account the estimate $\theta(\bar{x} - x_0) \le \lambda$ for the selected function $\theta(\cdot)$, the first stationary condition of the theorem.

To verify the second statement of the theorem under the assumption that $\widehat{\partial}^+\varphi(\bar{x}) \neq \emptyset$, we proceed as follows. Employing Corollary 2.13 gives us a vector $\bar{x} \in \mathbb{R}^n$ satisfying (2.23). Pick now any $v \in -\widehat{\partial}(-\varphi)(\bar{x})$, and apply to it the first smooth variational description in Theorem 1.27. This allows us to find a function ψ defined on a neighborhood of $\bar{x}$ such that ψ is Fréchet differentiable at $\bar{x}$ and obeys the conditions

$$\psi(\bar{x}) = \varphi(\bar{x}), \quad \nabla\psi(\bar{x}) = v, \text{ and } \psi(x) \ge \varphi(x) \text{ for all } x \in \mathbb{R}^n.$$

Combining it with (2.23) shows that the function $\phi(x) := \psi(x) + (\varepsilon/\lambda)\|x - \bar{x}\|$ attains a local minimum at $\bar{x}$. Then it follows from Proposition 1.30(i,ii) that

$$0 \in \widehat{\partial}\phi(\bar{x}) = \nabla\psi(\bar{x}) + \widehat{\partial}\Big(\frac{\varepsilon}{\lambda}\|\cdot - \bar{x}\|\Big)(\bar{x}) \subset v + \frac{\varepsilon}{\lambda}\mathbb{B},$$

which verifies that $\|v\| \le \varepsilon/\lambda$ and completes the proof of the theorem. $\triangle$

Observe that for functions φ differentiable at $\bar{x}$, both subdifferential stationary conditions in Theorem 2.14 reduce to (2.24).

Finally in this section, we show that the general variational principle of Theorem 2.12 implies the following *fixed point* result for set-valued mapping without standard continuity and contraction assumptions.

Proposition 2.15 (Fixed Points). *Let $F\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be a set-valued mapping with nonempty values, and let the functions φ, θ satisfy all the conditions of Theorem 2.12. Assume in addition that*

$$\text{for all } x \in \mathbb{R}^n \text{ there is } y \in F(x) \text{ with } \theta(y-x) \le \varphi(x) - \varphi(y). \tag{2.25}$$

Then there are points $\bar{x} \in \mathbb{R}^n$ and $\bar{y} \in F(\bar{x})$ such that $\theta(\bar{y}-\bar{x}) = 0$, which implies that F admits a fixed point $\bar{x} \in F(\bar{x})$ provided that $x = 0$ is the only root of the equation $\theta(x) = 0$. Furthermore, the validity of condition (2.25) for all $y \in F(x)$ ensures that $F(\bar{x}) = \{\bar{x}\}$.

Proof. Taking $\lambda = 2\varepsilon$, we get from (2.21) of Theorem 2.12 that

$$\theta(y-\bar{x}) \ge 2\big(\varphi(\bar{x}) - \varphi(y)\big) \text{ for all } y \in F(x).$$

This implies by assumption (2.25) with $x = \bar{x}$ that $\theta(\bar{x}-\bar{y}) = 0$ for some point $\bar{y} \in F(\bar{x})$, and hence we arrive at the fixed point statement of the corollary. Moreover, the fulfillment of (2.25) for any $y \in F(x)$ tells us that $\bar{y} = \bar{x}$ whenever $\bar{y} \in F(\bar{x})$ and thus completes the proof of the corollary. △

2.4 Basic Intersection Rule and Some Consequences

In this section, we first employ the *extremal principle* for systems of two closed sets in Theorem 2.3 to establish the fundamental *intersection rule* for limiting normals that plays the underlying role in deriving other calculus rules of generalized differentiation and their applications. Some of its direct consequences for normals and subgradients needed in what follows are also presented here.

2.4.1 Normals to Set Intersections and Additions

The following theorem on representing the normal cone to intersections of two closed sets is crucial for all the major results of generalized differential calculus involving the nonconvex robust constructions of Chapter 1.

Theorem 2.16 (Basic Intersection Rule). *Let $\Omega_1, \Omega_2 \subset \mathbb{R}^n$ be such that $\bar{x} \in \Omega_1 \cap \Omega_2$, and let the* NORMAL QUALIFICATION CONDITION

$$N(\bar{x};\Omega_1) \cap \big(-N(\bar{x};\Omega_2)\big) = \{0\} \tag{2.26}$$

be satisfied. Then we have the inclusion

$$N(\bar{x};\Omega_1\cap\Omega_2)\subset N(\bar{x};\Omega_1)+N(\bar{x};\Omega_2). \tag{2.27}$$

Furthermore, (2.27) *holds as equality, and the set* $\Omega_1\cap\Omega_2$ *is normally regular at* $\bar{x}$ *provided that both sets* Ω_1 *and* Ω_2 *are normally regular at this point.*

Proof. To verify (2.27), pick any $v\in N(\bar{x};\Omega_1\cap\Omega_2)$, and by the first representation of Theorem 1.6, find sequences $x_k\to\bar{x}$ and $v_k\to v$ such that

$$x_k\in\Omega_1\cap\Omega_2 \ \text{ and } \ v_k\in\widehat{N}(x_k;\Omega_1\cap\Omega_2) \ \text{ for all } \ k\in\mathbb{N}.$$

Select an arbitrary sequence of $\varepsilon_k\downarrow 0$ as $k\to\infty$, and for any fixed $k\in\mathbb{N}$, define two closed sets in $\mathbb{R}^{n+1}$ by

$$\begin{aligned}&\Lambda_1:=\Omega_1\times\mathbb{R}_+\quad\text{and}\\&\Lambda_{2k}:=\big\{(x,\alpha)\big|\ x\in\Omega_2,\ \langle v_k,x-x_k\rangle-\varepsilon_k\|x-x_k\|\ge\alpha\big\}.\end{aligned} \tag{2.28}$$

By the set construction in (2.28) and definition (1.5) of regular normals, we have that $(x_k,0)\in\Lambda_1\cap\Lambda_{2k}$ and that

$$\Lambda_1\cap\big(\Lambda_{2k}-(0,\nu)\big)\cap(U\times\mathbb{R})=\emptyset \ \text{ for all } \ \nu>0,$$

where U is a suitable neighborhood of x_k. This means that $(x_k,0)$ is a *locally extremal point* of the set system $\{\Lambda_1,\Lambda_{2k}\}$. Applying the extremal principle from Theorem 2.3 to this system at $(x_k,0)$ for each $k\in\mathbb{N}$ gives us pairs $(u_k,\lambda_k)\in\mathbb{R}^n\times\mathbb{R}$ with $\|(u_k,\lambda_k)\|=1$ satisfying the inclusions

$$(u_k,\lambda_k)\in N\big((x_k,0);\Lambda_1\big) \ \text{ and } \ (-u_k,-\lambda_k)\in N\big((x_k,0);\Lambda_{2k}\big). \tag{2.29}$$

By the compactness of the unit sphere in $\mathbb{R}^{n+1}$, we get without loss of generality that $(u_k,\lambda_k)\to(u,\lambda)$ as $k\to\infty$ for some pair $(u,\lambda)\in\mathbb{R}^n\times\mathbb{R}$ with $\|(u,\lambda)\|=1$. The robustness property of basic normals from Proposition 1.3 ensures by the first inclusion in (2.29) that $(u,\lambda)\in N((\bar{x},0);\Omega_1\times\mathbb{R}_+)$, which implies in turn by Proposition 1.4 that

$$u\in N(\bar{x};\Omega_1) \ \text{ and } \ \lambda\le 0. \tag{2.30}$$

Furthermore, using the structure of Λ_{2k} in (2.28) and both representations of basic normals in (1.7) allows us to conclude that

$$(-\lambda v-u,\lambda)\in N\big((\bar{x},0);\Omega_2\times\mathbb{R}_+\big). \tag{2.31}$$

To show next that $\lambda<0$, suppose on the contrary that $\lambda=0$, which implies by (2.30) and (2.31) that $0\ne u\in N(\bar{x};\Omega_1)\cap(-N(\bar{x};\Omega_2))$. This is impossible by the assumed qualification condition (2.26). Thus we can take $\lambda=-1$ and get from (2.31) that $w:=v-u\in N(\bar{x};\Omega_2)$, which verifies that the selected vector v belongs to the right-hand side of (2.27).

To prove the last statement of the theorem, observe first that the inclusion

$$\widehat{N}(\bar{x};\Omega_1\cap\Omega_2)\supset\widehat{N}(\bar{x};\Omega_1)+\widehat{N}(\bar{x};\Omega_2)$$

is always satisfied. Assuming now that both sets Ω_1 and Ω_2 are normally regular at $\bar{x}$ in the sense of (1.55), we get

$$N(\bar{x};\Omega_1)+N(\bar{x};\Omega_2)=\widehat{N}(\bar{x};\Omega_1)+\widehat{N}(\bar{x};\Omega_2)\subset\widehat{N}(\bar{x};\Omega_1\cap\Omega_2)\subset N(\bar{x};\Omega_1\cap\Omega_2).$$

which verifies the opposite inclusion in (2.27) and so completes the proof. △

As we'll see below, the obtained intersection rule is the key result of generalized differential calculus in variational analysis. Let us now present some of its rather straightforward consequences. The first one is an extension of the intersection rule to finitely many sets.

Corollary 2.17 (Normals to Finite Set Intersections). *Let $\Omega_1,\ldots,\Omega_s$ with $s\geq 2$ be subsets of $\mathbb{R}^n$ such that $\bar{x}\in\cap_{i=1}^s\Omega_i$, and let the system*

$$v_i\in N(\bar{x};\Omega_i),\ i=1,\ldots,s,\quad v_1+\ldots+v_s=0$$

has only the trivial solution $v_1=\ldots=v_s=0$. Then we have the inclusion

$$N\Big(\bar{x};\bigcap_{i=1}^s\Omega_i\Big)\subset N(\bar{x};\Omega_1)+\ldots+N(\bar{x};\Omega_s),\tag{2.32}$$

which holds as equality if all the sets Ω_i are normally regular at $\bar{x}$. In this case the intersection $\cap_{i=1}^s\Omega_i$ is also normally regular at $\bar{x}$.

Proof. Arguing by induction and having in hands the intersection rule for two sets, suppose now that this rule holds for $s-1$ sets, and then represent the intersection $\Omega=\Omega_1\cap\ldots\cap\Omega_s$ of $s>2$ sets as $\Omega=\Lambda_1\cap\Lambda_2$ with $\Lambda_1:=\cap_{i=1}^{s-1}\Omega_i$ and $\Lambda_2:=\Omega_s$. It is easy to check that the qualification condition imposed on $\{\Omega_1,\ldots,\Omega_s\}$ yields the validity of (2.26) for $\{\Lambda_1,\Lambda_2\}$. Thus applying Theorem 2.16 to the two-set intersection $\Lambda_1\cap\Lambda_2$ and using the induction assumption justify inclusion (2.32). We also get in this way the regularity and equality statements when all the sets Ω_i are normally regular at $\bar{x}$. △

The next consequence of Theorem 2.16 provides a useful *sum rule* for sets, which holds without imposing any qualification condition.

Corollary 2.18 (Normals to Sums of Sets). *Let $\Omega_1,\Omega_2\subset\mathbb{R}^n$, and let $\bar{x}\in\Omega_1+\Omega_2$. Assume that the set-valued mapping $S\colon\mathbb{R}^n\rightrightarrows\mathbb{R}^{2n}$ defined by*

$$S(x):=\Big\{(x_1,x_2)\in\mathbb{R}^{2n}\Big|\ x_1+x_2=x,\ x_1\in\Omega_1,\ x_2\in\Omega_2\Big\},\quad x\in\mathbb{R}^n,$$

is locally bounded around $\bar{x}$. Then we have the inclusion

$$N(\bar{x};\Omega_1+\Omega_2)\subset\bigcup_{(x_1,x_2)\in S(\bar{x})}N(x_1;\Omega_1)\cap N(x_2;\Omega_2).$$

Proof. Observe first that the closedness of Ω_1 and Ω_2 and the uniform boundedness of $S(x)$ around $\bar{x}$ ensure that the set $\Omega_1 + \Omega_2$ is locally closed around $\bar{x}$. Pick any $v \in N(\bar{x}; \Omega_1+\Omega_2)$, and by Theorem 1.6, find sequences $x_k \to \bar{x}$ with $x_k \in \Omega_1+\Omega_2$ and $v_k \to v$ such that $v_k \in \widehat{N}(x_k; \Omega_1 + \Omega_2)$. Considering the sets $\Lambda_1 := \Omega_1 \times \mathbb{R}^n$ and $\Lambda_2 := \mathbb{R}^n \times \Omega_2$, it is not hard to check that

$$(v_k, v_k) \in \widehat{N}\big((x_{1k}, x_{2k}); \Lambda_1 \cap \Lambda_2\big) \text{ whenever } (x_{1k}, x_{2k}) \in S(x_k) \tag{2.33}$$

for all $k \in \mathbb{N}$. Taking such a sequence of (x_{1k}, x_{2k}) and employing again the uniform boundedness of $S(x)$ around $\bar{x}$ give us some $(\bar{x}_1, \bar{x}_2) \in S(\bar{x})$ such that $(x_{1k}, x_{2k}) \to (\bar{x}_1, \bar{x}_2)$ along a subsequence. By passing to the limit in (2.33) as $k \to \infty$, we get vectors $u_1, u_2 \in \mathbb{R}^n$ with

$$(u_1, 0) \in N\big((\bar{x}_1, \bar{x}_2); \Lambda_1\big),\ (0, u_2) \in N\big((\bar{x}_1, \bar{x}_2); \Lambda_2\big),\ (v, v) = (u_1, 0) + (0, u_2),$$

which implies that $u_1 \in N(\bar{x}_1; \Omega_1)$, $u_2 \in N(\bar{x}_2; \Omega_1)$, and $u_1 = u_2 = v$. This verifies that $v \in N(\bar{x}_1; \Omega_1) \cap N(\bar{x}_2; \Omega_2)$ and thus completes the proof. $\triangle$

2.4.2 Subdifferential Sum Rules

Now we turn to *subgradients* of extended-real-valued l.s.c. functions and deduce directly from Theorem 2.16 the following *subdifferential sum rules* for both basic and singular subgradients in Definition 1.18. This theorem plays the underlying role in subdifferential calculus (see Section 4.1) as well as in deriving other results and various applications presented in the book.

Theorem 2.19 (Subdifferential Sum Rules for Two l.s.c. Functions). *Let $\varphi_1, \varphi_2\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ be such that $\bar{x} \in \operatorname{dom}\varphi_i$ for $i = 1, 2$, and let the (singular)* SUBDIFFERENTIAL QUALIFICATION CONDITION

$$\partial^\infty\varphi_1(\bar{x}) \cap \big(-\partial^\infty\varphi_2(\bar{x})\big) = \{0\} \tag{2.34}$$

be satisfied. Then we have the sum rule inclusions

$$\partial(\varphi_1 + \varphi_2)(\bar{x}) \subset \partial\varphi_1(\bar{x}) + \partial\varphi_2(\bar{x}), \tag{2.35}$$

$$\partial^\infty(\varphi_1 + \varphi_2)(\bar{x}) \subset \partial^\infty\varphi_1(\bar{x}) + \partial^\infty\varphi_2(\bar{x}). \tag{2.36}$$

If furthermore both functions φ_1, φ_2 are lower regular at $\bar{x}$, then the sum $\varphi_1 + \varphi_2$ also has this property and (2.35) holds as equality. The equality holds also in (2.36), and the function $\varphi_1 + \varphi_2$ is epigraphically regular at $\bar{x}$ if both functions φ_1, φ_2 are epigraphically regular at this point.

Proof. Let us derive inclusions (2.35) and (2.36) for basic and singular subgradients simultaneously by reducing both of them to Theorem 2.16. Taking v from either $\partial(\varphi_1+\varphi_2)(\bar{x})$ or $\partial^\infty(\varphi_1+\varphi_2)(\bar{x})$, we get by Definition 1.18 that

$$(v,-\lambda)\in N\big((\bar{x},(\varphi_1+\varphi_2)(\bar{x}));\mathrm{epi}\,(\varphi_1+\varphi_2)\big)\ \text{ with either }\ \lambda=1\ \text{ or }\ \lambda=0,$$

respectively. Denote $\bar{\alpha}_i:=\varphi_i(\bar{x})$ for $i=1,2$, and consider the sets

$$\Omega_i:=\big\{(x,\alpha_1,\alpha_2)\in\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}\big|\ \alpha_i\ge\varphi_i(x)\big\},\quad i=1,2.$$

It is easy to observe that $(v,-\lambda,-\lambda)\in N((\bar{x},\bar{\alpha}_1,\bar{\alpha}_2);\Omega_1\cap\Omega_2)$. Applying now the intersection rule of Theorem 2.16 to this set intersection with taking into account that the subdifferential qualification condition (2.34) ensures the validity of the normal one (2.26) for the sets Ω_i constructed above gives us pairs $(v_i,-\lambda_i)\in N((\bar{x},\bar{\alpha}_i);\mathrm{epi}\,\varphi_i)$ for $i=1,2$ such that

$$(v,-\lambda,-\lambda)=(v_1,-\lambda_1,0)+(v_2,0,-\lambda_2).$$

Thus we get $v=v_1+v_2$ with either $v_i\in\partial\varphi_i(\bar{x})$ or $v_i\in\partial^\infty\varphi_i(\bar{x})$ as $i=1,2$ depending on the choice of $\lambda=0,1$ in the arguments above. This verifies the sum rule inclusions in (2.35) and (2.36).

If both φ are lower regular at $\bar{x}$, in the sense of $\partial\varphi_i(\bar{x})=\widehat{\partial}\varphi_i(\bar{x})$ for $i=1,2$ (see Exercise 1.74), then the equality and regularity statements of the theorem follow from (2.35), and the inclusion

$$\widehat{\partial}(\varphi_1+\varphi_2)(\bar{x})\supset\widehat{\partial}\varphi_1(\bar{x})+\widehat{\partial}\varphi_2(\bar{x})$$

the validity of which for arbitrary functions φ_i can be immediately deduced from definition (1.33). The last statement of the theorem is verified similarly by using the second representation of Exercise 1.74(ii). △

We conclude this section with the following two direct corollaries of Theorem 2.19. The first one concerns *semi-Lipschitzian sum* $\mathcal{SL}(\bar{x})$, i.e., sums of two functions $\varphi_1+\varphi_2$ one of which is l.s.c. around $\bar{x}$, while the other is locally Lipschitzian around this point.

Corollary 2.20 (Semi-Lipschitzian Sum Rule for Basic Subgradients). *Let* $(\varphi_1,\varphi_2)\in\mathcal{SL}(\bar{x})$. *Then we have the basic subgradient inclusion* (2.35).

Proof. Follows from Theorem 2.19 due to $\partial^\infty\varphi(\bar{x})=\{0\}$ for locally Lipschitzian functions, which ensures the validity of (2.34) by Theorem 1.22. △

Note that for any pair $(\varphi_1,\varphi_2)\in\mathcal{SL}(\bar{x})$, the singular subdifferential inclusion (2.36) always holds as equality. This has been justified by the direct proof of Proposition 1.29 but can also be deduced from inclusion (2.36) by applying it to the sum $\varphi_2=(\varphi_1+\varphi_2)+(-\varphi_1)$ and using the characterization of the local Lipschitz continuity from Theorem 1.22.

The next corollary is an extension of Theorem 2.19 to finite sums.

Corollary 2.21 (Subgradients for Sums of Finitely Many l.s.c. Functions). *Let* $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ *for* $i = 1, \ldots, s$ *be such that* $\bar{x} \in \cap_{i=1}^{s} \mathrm{dom}\,\varphi_i$, *and let the following qualification condition be satisfied:*

$$\big[v_i \in \partial^\infty \varphi_i(\bar{x}),\ i = 1, \ldots, s \big|\ v_1 + \ldots + v_s = 0\big] \Longrightarrow v_1 = \ldots = v_s = 0, \quad (2.37)$$

which is surely the case where all but one of φ_i *are locally Lipschitzian around* $\bar{x}$. *Then we have the subdifferential sum rules*

$$\partial\big(\varphi_1 + \ldots + \varphi_s\big)(\bar{x}) \subset \partial\varphi_1(\bar{x}) + \ldots + \partial\varphi_s(\bar{x}), \quad (2.38)$$

$$\partial^\infty\big(\varphi_1 + \ldots + \varphi_s\big)(\bar{x}) \subset \partial^\infty\varphi_1(\bar{x}) + \ldots + \partial^\infty\varphi_s(\bar{x}), \quad (2.39)$$

where (2.38) *holds as equality if all* φ_i *are lower regular at* $\bar{x}$. *In this case the sum* $\varphi_1 + \ldots + \varphi_s$ *is lower regular at* $\bar{x}$ *as well. The equality also holds in* (2.39), *and the sum* $\varphi_1 + \ldots + \varphi_s$ *is epigraphically regular at* $\bar{x}$ *if all the functions* φ_i *are epigraphically regular at this point.*

Proof. From the case of $s = 2$ in Theorem 2.19, we can justify the general case of $s > 2$ by induction, where the qualification condition (2.37) at the current step is verified by using (2.39) at the previous step of induction. △

In the subsequent parts of the book (see, in particular, Sections 3.2 and 4.1 together with the exercises and commentary sections), we'll employ the basic intersection rule of Theorem 2.16 and its subdifferential consequences from Theorem 2.19 to deriving a number of calculus rules for coderivatives and subgradients of various compositions. To deal efficiently with set-valued and single-valued mappings, we study in the next chapter some fundamental *well-posedness* properties, which are of their own interest for numerous aspects of variational analysis and optimization while being used therein for developing and verifying a variety of results on generalized differential calculus.

2.5 Exercises for Chapter 2

Exercise 2.22 (Convex Separation for Finitely Many Sets). Deduce from the extremal principle of Theorem 2.3 the convex separation theorem for $s \geq 2$ sets in $\mathbb{R}^n$ under the relative interiority condition (2.10) in $\mathbb{R}^n$.

Exercise 2.23 (Interiors of Sets in Extremal Systems). Let $\Omega_1, \ldots, \Omega_s$ be subsets of a Banach space X such that the first $s-1$ of them has nonempty interiors. Show that if the system $\{\Omega_1, \ldots, \Omega_s, \bar{x}\}$ is locally extremal in the sense of Definition 2.1 considered in Banach spaces, then we have

$$\mathrm{int}\,\Omega_1 \cap \ldots \cap \mathrm{int}\,\Omega_{s-1} \cap \Omega_s \cap U = \emptyset.$$

When does the converse assertion hold?

Exercise 2.24 (Approximate Extremal Principle in Infinite Dimensions).

(i) Prove that the approximate extremal principle from Corollary 2.5 holds in *Fréchet smooth* spaces. *Hint:* Use an appropriate modification of the method of metric approximations, completeness of the space in question, and the Fréchet differentiability of an equivalent norm. Compare with the proof of [522, Theorem 2.10].

(ii) Check that the approximate extremal principle holds in *Asplund* spaces, being in fact a characterization of this class of Banach spaces. *Hint:* Use the method of separable reduction to reduce the Asplund space setting to the Fréchet smooth one in accordance with the proof of [522, Theorem 2.20].

Exercise 2.25 (Density Results). Let $\Omega \subset X$ be a proper (and closed) subset of an Asplund space X. Show that the approximate extremal principle in X yields the validity of the following statements:

(i) *Nonlinear Bishop-Phelps theorem*: the set

$$\Big\{x \in \operatorname{bd}\Omega \Big|\ \widehat{N}(x;\Omega) \neq \{0\}\Big\}$$

is dense on the boundary $\operatorname{bd}\Omega$ for any such Ω. *Hint:* Given any $\bar{x} \in \operatorname{bd}\Omega$, apply the approximate extremal principle from Exercise 2.24(ii) to the extremal system $\{\Omega, \{\bar{x}\}, \bar{x}\}$, and compare it with [522, Corollary 2.21] for this and other boundary characterizations of Asplund spaces.

(ii) Verify that for convex sets Ω, the density result from (i) reduces to the classical Bishop-Phelps theorem on the density of *support points* on the boundary of Ω (see, e.g., [638, Theorem 3.18]) while in the case of an Asplund space X.

(iii) *Density of regular subgradients*: the set

$$\{(x,\varphi(x)) \in X \times \mathbb{R} \Big|\ \widehat{\partial}\varphi(x) \neq \emptyset\Big\}$$

is dense on the graph of φ for every l.s.c. function $\varphi\colon X \to \overline{\mathbb{R}}$. *Hint:* Derive this from the approximate extremal principle, and compare with [522, Corollary 2.29].

Exercise 2.26 (Fuzzy Sum Rule from the Extremal Principle). Let $\varphi_1\colon X \to \mathbb{R}$ be locally Lipschitzian around $\bar{x}$, and let $\varphi_2\colon X \to \overline{\mathbb{R}}$ be l.s.c. around this point. Show that for any $\varepsilon > 0$, the following "fuzzy" sum rule holds:

$$\widehat{\partial}(\varphi_1+\varphi_2)(\bar{x}) \subset \bigcup_{x_i \in U(\varphi_i,\bar{x},\varepsilon)} \Big\{\widehat{\partial}\varphi_1(x_1)+\widehat{\partial}\varphi_2(x_2)\Big\} + \varepsilon\mathbb{B}^*, \tag{2.40}$$

where $U(\varphi,\bar{x},\varepsilon) := \{x \in X|\ \|x-\bar{x}\| < \varepsilon,\ |\varphi(x)-\varphi(\bar{x})| < \varepsilon\}$. *Hint:* Assuming without loss of generality that $\bar{x} = 0$ is a local minimizer of $\varphi_1+\varphi_2$ and that $\varphi_1(0) = \varphi_2(0) = 0$, consider the system of sets

$$\Omega_1 := \operatorname{epi}\varphi_1, \quad \Omega_2 := \Big\{(x,\alpha) \in X \times \mathbb{R} \Big|\ \varphi_2(x) \le -\alpha\Big\}.$$

which is locally extremal at $(0,0)$. Apply to it the approximate extremal principle, and compare with [522, Lemma 2.32].

Exercise 2.27 (Weak Fuzzy Sum Rule). Let X be an Asplund space, and let $\varphi_1,\ldots,\varphi_s\colon X \to \overline{\mathbb{R}}$ be l.s.c. functions on X. Prove that for any $\bar{x} \in X$, $\varepsilon > 0$, $x^* \in \widehat{\partial}(\varphi_1+\ldots+\varphi_s)(\bar{x})$, and any weak* neighborhood V^* of $0 \in X^*$ there are $x_i \in \bar{x}+\varepsilon\mathbb{B}$ and $x_i^* \in \widehat{\partial}\varphi_i(x_i)$ such that $|\varphi_i(x_i)-\varphi_i(\bar{x})| \le \varepsilon$ for all $i = 1,\ldots,s$ and

$$x^* \in \sum_{i=1}^{s} x_i^* + V^*.$$

Hint: Use the density subdifferential result from Exercise 2.25(ii) and properties of infinite convolutions; cf. [254, Theorem 2].

Exercise 2.28 (Sequential Normal Compactness of Sets). Let Ω be a subset of a Banach space X, and let $\bar{x} \in \Omega$. We say that Ω is *sequentially normally compact* (SNC) at $\bar{x}$ if for any sequences $(x_k, x_k^*, \varepsilon_k) \subset X \times X^* \times \mathbb{R}_+$ we have

$$\Big[x_k \overset{\Omega}{\to} \bar{x},\ x_k^* \overset{w^*}{\to} 0,\ \varepsilon_k \downarrow 0,\ x_k^* \in \widehat{N}_{\varepsilon_k}(x_k; \Omega)\Big] \Longrightarrow \|x_k^*\| \to 0 \ \text{ as } \ k \to \infty. \tag{2.41}$$

(i) Show that we can equivalently put $\varepsilon_k \equiv 0$ in (2.41) if X is Asplund (and Ω is locally closed around $\bar{x}$ in our standing assumption). *Hint:* Employ Exercise 1.42.

(ii) The *affine hull* of Ω, $\operatorname{aff}\Omega$, is the smallest affine set containing Ω; its closure is denoted by $\overline{\operatorname{aff}}\,\Omega$. The *codimension* $\operatorname{codim}(\overline{\operatorname{aff}}\,\Omega)$ of $\overline{\operatorname{aff}}\,\Omega$ is the dimension of the quotient space $X \setminus (\overline{\operatorname{aff}}\,\Omega - x)$, which is independent of $x \in \overline{\operatorname{aff}}\,\Omega$. Show that the SNC property of Ω at $\bar{x}$ implies that the subspace $\overline{\operatorname{aff}}\,(\Omega \cap U)$ is of *finite codimension* for any neighborhood U of $\bar{x}$. In particular, a singleton in X is SNC if and only if $\dim X < \infty$. *Hint:* Use the fundamental Josefson-Nissenzweig theorem telling us that for any infinite-dimensional Banach space X, there is a sequence of unit vectors $x_k^* \in X^*$ that weak* converges to zero; see [207, Chapter 12].

(iii) The *relative interior* of Ω, $\operatorname{ri}\Omega$, is the interior of Ω with respect to $\overline{\operatorname{aff}}\,\Omega$. Prove that for convex sets Ω with $\operatorname{ri}\Omega \neq \emptyset$, the SNC property of Ω at every $\bar{x} \in \Omega$ is *equivalent* to $\operatorname{codim}(\overline{\operatorname{aff}}\,\Omega) < \infty$. *Hint:* Apply the representation of ε-normals to convex sets from Proposition 1.7 (which holds in any Banach space) to the given set Ω at $\bar{x}$ with respect to the subspace $\overline{\operatorname{aff}}\,\Omega$; see [522, Theorem 1.21].

Exercise 2.29 (Epi-Lipschitzian and Compactly Epi-Lipschitzian Sets). We say that $\Omega \subset X$ is COMPACTLY EPI-LIPSCHITZIAN (CEL) around $\bar{x} \in \Omega$ if there are a compact set $C \subset X$, neighborhoods U of $\bar{x}$, O of $0 \in X$, and $\gamma > 0$ such that

$$\Omega \cap U + tO \subset \Omega + tC \ \text{ for all } \ t \in (0, \gamma). \tag{2.42}$$

The set Ω is said to be EPI-LIPSCHITZIAN around $\bar{x}$ if C in (2.42) can be selected as a singleton. Verify the following statements, where X is an arbitrary Banach space unless otherwise stated:

(i) If the set Ω is CEL around $\bar{x}$, then it is SNC at this point. *Hint:* Compare it with the proof of [522, Theorem 1.26].

(ii) The SNC property is *strictly weaker* than the CEL one in *every* X for which the dual ball $\mathbb{B}^*$ is *not* weak* sequentially compact, in particular in the classical spaces l^∞ and $L^\infty[0,1]$. *Hint:* Find this in [259].

(iii) There is a nonseparable Asplund space X admitting a $\mathcal{C}^\infty$-smooth renorm and a closed *convex cone* $\Omega \subset X$ such that Ω is SNC at the origin but not CEL around $\bar{x} = 0$. *Hint:* Compare it with [259] and [522, Example 3.6].

(iv) A convex set Ω is epi-Lipschitzian around any $\bar{x} \in \Omega$ if and only if $\operatorname{int}\Omega \neq \emptyset$. *Hint:* Compare it with the proof of [522, Proposition 1.25].

Exercise 2.30 (SNC Property for Inverse Images of Sets Under Differentiable Mappings Between Banach Spaces). Let $f\colon X \to Y$ be a between Banach spaces that is strictly differentiable at $\bar{x}$ with the surjective derivative $\nabla f(\bar{x})$, and let Θ be a subset of Y containing $\bar{y} := f(\bar{x})$. Show that the set $f^{-1}(\Theta)$ is SNC at $\bar{x}$ if and only if Θ is SNC at $\bar{y}$. *Hint:* Use the classical open mapping theorem together with the result of Exercise 1.53, and compare it with the proofs of [522, Lemma 1.16 and Theorem 1.22].

Exercise 2.31 (Exact Extremal Principle in Infinite Dimensions).

(i) Use the approximate extremal principle to show that the exact/pointbased extremal principle of Theorem 2.3 holds provided that the dual unit ball $\mathbb{B}^* \subset X^*$ is *sequentially weak* compact* (as in the case of Asplund spaces; see Exercise 1.41(iii)) and that *all but one* sets Ω_i, $i = 1, \ldots, s$, are *SNC* at their locally extremal point $\bar{x}$. *Hint:* Compare it with [522, Theorem 2.22].

(ii) Show that any infinite-dimensional separable Banach space contains an extremal system of two convex compact sets, which are not SNC and for which the relationships of the exact extremal principle fail.

Exercise 2.32 (Nontriviality of Basic Normals and Subgradients from the Extremal Principle). Derive from the exact extremal principle the following statements in any Banach space X: ,

(i) Let $\Omega \subset X$ be proper, closed, and SNC at $\bar{x} \in \operatorname{bd}\Omega$. Then $N(\bar{x};\Omega) \neq \{0\}$.
(ii) Let $\varphi\colon X \to \mathbb{R}$ be locally Lipschitzian around $\bar{x}$. Then $\partial\varphi(\bar{x}) \neq \emptyset$.

Exercise 2.33 (Global Set Extremality and Separation). We say that two nonempty subsets Ω_1, Ω_2 of a locally convex topological vector space X form a (globally) *extremal system* if for any neighborhood V of the origin in X there exists a vector $a \in X$ such that

$$a \in V \quad \text{and} \quad (\Omega_1 + a) \cap \Omega_2 = \emptyset. \tag{2.43}$$

(i) Compare this notion with the local set extremality from Definition 2.1.

(ii) Verify that the sets Ω_1 and Ω_2 form an extremal system in the sense of (2.43) if and only if $0 \notin \operatorname{int}(\Omega_1 - \Omega_2)$. Show furthermore that the extremality of Ω_1, Ω_2 implies that $(\operatorname{int}\Omega_1) \cap \Omega_2 = \emptyset$ while the opposite implication fails.

(iii) Prove that the global extremality of convex sets Ω_1, Ω_2 together with the difference interiority condition $\operatorname{int}(\Omega_1 - \Omega_2) \neq \emptyset$ yields the separation property

$$\sup_{x \in \Omega_1} \langle x^*, x\rangle \le \inf_{x \in \Omega_2} \langle x^*, x\rangle \quad \text{for some} \quad x^* \neq 0. \tag{2.44}$$

(iv) Show that the separation property (2.44) always implies the global set extremality (2.43), without imposing either the convexity of Ω_1, Ω_2 or the difference interiority condition $\operatorname{int}(\Omega_1 - \Omega_2) \neq \emptyset$ as in (iii).

Hint: Use the definitions, and apply the convex separation theorem to the sets $\Lambda_1 := \Omega_1 - \Omega_2$ and $\Lambda_2 := \{0\}$ in (iii). Compare it with the proof of [538, Theorem 2.2].

Exercise 2.34 (Approximate and Exact Versions of the Convex Extremal Principle in Banach Spaces). Let Ω_1 and Ω_2 be closed and convex subsets of a Banach space X, and let $\bar{x}$ be any common point of the sets Ω_1, Ω_2.

(i) Show that the extremality of Ω_1, Ω_2 in the sense of (2.43) yields the validity of the approximate extremal principle relationships: for any $\varepsilon > 0$, there exist $x_i \in B_\varepsilon(\bar{x}) \cap \Omega_i$ and $x_i^* \in N(x_{i\varepsilon};\Omega_i) + \varepsilon\mathbb{B}^*$ as $i = 1, 2$ such that

$$x_1^* + x_2^* = 0 \quad \text{and} \quad \|x_1^*\| = \|x_2^*\| = 1.$$

(ii) Assuming that one of the sets Ω_1, Ω_2 is SNC at $\bar{x}$, prove that the above extremality of these sets is equivalent to the approximate extremal principle conditions in (ii) as well as to the separation property (2.44).

(iii) Deduce from (ii) the seminal Bishop-Phelps theorem on the density of the support points on boundaries of closed and convex subsets of general Banach spaces.

Hint: To verify (i), invoke the Ekeland's variational principle, and compare this with the proof of [538, Theorem 2.5].

Exercise 2.35 (Violation of the Conic Extremal Principle in Hilbert Spaces). Let X be an arbitrary Hilbert space with $\dim X = \infty$. Give an example of *half-spaces* $\{\Lambda_i\}_{i \in \mathbb{N}}$ satisfying the assumptions of Theorem 2.9 for which CEP fails.

Exercise 2.36 (Weak Contingent Extremal Principle in Reflexive Spaces). We say that $\bar{x} \in \cap_{i=1}^s \Omega_i$ is a *weak contingent locally extremal point* of the set systems $\{\Omega_1, \ldots, \Omega_s\}$ in X if the system of weak contingent cones $\{T_W(\bar{x};\Omega_i)\}$, $i = 1, \ldots, s$, is extremal at the origin in the sense of Definition 2.6(a). Assume that $\bar{x}$ is such a point and that the space X is reflexive.

(i) Show that the approximate extremal principle holds at $\bar{x}$.

(ii) Assume in addition that all but one of the sets Ω_i, $i = 1, \ldots, s$, are SNC at $\bar{x}$, and show that in this case the exact extremal principle holds at $\bar{x}$.

Hint: Verify first that the SNC property of Ω_i at $\bar{x}$ yields this property for $T_W(\bar{x};\Omega_i)$ at the origin, and then proceed as in the proof of [568, Theorem 7.3].

Exercise 2.37 (Rated Extremal Principle in Finite Dimensions). We say that $\bar{x}\in\cap_{i=1}^s\Omega_i$ is a (locally) *rated extremal point* of rank $\alpha\in[0,1)$ for the set system $\{\Omega_1,\ldots,\Omega_s\}$ in a Banach space X if there are sequences $\{a_{ik}\}\subset X$, $i=1,\ldots,s$, and a positive number γ such that $r_k:=\max_i\|a_{ik}\|\to 0$ as $k\to\infty$ and

$$\bigcap_{i=1}^s\big(\Omega_i-a_{ik}\big)\cap\big(\bar{x}+\gamma r_k^\alpha\mathbb{B}\big)=\emptyset\ \text{ for large }\ k\in\mathbb{N}.$$

(i) Give an example of two sets in $\mathbb{R}^2$ for which $\bar{x}=(0,0)$ is a rated extremal point with $\alpha=0.5$ while not being locally extremal in the sense of Definition 2.1.

(ii) Show by using the method of metric approximations that any rated extremal point of rank $\alpha\in[0,1)$ for systems of finitely many (closed) sets in $\mathbb{R}^n$ satisfies the relationships of the basic/exact extremal principle.

(iii) Give an example illustrating the failure of this result for $\alpha=1$.

(iv) Show that a rated extremal point $\bar{x}$ of rank $\alpha\in[0,1)$ satisfies the relationships of the approximate extremal principle in Asplund spaces and the relationships of the exact extremal principle if all but one of the sets Ω_i are SNC at $\bar{x}$.

(v) Provide an extension of the rated extremal principle for infinitely many sets under an appropriate growth condition of the rate rank.

Hint: Proceed as in the case of $\alpha=0$, and compare with [567].

Exercise 2.38 (Ekeland's Variational Principle in Metric Spaces). Let (X,d) be a metric space. Show that the conditions of Ekeland's variational principle formulated in Corollary 2.13 with the norm $\|\cdot\|$ replaced by the distance function $d(\cdot,\cdot)$ hold under the completeness of the space X. Furthermore, the validity of these conditions characterize the completeness of (X,d).

Hint: Starting with the given point x_0 and assuming that $\varepsilon=\lambda=1$ without loss of generality, construct the iterates $\{x_k\}$ by

$$x_{k+1}\in T(x_k)\ \text{ and }\ \varphi(x_{k+1})<\inf_{x\in T(x_k)}\varphi(x)+\frac{1}{k},\quad k\in\mathbb{N},$$

where $T(x):=\{u\in X|\ \varphi(u)+d(x,u)\le\varphi(x)\}$. Observing that the sets $T(x_k)$ are nonempty and closed with $T(x_{k+1}\subset T(x_k)$ and $\operatorname{diam}T(x_k)\to 0$ as $k\to\infty$, conclude by the completeness of X that $\cap_{k=1}^\infty T(x_k)=\{\bar{x}\}$ for some $\bar{x}\in X$, which is actually the required point. Compare this with [522, Theorem 2.26], where the converse statement is also verified.

Exercise 2.39 (Lower Subdifferential Variational Principle). Prove that for every (l.s.c. and) bounded from below function $\varphi\colon X\to\overline{\mathbb{R}}$ on an *Asplund* space X, for any $\varepsilon,\lambda>0$ and $x_0\in X$ with $\varphi(x_0)<\inf_X\varphi+\varepsilon$, there are $\bar{x}\in X$ and

$$x^*\in\widehat{\partial}\varphi(\bar{x})\ \text{ with }\ \|\bar{x}-x_0\|<\lambda,\ \varphi(\bar{x})<\inf_X\varphi+\varepsilon,\ \|x^*\|<\varepsilon/\lambda.$$

Hint: Employ Ekeland's variational principle and then the approximate extremal principle in Asplund spaces; compare it with [522, Theorem 2.28].

Exercise 2.40 (Upper Subdifferential Variational Principle). Prove that for every (l.s.c. and) bounded from below function $\varphi\colon X\to\overline{\mathbb{R}}$ on a *Banach* space X, for any $\varepsilon,\lambda>0$ and $x_0\in X$ with $\varphi(x_0)<\inf_X\varphi+\varepsilon$, there exists $\bar{x}\in X$ with $\|\bar{x}-x_0\|<\lambda$ and $\varphi(\bar{x})<\inf_X\varphi+\varepsilon$ such that

$$\|x^*\|<\varepsilon/\lambda\ \text{ for all }\ x^*\in\widehat{\partial}^+\varphi(\bar{x}).$$

Hint: Combine the usage of Ekeland's variational principle with the first part of Theorem 1.27 (smooth variational description) and Proposition 1.30(ii), which both hold in general Banach spaces; compare it with [522, Theorem 2.30] and the proof of the second part of Theorem 2.14.

Exercise 2.41 (Smooth Variational Principles in Asplund Spaces).

(i) Prove the following smooth variational principle, which is an enhanced version of the Borwein-Preiss one: Given a (l.s.c. and) bounded from below function $\varphi\colon X \to \overline{\mathbb{R}}$ on a *Fréchet smooth* space X, for any $\varepsilon, \lambda > 0$ and $x_0 \in X$ with $\varphi(x_0) < \inf_X \varphi + \varepsilon$, there is a *concave* Fréchet smooth function $\psi\colon X \to \mathbb{R}$ such that $\|\bar{x} - x_0\| < \lambda$, $\varphi(\bar{x}) < \inf_X \varphi + \varepsilon$, $\|\nabla\psi(\bar{x})\| < \varepsilon/\lambda$, and

$$\varphi(\bar{x}) = \psi(\bar{x}), \quad \varphi(x) \geq \psi(x) + \|x - \bar{x}\|^2 \quad \text{whenever } x \in X. \tag{2.45}$$

The Fréchet smoothness of X is also *necessary* for the *concavity* of ψ in (2.45).

Hint: To verify the sufficient part of this statement, use the proofs of Theorem 1.10(ii) and its subgradient counterpart in Theorem 1.27 holding in any Fréchet smooth space; see Exercise 1.51(ii). To justify the necessity part, apply (2.45) to $\varphi(x) := 1/\|x\|$, find the corresponding function ψ, form the convex and Fréchet smooth function $p(x) := -\psi(x + v) + 1/\|v\|$, and consider the Minkowski gauge

$$g(x) := \inf\big\{\lambda > 0 \big|\ x \in \lambda\Omega\big\} \ \text{ with } \ \Omega := \big\{x \in X \big|\ p(x) \leq 1/(2\|v\|)\big\},$$

which defines the equivalent norm $n(x) := g(x) + g(-x)$ on X. Since p is of class $\mathcal{C}^1$ and convex, the Fréchet differentiability of g on $X \setminus \{0\}$ is equivalent to the Gâteaux one, and thus it remains to check that $\partial g(x)$ is a singleton at nonzero points as in the proof of the corresponding parts of [522, Theorem 2.31].

(ii) Derive the $\mathcal{S}$-*smooth* versions (while without the concavity property of ψ) of (i) for Asplund spaces admitting $\mathcal{S}$-smooth bump functions of the classes listed in Exercise 1.51(iii). *Hint:* Compare this with [522, Theorem 2.31](ii).

Exercise 2.42 (Regular Normals to Set Intersections via the Extremal Principle). Let Ω_1, Ω_2 be (closed) subsets of an Asplund space X, and let $\bar{x} \in \Omega_1 \cap \Omega_2$.

(i) Show that for any $x^* \in \widehat{N}(\bar{x}; \Omega_1 \cap \Omega_2)$ and $\varepsilon > 0$, there exist $\lambda \geq 0$, $x_i \in \Omega_i \cap (\bar{x} + \varepsilon\mathbb{B})$, and $x_i^* \in \widehat{N}(x_i; \Omega_i) + \varepsilon\mathbb{B}^*$, $i = 1, 2$, such that

$$\lambda x^* = x_1^* + x_2^*, \quad \max\big\{\lambda, \|x^*\|\big\} = 1. \tag{2.46}$$

Hint: Proceed similarly to the proof of Theorem 2.16 with applying the approximate extremal principle instead of the exact one by using the sum norm (1.18) on $X \times \mathbb{R}$. Compare this with [522, Lemma 3.1].

(ii) Obtain conditions ensuring that $\lambda \neq 0$ in (2.46). *Hint:* Consult with [583] for various results of this type and their uniform versions.

Exercise 2.43 (Intersection Rules for Basic Normals to Nonconvex Sets in Asplund Spaces). We say that the sets $\{\Omega_1, \Omega_2\}$ in a Banach space X satisfies the *limiting qualification condition* at $\bar{x} \in \Omega_1 \cap \Omega_2$ if

$$\big[\|x_{1k}^* + x_{2k}^*\| \to 0 \ \text{ as } \ k \to \infty\big] \Longrightarrow x_1^* = x_2^* = 0$$

for any sequences $x_{ik} \xrightarrow{\Omega_i} \bar{x}$, $x_{ik}^* \xrightarrow{w^*} x_i^*$, and $\varepsilon_k \downarrow 0$ with $x_{ik}^* \in \widehat{N}_{\varepsilon_k}(x_{ik}; \Omega_i)$, $i = 1, 2$.

(i) Let X be Asplund. Based on (1.59), show that ε_k can be dropped in the definition above and that the limiting qualification condition is implied by the normal one (2.26). Give an example of sets for which the reverse implication fails.

(ii) Prove the validity of the basic intersection rule (2.16) in Asplund spaces provided that the limiting qualification condition holds and *one* of the sets Ω_i is SNC at $\bar{x}$. *Hint*: Pass to the limit from the fuzzy intersection rule of Exercise 2.42, and compare with the proof of a more general result in [522, Theorem 3.4].

(iii) Give an example showing that the SNC assumption is essential for the validity of the intersection rule even in the Hilbert space setting.

(iv) Consider the intersection of finitely many subsets of an Asplund space with $\bar{x} \in \Omega := \Omega_1 \cap \ldots \cap \Omega_s$, and verify the inclusion

$$N(\bar{x}; \Omega) \subset N(\bar{x}; \Omega_1) + \ldots + N(\bar{x}; \Omega_s) \tag{2.47}$$

provided that all but one of the sets Ω_i are SNC at $\bar{x}$ and that the following *normal qualification condition* for *finitely many* sets is satisfied:

$$\big[x_1^* + \ldots + x_s^* = 0,\ x_i^* \in N(\bar{x}; \Omega_i)\big] \Longrightarrow x_i^* = 0, \quad i = 1, \ldots, s.$$

Show that Ω is normally regular at $\bar{x}$ and (2.47) holds as equality if all Ω_i are normally regular at $\bar{x}$. *Hint:* Proceed by induction with the usage of (2.47) for $s = m$, $m \geq 2$, to verify the validity of the normal qualification condition for $s = m + 1$.

Exercise 2.44 (Normals to Intersections of Convex Sets in Locally Convex Topological Vector Spaces). Let Ω_1 and Ω_2 be nonempty convex subsets of a LCTV space X, and let $\bar{x} \in \Omega_1 \cap \Omega_2$.

(i) Assuming that there is a bounded convex neighborhood V of $\bar{x}$ such that

$$0 \in \operatorname{int}\big(\Omega_1 - (\Omega_2 \cap V)\big), \tag{2.48}$$

prove the precise normal cone intersection formula

$$N(\bar{x}; \Omega_1 \cap \Omega_2) = N(\bar{x}; \Omega_1) + N(\bar{x}; \Omega_2). \tag{2.49}$$

Hint: Show that the convex sets

$$\Theta_1 := \Omega_1 \times [0, \infty) \ \text{ and } \ \Theta_2 := \big\{(x, \mu) \in X \times \mathbb{R} \big| \ x \in \Omega_1 \cap V,\ \mu \leq \langle x^*, x - \bar{x}\rangle\big\}$$

form an extremal system (2.1), and then proceed by applying the convex extremal principle from Exercise 2.34; compare it with the proof of [538, Theorem 3.1].

(ii) Establish relationships between (2.48), the condition $0 \in \operatorname{int}(\Omega_1 - \Omega_2)$, and the classical qualification condition $\Omega_1 \cap (\operatorname{int} \Omega_2) \neq \emptyset$ for the validity of the normal cone formula (2.49) in general LCTV spaces and also in normed spaces.

(iii) Assuming that X is Banach, that both sets Ω_1, Ω_2 are closed, and that $\operatorname{int}(\Omega_1 - \Omega_2) \neq \emptyset$, prove the equivalence

$$\big[0 \in \operatorname{core}(\Omega_1 - \Omega_2)\big] \Longleftrightarrow \big[0 \in \operatorname{int}(\Omega_1 - \Omega_2)\big],$$

where the symbol "core" stands for the *algebraic core* of a set defined by

$$\operatorname{core}\Omega := \big\{x \in \Omega \big| \ \forall v \in X\ \exists \gamma > 0 \text{ such that } \ x + tv \in \Omega \text{ whenever } \ |t| < \gamma\big\}.$$

Hint: Use the equality $\operatorname{int}\Omega = \operatorname{core}\Omega$ that holds for closed and convex subsets of Banach spaces; see, e.g., [114, Theorem 4.1.8].

Exercise 2.45 (Preservation of the SNC Property for Set Intersections). Let Ω_1 and Ω_2 be subsets of an Asplund space X, and let $\bar{x} \in \Omega_1 \cap \Omega_2$.

(i) Prove that if both Ω_i are SNC at $\bar{x}$ and the normal qualification condition (2.26) holds, then $\Omega_1 \cap \Omega_2$ is also SNC at $\bar{x}$. *Hint*: Apply the result from Exercise 2.42 based on the extremal principle, and compare with the proof of a more general statement in [522, Theorem 3.79].

(ii) Show that the normal qualification condition (2.26) is essential in infinite dimensions. Could it be replaced by the limiting qualification condition?

(iii) Derive an extension of (i) to the case of finitely many sets.

Exercise 2.46 (Inner Semicontinuity on the Graph and Inner Semicompactness on the Domain). Let $F\colon X \rightrightarrows Y$ be a set-valued mapping between Banach spaces, and let $(\bar{x}, \bar{y}) \in \text{gph}\, F$. We say that F is *inner semicontinuous* at the *graph* point $(\bar{x}, \bar{y})$ if for every sequence $x_k \stackrel{\text{dom}\, F}{\longrightarrow} \bar{x}$, there is a sequence $y_k \in F(x_k)$ that converges to $\bar{y}$ as $k \to \infty$. The mapping F is *inner semicompact* at the *domain* point $\bar{x}$ if for every sequence $x_k \stackrel{\text{dom}\, F}{\longrightarrow} \bar{x}$, there is a sequence $y_k \in F(x_k)$ that contains a convergent subsequence as $k \to \infty$.

(i) Observe that the inner semicontinuity of F at $(\bar{x}, \bar{y})$ for *every* $\bar{y} \in F(\bar{x})$ reduces to the inner semicontinuity of F at the domain point $\bar{x}$ from (1.20).

(ii) Check that if F is a *locally compact* near $\bar{x}$ (i.e., the image of a neighborhood U of $\bar{x}$ under F is enclosed into a compact set; this corresponds to the local boundedness of F when $\dim Y < \infty$ as defined in Subsection 1.2.1), then F is inner semicompact *around* this point, i.e., for each $x \in U$.

(iii) Give an example showing that, in contrast to the inner semicontinuity above, the inner semicompactness can't be equivalently formulated via the convergence of the entire sequence $\{y_k\}$, $k \in \mathbb{N}$ and requires passing to a *subsequence*.

Exercise 2.47 (Normals to Sums of Sets in Infinite Dimensions). Let $\Omega_1, \Omega_2 \subset X$ for an Asplund space X with $\bar{x} \in \Omega_1 + \Omega_2$, and let $S\colon X \rightrightarrows X^2$ be defined by

$$S(x) := \big\{(x_1, x_2) \in X \times X \big|\, x_1 + x_2 = x,\ x_1 \in \Omega_1,\ x_2 \in \Omega_2\big\}. \tag{2.50}$$

Verify the following sum rules for basic normals:

(i) If the mapping S in (2.50) is inner semicompact at $\bar{x}$, then

$$N(\bar{x}; \Omega_1 + \Omega_2) \subset \bigcup_{(x_1, x_2) \in S(\bar{x})} N(x_1; \Omega_1) \cap N(x_2; \Omega_2).$$

(ii) If S is inner semicontinuous at $(\bar{x}, \bar{x}_1, \bar{x}_2)$ for some $(\bar{x}_1, \bar{x}_2) \in S(\bar{x})$, then

$$N(\bar{x}; \Omega_1 + \Omega_2) \subset N(\bar{x}_1; \Omega_1) \cap N(\bar{x}_2; \Omega_2).$$

Hint: Reduce it to the intersection rule from Exercise 2.43 for the sets $\widetilde{\Omega}_1 := \Omega_1 \times X$ and $\widetilde{\Omega}_2 := X \times \Omega_2$ in the Asplund space X^2; compare with [522, Theorem 3.7].

Exercise 2.48 (SNC Property Under Set Additions). Let X be Asplund, and let $\Omega_1, \Omega_2 \subset X$ with $\bar{x} \in \Omega_1 + \Omega_2$. Define a set-valued mapping $S\colon X \rightrightarrows X^2$ by (2.50), and prove that the set $\Omega_1 + \Omega_2$ is SNC at $\bar{x}$ if either

(a) S is inner semicompact at $\bar{x}$, and for each $(x_1, x_2) \in S(\bar{x})$, one of the sets Ω_1, Ω_2 is SNC at x_1 and x_2, respectively, or

(b) S is inner semicontinuous at $(\bar{x}_1, \bar{x}_2, \bar{x})$ with some $(\bar{x}_1, \bar{x}_2) \in S(\bar{x})$, and one of the sets Ω_1, Ω_2 is SNC at $\bar{x}_1$ and $\bar{x}_2$, respectively.

Hint: Check the SNC property of the sum $\Omega_1 + \Omega_2$ by reducing it to that for the intersection of $\widetilde{\Omega}_1, \widetilde{\Omega}_2 \subset X^2$ as in Exercise 2.47; compare with [522, Theorem 3.73].

Exercise 2.49 (SNEC Property of Extended-Real-Valued Functions). A function $\varphi\colon X \to \overline{\mathbb{R}}$ on a Banach space X is *sequentially normally epicompact* (SNEC) at $\bar{x} \in \text{dom}\, \varphi$ if its epigraphical set is SNC at $(\bar{x}, \varphi(\bar{x}))$.

(i) Show that the SNC property of φ at $\bar{x}$ (i.e., of its graph at $(\bar{x}, \varphi(\bar{x}))$ implies that both φ and $-\varphi$ are SNEC at this point. Does the reverse implication hold?

(ii) Show that the local Lipschitz continuity of φ around $\bar{x}$ implies both SNC and SNEC properties of φ at this point.

Hint: These properties are epigraphical and graphical specifications of the relationships in Exercise 2.29 for the case of extended-real-valued functions.

Exercise 2.50 (Subgradient Description of the SNEC Property). Let X be Asplund. Then the SNEC property of any (l.s.c.) function $\varphi\colon X \to \overline{\mathbb{R}}$ at $\bar{x} \in \operatorname{dom}\varphi$ admits the following subgradient description: for every sequences $x \xrightarrow{\varphi} \bar{x}$, $\lambda_k \downarrow 0$, and $x_k^* \in \lambda_k \widehat{\partial}\varphi(x_k)$, we have the implication

$$x_k^* \xrightarrow{w^*} 0 \Longrightarrow \|x_k^*\| \to 0 \ \text{ as } \ k \to \infty.$$

Hint: Use description (1.71) of singular subgradients in Asplund spaces, and compare the proof with the one given in [522, Corollary 2.39].

Exercise 2.51 (Basic Normals and the SNC Property for Sets Defined by Inequality Constraints). Let X be an Asplund space.

(i) Consider the level set $\Omega := \{x \in X|\ \varphi(x) \le 0\}$, where $\varphi\colon X \to \overline{\mathbb{R}}$ is merely l.s.c. around $\bar{x}$ with $\varphi(\bar{x}) = 0$. Assume that $0 \notin \partial\varphi(\bar{x})$ and that φ is SNEC at $\bar{x}$. Show that Ω is SNC at the reference point and that

$$N(\bar{x};\Omega) \subset \big[\operatorname{cone}\partial\varphi(\bar{x})\big] \cup \partial^\infty\varphi(\bar{x}),$$

where the equality holds if φ is epigraphically regular at $\bar{x}$. *Hint*: To verify the SNC property of Ω at $\bar{x}$, apply the result of Exercise 2.45 to the intersection of $\Omega_1 := \operatorname{epi}\varphi$ and $\Omega_2 := \{(x,\alpha) \in X \times \mathbb{R}|\ \alpha = 0\}$. In this way the claimed normal cone representations can be deduced from Exercise 2.43(ii) in Asplund spaces and from Theorem 2.16 in finite dimensions.

(ii) Consider the set $\Omega := \{x \in X|\ \varphi_i(x) \le 0,\ i = 1,\ldots,m\}$, and denote by

$$I(\bar{x}) := \big\{i \in \{1,\ldots,m\}\big|\ \varphi_i(\bar{x}) = 0\big\} \tag{2.51}$$

the set of *active constraint indices*. Assume that the functions φ_i are locally Lipschitzian around $\bar{x}$ for $i \in I(\bar{x})$ and *upper* semicontinuous for $i \in \{1,\ldots,m\} \setminus I(\bar{x})$. Show that the constraint qualification condition

$$0 \notin \operatorname{co}\big[\partial\varphi_i(\bar{x})\big|\ i \in I(\bar{x})\big]$$

ensures the simultaneous validity of the SNC property of Ω at $\bar{x}$ and the inclusion

$$N(\bar{x};\Omega) \subset \bigcup\Big\{\sum \lambda_i\partial\varphi_i(\bar{x})\Big|\ \lambda_i \ge 0,\ \lambda_i\varphi_i(\bar{x}) = 0,\ i = 1,\ldots,m\Big\},$$

which holds as equality if φ_i are lower regular at $\bar{x}$ for all $i \in I(\bar{x})$. In this case the set Ω is normally regular at $\bar{x}$. *Hint*: Use the results from (i) and the intersection rules for the normal cone and SNC property from Exercises 2.43 and 2.45.

(iii) Obtain extensions of (ii) to the case where φ_i are merely l.s.c. for $i \in I(\bar{x})$.

Exercise 2.52 (Basic Normals and the SNC Property for Sets Defined by Equality Constraints). Let X be an Asplund space.

(i) Consider the set $\Omega := \{x \in X|\ \varphi(x) = 0\}$, where $\varphi\colon X \to \overline{\mathbb{R}}$ is continuous around $\bar{x} \in \Omega$. Show that the condition $0 \notin \partial\varphi(\bar{x}) \cup \partial(-\varphi)(\bar{x})$ ensures that the set Ω is SNC at $\bar{x}$ and that the inclusion

$$N(\bar{x};\Omega) \subset \big[\operatorname{cone}\big\{\partial\varphi(\bar{x}) \cup \partial(-\varphi)(\bar{x})\big\}\big] \cup \big[\partial^\infty\varphi(\bar{x}) \cup \partial^\infty(-\varphi)(\bar{x})\big]$$

holds with the equality and normal regularity of Ω therein if φ is strictly differentiable at $\bar{x}$. *Hint*: Apply the result from Exercise 2.45(i) to the intersection of the sets $\Omega_1 := \operatorname{gph}\varphi$ and $\Omega_2 := \{(x,\alpha) \in X \times \mathbb{R}|\ \alpha = 0\}$ to verify the SNC property of Ω at $\bar{x}$ and then Exercise 2.45(ii) to get the claimed normal cone representations.

(ii) Let $\Omega := \{x \in X|\ \varphi_i(x) = 0,\ i = 1,\ldots,m\}$, where all the functions φ_i are locally Lipschitzian around $\bar{x}$. Assume the validity of the constraint qualification condition $0 \notin \operatorname{co}\big\{\partial\varphi_i(\bar{x}) \cup \partial(-\varphi_i)(\bar{x})|\ i = 1,\ldots,m\big\}$. Then the set Ω is SNC at $\bar{x}$, and we have the inclusion

$$N(\bar{x};\Omega) \subset \Big\{\sum \lambda_i\big[\partial\varphi_i(\bar{x}) \cup \partial(-\varphi_i)(\bar{x})\big]\Big|\ \lambda_i \ge 0,\ i = 1,\ldots,m\Big\},$$

which holds as equality with the normal regularity of Ω if φ_i are strictly differentiable at $\bar{x}$. *Hint*: Combine the result from (i) with those in Exercises 2.43 and 2.45.

(iii) Extend (ii) to the case where φ_i are merely continuous around $\bar{x}$.

Exercise 2.53 (Basic Normals and the SNC Property of Constraint Systems in Nonlinear Programming). Let X be an Asplund space. Consider the set

$$\Omega := \big\{x \in X \big|\, \varphi_i(x) \le 0,\ i = 1, \ldots, m, \ \text{ and } \ \varphi_i(x) = 0,\ i = m+1, \ldots, m+r\big\}.$$

(i) Assume that all the functions φ_i are strictly differentiable at $\bar{x}$, and impose the *Mangasarian-Fromovitz constraint qualification* (MFCQ):

(a) $\nabla\varphi_{m+1}(\bar{x}), \ldots, \nabla\varphi_{m+r}(\bar{x})$ are linearly independent;

(b) there is $u \in X$ satisfying the conditions

$$\langle\nabla\varphi_i(\bar{x}), u\rangle < 0,\ i \in I(\bar{x}), \ \text{ and } \ \langle\nabla\varphi_i(\bar{x}), u\rangle = 0,\ i = m+1, \ldots, m+r,$$

where $I(\bar{x})$ is defined in (2.51). Show that in this case, the set Ω is SNC and normally regular at $\bar{x}$, and we have the normal cone representation

$$N(\bar{x}; \Omega) = \Big\{\sum_{i=1}^{m+r} \lambda_i \nabla\varphi_i(\bar{x}) \Big|\ \lambda_i \ge 0,\ \lambda_i\varphi_i(\bar{x}) = 0 \ \text{ for } \ i = 1, \ldots, m,$$
$$\text{and } \ \lambda_i \in \mathbb{R} \ \text{ for } \ i = m+1, \ldots, m+r\Big\}.$$

Hint: Deduce these results from the previous exercises and the fact that $\partial\varphi(\bar{x}) = \{\nabla\varphi(\bar{x})\}$ for strictly differentiable functions.

(ii) Assume that all the functions φ_i are locally Lipschitzian around $\bar{x}$. Formulate the corresponding generalized version of MFCQ in this case, and derive extensions of the normal cone representation and SNC results from (i) to the nondifferentiable case. *Hint:* Compare these results with [523, Theorem 3.86].

Exercise 2.54 (Subdifferential and SNEC Sum Rules for Functions Defined on Infinite-Dimensional Spaces).

(i) Extend the subdifferential sum rules of Theorem 2.19 for (locally l.s.c.) functions $\varphi_1, \varphi_2\colon X \to \overline{\mathbb{R}}$ on an Asplund space X provided that one of them is SNEC at $\bar{x}$. *Hint:* Proceed as in the proof of Theorem 2.19 by the reduction to the normal cone intersection rule from Exercise 2.43 under the normal qualification condition.

(ii) Show that the sum $\varphi_1 + \varphi_2$ is SNEC at $\bar{x} \in \operatorname{dom}\varphi_1 \cap \operatorname{dom}\varphi_2$ if both functions φ_i have this property and the qualification condition (2.34) holds. *Hint:* Reduce this to the SNC result for sets from Exercise 2.45(i).

(iii) Let $\varphi_1, \varphi_2\colon X \to \overline{\mathbb{R}}$ be convex functions on a LCTV space X. Using the geometric approach implemented in the proof of Theorem 2.19, derive the convex subdifferential sum rule from the intersection rule given in Exercise 2.44(i).

Exercise 2.55 (Minimality of the Basic Subdifferential). Let $\widehat{\partial}^\bullet\varphi\colon X \rightrightarrows X^*$ be an abstract presubdifferential on the class of l.s.c. functions $\varphi\colon X \to \overline{\mathbb{R}}$ with $\varphi(\bar{x}) < \infty$ defined on a Banach space X and satisfied the following properties:

(a) $\widehat{\partial}^\bullet\phi(u) = \widehat{\partial}^\bullet\varphi(x+u)$ for $\phi(u) := \varphi(x+u)$ and $x, u \in X$.

(b) $\widehat{\partial}^\bullet\varphi(x)$ is contained in the subdifferential of convex analysis for convex continuous functions represented in the form

$$\varphi(x) := \langle x^*, x\rangle + \varepsilon\|x\| \ \text{ whenever } \ x^* \in X^*,\ \varepsilon > 0. \tag{2.52}$$

(c) For any $\eta > 0$ and any functions φ_i, $i = 1, 2$, such that φ_1 is of type (2.52) and the sum $\varphi_1 + \varphi_2$ attains its local minimum at $x = 0$, there are $x_1, x_2 \in \eta\mathbb{B}$ satisfying the conditions $|\varphi_2(x_2) - \varphi_2(0)| \le \eta$ and

$$0 \in \widehat{\partial}^{\bullet}\varphi_1(x_1) + \widehat{\partial}^{\bullet}\varphi_2(x_2) + \eta\mathbb{B}^*.$$

Prove that we always have the inclusion

$$\partial\varphi(\bar{x}) \subset \mathop{\mathrm{Lim\,sup}}_{x \xrightarrow{\varphi} \bar{x}} \widehat{\partial}^{\bullet}\varphi(x)$$

for the basic subdifferential $\partial\varphi(\bar{x})$ via the sequential weak* limit of $\widehat{\partial}^{\bullet}\varphi(x)$.

Hint: Proceed by using the analytic representation (1.68) of the basic subdifferential, and check that the presubdifferential $\widehat{\partial}\varphi$ from (1.33) satisfies all the properties listed in (a)–(c) in the case of Asplund spaces. Compare it with [580, Theorem 9.7].

2.6 Commentaries to Chapter 2

Section 2.1. The single most important conceptual ingredient of our geometric dual-space approach to variational analysis is the *extremal principle* for systems of sets as well as its proof based on the *method of metric approximations* (MMA) initiated by the author [502, 504] in the context of general problems of optimization and control. Recall that the very notion of the (basic) normal cone (1.4) appears in [502] as a by-product of the method of metric approximations.

The term "extremal principles" for geometric variational principles of the type presented in Section 2.1 was coined by the author in [511], while the result of Theorem 2.3 has been derived earlier via the MMA in finite-dimensional spaces in the joint papers with Kruger [440, 441] under the name of "generalized Euler equation." It has also been extended therein to Fréchet smooth spaces in an approximate form by involving ε-normals (1.6) as $\varepsilon > 0$. The Euler equation terminology came from the analogy with the "abstract Euler equation" used by Dubovitskii and Milyutin [234] to describe the result of conic convex separation in their scheme of obtaining necessary optimality conditions in problems of optimization and control. As proved in [579], the ε-version of the extremal principle from [440, 441] happened to be equivalent to the "fuzzy sum rule" suggested later by Ioffe [367].

The *approximate extremal principle* in the enhanced form of Corollary 2.5, playing a crucial role in infinite-dimensional spaces, was established by Mordukhovich and Shao [579] as a *characterization of Asplund spaces* via variational arguments involving Fréchet-like subgradients. Other proofs of this result are given in [580] by using the characterizations of *well-posedness* from [578] (cf. Section 3.1) and in [258, 522] by employing the *method of separable reduction*; see the cited publications for more details, discussions, and references. The state of the art of this method and its relationships with the approximate extremal principle can be found in the recent paper by Cúth and Fabian [187]. The line of equivalences from [579], with adding more results therein, was extended by Zhu [786] to Banach spaces with bornological smooth renorms. In parallel Borwein, Mordukhovich, and Shao [107] established the equivalence of bornological versions of the approximate extremal principle in Banach spaces with smooth renorms (resp. bump functions) to the smooth variational principles by Borwein and Preiss [108] (resp. by Deville, Godefroy and Zizler [205]). Some versions of the extremal principle and related results in terms of abstract normal cones and subdifferentials in Banach spaces can be found in [468, 515, 522].

The exact/limiting form (2.4)–(2.5) of the extremal principle holds in any Asplund space [580] provided that *all but one* set Ω_i are *sequentially normally compact* (SNC) at $\bar{x}$ in the sense of (2.41) introduced by the author and Shao in [582] (preprint of 1994) together with its *partial* counterpart (PSNC) for mappings $F\colon X \rightrightarrows Y$ as in (3.65). Then these properties were further developed and applied in [581] and subsequent publications. It turns out that the SNC property holds in any Banach spaces for sets that are *compactly epi-Lipschitzian* (CEL) in the sense of Borwein and Strójwas [109], which extends the *epi-Lipschitzian* property by Rockafellar [669]; see Exercise 2.29 for the definitions and more references. On the other hand, its PSNC counterpart is valid for *any Lipschitz-like* multifunction between Banach spaces as follows from the coderivative criterion for the Lipschitz-like property discussed in [522] and Chapter 3 below. While both SNC and

CEL properties are automatic in finite dimensions (and PSNC for $F\colon X \rightrightarrows Y$ is automatic when $\dim X < \infty$), well-developed *calculus/preservation rules* are available in [522, 589], the proofs of which are based on the *extremal principle*; see Sections 2.5 and 3.4. Note that *topological* counterparts of these properties (nets instead of sequences in the weak* convergence on dual spaces) were developed by Penot [635] (preprint of 1995). We refer the reader to [106] for complete characterizations of the CEL property of convex sets, to [371] for comparisons between the SNC and CEL properties for closed sets in Banach spaces, and to [259] for relationships between the SNC property and its topological counterpart; see [522, Remark 1.27] for a detailed summary. Further results and applications in this direction can be found in [731, 732].

A certain modification of global extremality for sets in LCTV spaces, which doesn't require the set closedness and nonempty intersection and occurs to be especially useful in the study of convex sets, has been recently suggested and investigated by the author and Nam [538]. Enhanced versions of the extremal principle in both approximate and exact forms were obtained in [538] in LCTV and normed spaces frameworks and then applied to generalized differential and conjugate calculi of convex sets and functions via a variational geometric approach; see [538, 541] and also Exercises 2.33 and 2.34 for some results and discussions.

Extended versions of extremal principle in both approximate and exact forms, involving *nonlinear deformations* of sets and set-valued mappings defined on metric spaces, were introduced and developed by Mordukhovich, Treiman, and Zhu [586] being particularly motivated by applications to some problems of multiobjective optimization; see Chapter 9. Another version of the nonconvex separation theorem for sets was established by Borwein and Jofré [102]. Further developments and applications in this direction can be found in [50, 114, 265, 433, 523, 685, 773, 774, 777, 787] along with other publications. We also mention here important results on the so-called *nonlinear separation* that were initiated by Gerstewitz (Tammer) [278] who was motivated by developing new scalarization techniques in vector optimization. Her idea was greatly elaborated and applied in many subsequent works; see, Eqs., [245, 279, 300, 321, 385, 389, 407, 409] and the references therein.

Section 2.2. The material of this section is rather fresh and has never appeared in the monographic literature. It concerns extremality notions and various extensions of the extremal principle to *infinite* (actually *countable*) systems of sets in finite-dimensional spaces. Besides being of undoubted mathematical interest for their own sake, this topic has been motivated by applications to optimization problems of semi-infinite programming considered in Chapters 7 and 8 below. Section 2.2 mostly follows the recent papers by Mordukhovich and Phan [568, 569] in the case of finite dimensions, while we present infinite-dimensional extensions and the rated version of the extremal principle from [567] in Section 2.5. The reader is referred to the subsequent papers by Kruger and López [436, 437] for further developments and applications in this direction based on somewhat different ideas.

Section 2.3. Ekeland's variational principle formulated in Corollary 2.13 is one of the first and most powerful results of modern variational analysis. From the very beginning [248, 249], it has been proved in *complete metric* spaces (characterizing in fact their completeness) by a rather complicated device involving transfinite induction and the Zorn lemma. A constructive proof in complete metric spaces was presented in [250] based on a personal communication with Michael Crandall; see Exercise 2.38. Observe that, being a metric space results, Ekeland's principle brought new and very important information in finite-dimensional spaces as well. Its short proof in $\mathbb{R}^n$ was given by Hiriart-Urruty [349].

The finite-dimensional geometry allows us to obtain a variational result in the general form of Theorem 2.12 taken from the author's early book [507]. By the choice of the function θ therein we can unify, in particular, Ekeland's principle and various smooth variational principles. This is useful in several applications as shown, e.g., in Theorem 2.14 and Proposition 2.15 also taken from [507]. Besides the Borwein-Preiss and Deville-Godefroy-Zizler variational principles and their enhanced forms, other smooth variational principles were obtained in the author's joint paper with Fabian [257]; see Exercise 2.41 for some results in this direction. It is remarkable to see, e.g., that the Fréchet renorming of Banach spaces is not only sufficient but also *necessary* for the *smoothness* and *concavity* of perturbations as in Exercise 2.41(i). The reader can find more information about various smooth variational principles in appropriate infinite-dimensional spaces and their relationships with the extremal principles in [522, Chapter 2] and the references therein.

In [522, Subsection 2.3.2] the reader can also find other, less known albeit very useful, types of variational principles. The first one, the *lower subdifferential variational principle*, was established by Mordukhovich and Wang [587] as yet another characterization of Asplund spaces. This result has the same form as Ekeland's variational principle while replacing the minimization condition in (2.23) by the subgradient estimate $\|x^*\| \le \varepsilon/\lambda$ for *some* $x^* \in \widehat{\partial}\varphi(\bar{x})$ at the suboptimal point $\bar{x}$, which is a nonsmooth extension of the almost stationary condition (2.24); see Exercise 2.39. The other result, obtained in the author's joint paper with Nam and Yen [546] and named there the *upper subdifferential variational principle*, justifies the validity of the latter estimate for *all* $x^* \in \widehat{\partial}^+\varphi(\bar{x})$ in an arbitrary Banach space provided that $\widehat{\partial}^+\varphi(\bar{x}) \ne \emptyset$; see Exercise 2.40. This also reduces to the almost stationary condition (2.24) when φ is Fréchet differentiable at $\bar{x}$.

Section 2.4. The results of this section (except Corollary 2.18 that is taken from [678]) are based on the author's paper [505], where the normal and subdifferential qualification conditions (2.26) and (2.34) were first introduced and the underlying intersection and sum rules (2.27), (2.35) were derived by the method of metric approximations; see [507, 522] for comprehensive accounts. Some (directionally) Lipschitzian versions of these results can be found in Ioffe [365] and Kruger [428, 430]. Another paper by Ioffe [368] dealt, by using a penalty function method clearly inspired by the MMA (which was acknowledged therein as well as in [364, 365]), with the case of l.s.c. functions on $\mathbb{R}^n$ under essentially more restrictive *tangential* qualification conditions formulated in terms of the directional derivative construction by Rockafellar [670] and the tangent cone by Clarke. More recent finite-dimensional results in this direction were established by Ioffe and Outrata [376] under certain calmness and metric qualification conditions. Various infinite-dimensional versions of the presented results were given in [114, 369, 375, 398, 399, 518, 522, 533, 580, 588, 610, 637, 685] and the references therein. Some of these and related results can be found in exercises to Chapters 2–4. Note, in particular, the validity of the comprehensive extensions of Theorems 2.16, 2.19 and their corollaries to the case of Asplund spaces under SNC assumptions of type (2.41) as well as its functional counterparts. As mentioned above such properties are automatic in finite dimensions and also hold in (generalized) Lipschitzian settings of Banach spaces; see [522] for more details.

The major qualification condition (2.26) was introduced in [505] under the name of the "generalized nonseparation property" for nonconvex sets in finite dimensions in order to derive the basic intersection rule in Theorem 2.16 by using the method of metric approximations. Its negation, which amounts to the relationships of the exact/basic extremal principle, was called in [505] the "generalized separation property." Both names reflect the fact that these properties are nonconvex generalizations of the corresponding ones for convex sets; see [507] for more details and discussions. Condition (2.26) was studied and applied in [522] under the name of the "normal qualification condition," which allowed us to derive the intersection rule for basic normals in finite and infinite dimensions. However, the weaker "limiting qualification condition" was proved to be sufficient for deriving the intersection rule in Asplund spaces under appropriate SNC/PSNC assumptions; see [522, Theorem 3.4].

More recently, another line of impressive applications of the qualification condition (2.26) has been developed in *algorithmic* aspects of feasibility and optimization for nonconvex problems. A pioneering work in this direction was done by Lewis, Luke, and Malick [458] who established a linear rate of local convergence of a nonconvex version of the (von Neumann) alternating projection algorithm and its averaged projection modification in the problem of finding an intersection point of two (and finitely many) nonconvex sets in $\mathbb{R}^n$ under the qualification condition (2.26) and its version for finitely many sets presented in Corollary 2.17. Analyzing the connection with the original development in (2.26) and the exact extremal principle while giving its algorithmic description, the authors of [458] interpreted the basic qualification condition (2.26) as the "linear regular intersection" of closed sets and made connections with metric regularity notions considered below in Chapter 3.

Further striking developments of the variational analysis approach to algorithmic issues were made by Bauschke et al. in [71, 72]. The authors of [71, 72] introduce a new notion of the "restricted normal cone" to closed sets in $\mathbb{R}^n$, which is an extension of our basic normal cone (1.4) while allowing them to essentially weaken the assumptions of [458] to obtain a local linear convergence of the alternating projection algorithm. The other paper [73] develops these ideas to solve numerically a sparsity optimization problem with affine constraints.

Lewis and Malick [459] were the first to establish an equivalence between the qualification condition (2.26) and the classical *transversality condition* in differential geometry for $\mathcal{C}^2$-smooth manifolds by using the coderivative criterion for metric regularity from Theorem 3.3(ii); the latter actually ensures such an equivalence for general closed sets in finite dimensions. Subsequent results on a local linear convergence of alternative projections and related algorithms for nonsmooth and nonconvex problems of feasibility and optimization have been recently developed in [229, 438, 480, 481, 617] and other publications with certain modifications of the normal qualification/transversality condition (2.26) defined under the names of "intrinsic transversality, separable intersection, subtransversality," etc. The transversality language was used in the recent book [375] without mentioning the introduction of (2.26) and the original derivation of the intersection rule under this qualification condition in [505] as well as omitting the references to the original paper [459] on transversality in the alternating projection algorithm and to a major contribution in this direction developed by Noll and Rondepierre [617] (preprint of 2013) concurrently to the paper by Drusvyatskiy, Ioffe, and Lewis [229].

The subdifferential qualification condition (2.34) was also first introduced in the author's paper [505] for establishing the basic subdifferential sum rule in Theorem 2.19. This result plays a crucial role in deriving other rules of subdifferential calculus and is deduced from the basic intersection rule of Theorem 2.16. The subdifferential qualification condition parallel to (2.34) while expressed via singular subgradients generated by the convexified normal cone was independently introduced by Rockafellar [675] who used it to obtain major calculus rules for Clarke's subgradients of extended-real-valued l.s.c. functions on finite-dimensional spaces. Conditions of these types and their indicator function versions as in (2.26) were the first ones to express qualification requirements in subdifferential calculus and constraint qualifications in nonsmooth optimization in the same (dual) terms as calculus rules and necessary optimality conditions. That was the reason to label such conditions in Ioffe [369] as well as in [376] and other publications as the "Mordukhovich-Rockafellar (MR) subdifferential qualification conditions." This name and any related discussions on (2.26) and (2.34) with the references to [505, 675] were not presented in [375], while the known qualification conditions for calculus rules were basically reformulated therein by using the transversality-related terminology.

Section 2.5. Most of the exercises presented in this section have hints and references to the publications, where the reader can find more details and sources. We comment only on the minimality result given in Exercise 2.55, which is taken from [580, Theorem 9.7] and [522, Proposition 2.45]. The origin of it should be traced to [368, Theorem 9] and [505, Theorem 4], where the minimality property was proved under somewhat different subdifferential requirements in finite dimensions. Note that the subsequent result by Ioffe [369, Proposition 8.2] doesn't imply that the nucleus of his G-subdifferential is smaller than our basic subdifferential $\partial\varphi(\bar{x})$ as mistakenly claimed therein. The mistake is due to the fact that the mapping $x \mapsto \partial\varphi(x)$ may not be of closed-graph in the norm$\times$weak* topology of $X \times X^*$ even for Lipschitzian functions on Asplund spaces. As the reader can see, the result presented in Exercise 2.55 yields that the basic subdifferential $\partial\varphi(\bar{x})$ is the smallest among all natural subdifferential constructions that are *sequentially* outer/upper semicontinuous on $\operatorname{gph}\varphi$. This includes, in particularly, all the "approximate" subdifferentials.

Chapter 3
Well-Posedness and Coderivative Calculus

This chapter concerns the study of two important topics in variational analysis that don't seem to be related to each other at the first glance. The first topic revolves around certain *well-posedness* issues for set-valued mappings/multifunctions, which constitute a large area of great significance for variational theory and its numerous applications. The area of well-posedness covers "good" properties of multifunctions that are desired to get achieved in the framework of variational analysis, optimization, equilibria, control, etc. It has been undoubtedly recognized from the viewpoints of both variational theory and applications that such properties include those known as *Lipschitzian stability*, *metric regularity*, and *covering/linear openness*, which are fundamental in fact for the whole field of nonlinear analysis, not only for its variational aspects. Properties of this type are defined in terms of a given multifunction and have nothing to do with notions of (generalized) differentiation.

It occurs nevertheless that the aforementioned properties admit complete *qualitative* and *quantitative characterizations* via our basic *coderivative* of multifunctions calculated exactly at the reference points. Such *pointbased* (i.e., expressed entirely at the point in question) *coderivative criteria* for general closed-graph multifunctions are derived in this chapter. However, applying them efficiently to particular models of optimization, equilibria, control, etc. requires comprehensive calculus rules, which open the gate to deal with structural mappings. The required *pointbased coderivative calculus* is presented below under certain *pointbased qualification conditions*. On the other side of developments, the obtained coderivative characterizations of the well-posedness properties allow us to verify that the imposed qualification conditions automatically hold for large classes of multifunctions satisfying these properties. Furthermore, involving the coderivative characterizations and calculus rules brings us to a rather surprising conclusion that the property of metric regularity *fails* to fulfill for major classes of *variational systems* given as solution maps to parametric generalized equations, variational inequalities , etc. Thus the results presented in this chapter fully justify *two-sided* relationships between well-posedness and pointbased coderivative calculus. Many other related results and well-posedness properties in finite and infinite dimensions are presented and largely discussed in the exercise and commentary sections.

B. S. Mordukhovich, *Variational Analysis and Applications*, Springer Monographs in Mathematics, https://doi.org/10.1007/978-3-319-92775-6_3

3.1 Well-Posedness Properties of Multifunctions

We start by formulating the fundamental well-posedness properties of multifunctions, which will be then characterized in terms of the coderivative (1.15).

3.1.1 Paradigm in Well-Posedness

In this chapter we mainly deal with multifunctions $F\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ between finite-dimensional spaces with discussing infinite-dimensional issues in the exercise and commentary sections. While multifunctions under consideration are generally set-valued, it doesn't exclude of course a single-valued case where F is denoted by $F = f\colon \mathbb{R}^n \to \mathbb{R}^m$ for notational convenience.

Definition 3.1 (Well-Posedness Properties). *Let $F\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, and let $(\bar{x}, \bar{y}) \in$* gph F *be the reference point. We say that:*

(a) *F has the* COVERING PROPERTY *around $(\bar{x}, \bar{y})$ with modulus $\kappa > 0$ if there are neighborhoods U of $\bar{x}$ and V of $\bar{y}$ such that*

$$F(x) \cap V + \kappa r\mathbb{B} \subset F(x + r\mathbb{B}) \;\; \textit{whenever} \;\; x + r\mathbb{B} \subset U \;\; \textit{as} \;\; r > 0. \tag{3.1}$$

The supremum of all the moduli $\{\kappa\}$ for which (3.1) holds with some neighborhoods U and V is called the EXACT COVERING BOUND *of F around $(\bar{x}, \bar{y})$ and is denoted by* cov $F(\bar{x}, \bar{y})$.

(b) *F is* METRICALLY REGULAR *around $(\bar{x}, \bar{y})$ with modulus $\mu > 0$ if there are neighborhoods U of $\bar{x}$ and V of $\bar{y}$ such that*

$$\operatorname{dist}\big(x; F^{-1}(y)\big) \le \mu \operatorname{dist}\big(y; F(x)\big) \;\; \textit{for all} \;\; x \in U, \;\; y \in V. \tag{3.2}$$

The infimum of all the moduli $\{\mu\}$ for which (3.2) holds with some neighborhoods U and V is called the EXACT REGULARITY BOUND *of F around $(\bar{x}, \bar{y})$ and is denoted by* reg $F(\bar{x}, \bar{y})$.

(c) *F is* LIPSCHITZ-LIKE *around $(\bar{x}, \bar{y})$ with modulus $\ell \ge 0$ if there are neighborhoods U of $\bar{x}$ and V of $\bar{y}$ such that*

$$F(x) \cap V \subset F(u) + \ell\|x - u\|\mathbb{B} \;\; \textit{for all} \;\; x, u \in U. \tag{3.3}$$

The infimum of all the moduli $\{\ell\}$ for which (3.3) holds with some neighborhoods U and V is called the EXACT LIPSCHITZIAN BOUND *of F around $(\bar{x}, \bar{y})$ and is denoted by* lip $F(\bar{x}, \bar{y})$.

All the three properties in Definition 3.1 are *stable/robust* with respect to small perturbations of the reference point $(\bar{x}, \bar{y})$. They postulate a "good behavior" of F *around* $(\bar{x}, \bar{y})$ and are highly interconnected; see Theorem 3.2.

The covering property is also known as *openness with linear rate* or *linear openness* of F around $(\bar{x}, \bar{y})$. For single-valued mappings f it somewhat relates, while being essentially different for nonlinear mappings, to a conventional *openness property* of f at $\bar{x}$ meaning that the f-image of every neighborhood of $\bar{x}$ contains/covers a neighborhood of $f(\bar{x})$ or equivalently

$$f(\bar{x}) \in \operatorname{int} f(U) \quad \text{for any neighborhood } U \text{ of } \bar{x}. \tag{3.4}$$

Property (3.1) postulates more, even for single-valued mappings: it ensures the *uniformity* of covering *around* $\bar{x}$ with *linear rate* quantified by κ. The cubic function $f(x) = x^3$ on $\mathbb{R}$ gives a simple example of a mapping having the openness property (3.4) at $\bar{x} = 0$ while not that with linear rate; see Fig. 3.1.

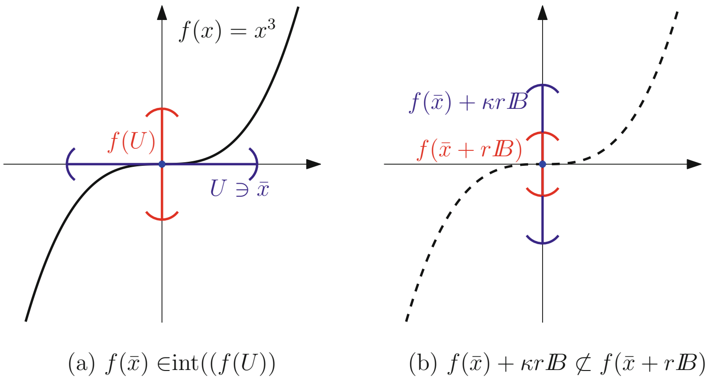


Fig. 3.1 Openness (a) but not linear openness (b).

Metric regularity (3.2) provides, for any (x, y) near $(\bar{x}, \bar{y})$, a *linear estimate* of the distance between x and the set of solutions to the (generalized) equation $y \in F(u)$ through the distance between y and $F(x)$, which is much easier to calculate; see Fig. 3.2. In particular settings it closely relates to the (local) *error bound* property that plays a significant role in theoretical and numerical aspects of optimization and its applications.

For single-valued mappings f the Lipschitz-like property (3.3) goes back to the classical local Lipschitzian behavior (1.26), while in the compact-valued case with $V = \mathbb{R}^m$ in (3.3), it reduces to the standard (Hausdorff) *local Lipschitzian* property of multifunctions. In the general case of V in (3.3), this condition is also known as the *pseudo-Lipschitz* or *Aubin property*, which is a graphical localization of Lipschitzian behavior for set-valued mappings.

The following result shows that all the properties from Definition 3.1 are in fact *equivalent* with the precise relationships between their *exact bounds*.

Theorem 3.2 (Equivalence Between Well-Posedness Properties). *Let* $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ *with* $(\bar{x}, \bar{y}) \in \operatorname{gph} F$*. The following are equivalent:*

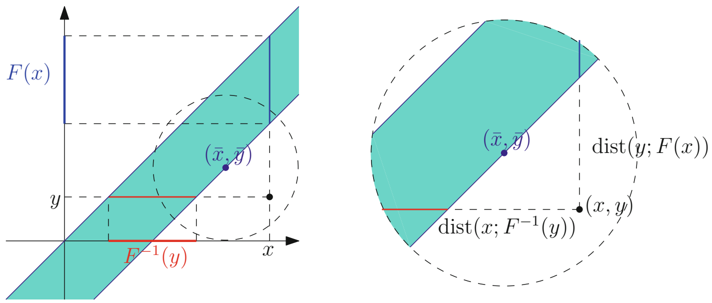


Fig. 3.2 Metric regularity.

(i) *F enjoys the covering property around* $(\bar{x}, \bar{y})$ *if and only if it is metrically regular around this point. In this case we have*

$$\mathrm{cov}\, F(\bar{x}, \bar{y}) = \big(\mathrm{reg}\, F(\bar{x}, \bar{y})\big)^{-1}.$$

(ii) *F is Lipschitz-like around* $(\bar{x}, \bar{y})$ *if and only if the inverse mapping* $F^{-1}\colon \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ *is metrically regular around* $(\bar{y}, \bar{x})$*. In this case we have*

$$\mathrm{lip}\, F(\bar{x}, \bar{y}) = \mathrm{reg}\, F^{-1}(\bar{y}, \bar{x}).$$

Proof. Let us give the proof of these equivalences, which holds also with small changes for appropriate *semilocal* and modifications of the local notions from Definition 3.1 in both finite and infinite dimensions; see Sections 3.4 and 3.5. We split the proof into several steps of their own interest.

Step 1: *Metric regularity in* (3.2) *can be equivalently verified only for vectors* $(x, y) \in U \times V$ *satisfying the estimate* $\mathrm{dist}(y; F(x)) \le \gamma$ *for some* $\gamma > 0$.

To verify this, let us show that for any $\eta, \gamma > 0$ there is $\nu > 0$ such that (3.2) holds for all $x \in \bar{x} + \nu\mathbb{B}$ and $y \in \bar{y} + \nu\mathbb{B}$ provided that it is valid for $x \in \bar{x} + \eta\mathbb{B}$ and $y \in \bar{y} + \eta\mathbb{B}$ with $\mathrm{dist}(y; F(x)) \le \gamma$. Given (μ, η, γ), denote $\nu := \min\{\eta,\ \gamma\mu/(\mu+1)\}$ and check that (3.2) holds for all $x \in \bar{x} + \nu\mathbb{B}$ and $y \in \bar{y} + \nu\mathbb{B}$ with $\mathrm{dist}(y; F(x)) > \gamma$. Observe that $\mathrm{dist}(\bar{x}; F^{-1}(y)) \le \mu\, \mathrm{dist}(y; F(\bar{x}))$ for such x, y due to $\mathrm{dist}(y; F(\bar{x})) \le \|y - \bar{y}\| \le \nu \le \gamma$. This gives us

$$\begin{aligned}\mathrm{dist}\big(x; F^{-1}(y)\big) &\le \mathrm{dist}\big(\bar{x}; F^{-1}(y)\big) + \|x - \bar{x}\| \le \mu\, \mathrm{dist}\big(y; F(\bar{x})\big) + \|x - \bar{x}\| \\ &\le \mu\, \|y - \bar{y}\| + \|x - \bar{x}\| \le \nu(\mu + 1) \le \gamma\mu < \mu\, \mathrm{dist}\big(y; F(x)\big)\end{aligned}$$

by the choice of ν and thus verifies the statement of Step 1.

Step 2: *Metric regularity implies covering with* $\mathrm{cov}\, F(\bar{x}, \bar{y}) \ge (\mathrm{reg}\, F(\bar{x}, \bar{y}))^{-1}$.

Take $\eta, \mu > 0$ such that (3.2) holds for $x \in U := \mathrm{int}\,(\bar{x} + \eta\mathbb{B})$ and $y \in V$ with some V. Define $\nu := \min\{\eta, \mu\}$, $\widetilde{U} := \mathrm{int}\,(\bar{x} + \nu\mathbb{B})$ and then pick

$$v \in \mathrm{int}\,\big(F(x) \cap V + (r/\mu)\mathbb{B}\big) \;\text{ with }\; x + r\mathbb{B} \subset \widetilde{U} \;\text{ for some }\; r > 0.$$

By these constructions and the assumed estimate (3.2), we conclude that $\text{dist}(x; F^{-1}(v)) < r$ for such (x, v, r), which allows us to choose $u \in F^{-1}(v)$ with $u \in \imath(x + r\mathbb{B})$ and $v \in F(u) \subset F(\imath(x + r\mathbb{B}))$. This ensures that

$$\text{int}\big(F(x) \cap V + \mu^{-1} r\mathbb{B}\big) \subset F\big(\text{int}\,(x + r\mathbb{B})\big) \;\text{ whenever }\; x + r\mathbb{B} \subset \widetilde{U}.$$

Taking now an arbitrary small $\varepsilon > 0$ gives us the inclusions

$$F(x) \cap V + (\mu+\varepsilon)^{-1} r\mathbb{B} \subset \text{int}\big(F(x) \cap V + \mu^{-1} r\mathbb{B}\big) \subset F\big(\text{int}\,(x + r\mathbb{B})\big) \subset F(x + r\mathbb{B})$$

when $x + r\mathbb{B} \subset \widetilde{U}$. This justifies the covering property with $\text{cov}\, F(\bar{x}, \bar{y}) \geq (\text{reg}\, F(\bar{x}, \bar{y}))^{-1}$ while the case of $\text{reg}\, F(\bar{x}, \bar{y}) = 0$ is trivial.

Step 3: *Covering implies metric regularity with* $\text{cov}\, F(\bar{x}, \bar{y}) \leq (\text{reg}\, F(\bar{x}, \bar{y}))^{-1}$.

Indeed, by the covering property, we find $\kappa, \eta > 0$ such that

$$F(x) \cap V + \kappa r\mathbb{B} \subset F(x + r\mathbb{B}) \;\text{ whenever }\; x + r\mathbb{B} \subset U := \text{int}\,(\bar{x} + \eta\mathbb{B}),\; r > 0$$

for some neighborhood V of $\bar{y}$. Denote $\nu := \eta/2$, $\widetilde{U} := \text{int}\,(\bar{x} + \nu\mathbb{B})$, $\gamma := \kappa\eta/2$, and show that (3.2) holds for all $x \in \widetilde{U}$ and $y \in V$ with $\text{dist}(y; F(x)) \leq \gamma/2$. This is sufficient for metric regularity due to Step 1. To proceed, fix such a pair (x, y) and consider any number α satisfying $\text{dist}(y; F(x)) < \alpha < \gamma$. Then

$$y \in F(x) \cap V + \kappa r\mathbb{B} \;\text{ and }\; x + r\mathbb{B} \subset U \;\text{ with }\; r := \alpha/\kappa.$$

The covering property ensures the existence of $u \in x + r\mathbb{B}$ with $u \in F^{-1}(y)$, which implies that $\text{dist}(x; F^{-1}(y)) \leq \|x - u\| \leq r = \alpha/\kappa$. Passing now to the limit as $\alpha \downarrow \text{dist}(y; F(x))$ gives us the following estimate:

$$\text{dist}\big(x; F^{-1}(y)\big) \leq \kappa^{-1} \text{dist}\big(y; F(x)\big) \;\text{ for those }\; x \in \widetilde{U} \;\text{ and }\; y \in V$$

that satisfy $\text{dist}(y; F(x)) \leq \gamma$. It justifies the statement of this step with $\text{cov}\, F(\bar{x}, \bar{y}) \leq (\text{reg}\, F(\bar{x}, \bar{y}))^{-1}$ and thus completes the proof of assertion (i).

Step 4: *Lipschitz-like property of* F *implies metric regularity of* F^{-1} *with the estimate* $\text{reg}\, F^{-1}(\bar{y}, \bar{x}) \leq \text{lip}\, F(\bar{x}, \bar{y})$.

To verify it, let $\overline{\ell} := \text{lip}\, F(\bar{x}, \bar{y})$ and for any $\varepsilon > 0$ get

$$F(x) \cap V \subset F(u) + (\overline{\ell} + \varepsilon)\|x - u\|\mathbb{B} \;\text{ whenever }\; x, u \in U$$

with some neighborhoods U of $\bar{x}$ and V of $\bar{y}$. This tells us that

$$\text{dist}\big(y; F(u)\big) \leq (\overline{\ell} + \varepsilon)\|x - u\| \;\text{ if }\; y \in F(x) \cap V \;\text{ and }\; x, u \in U.$$

Take $r > 0$ with $\bar{x} + r\mathbb{B} \subset U$ and observe from the above that

$$\text{dist}\big(y; F(u)\big) \leq (\overline{\ell} + \varepsilon)\, \text{dist}\big(u; F^{-1}(y)\big) \tag{3.5}$$

whenever $u \in \bar{x} + r\mathbb{B}$, $y \in V$, and $F^{-1}(y) \cap (\bar{x} + r\mathbb{B}) \neq \emptyset$. It is easy to check that this ensures the validity of (3.5) for all $u \in \widetilde{U} := \bar{x} + (r/3)\mathbb{B}$ and $y \in V$ satisfying $\text{dist}(u; F^{-1}(y)) \le \gamma := r$. Taking into account the statement proved in Step 1 and the arbitrary choice of $\varepsilon > 0$, we conclude that F^{-1} is metrically regular around $(\bar{y}, \bar{x})$ with $\text{reg}\, F^{-1}(\bar{y}, \bar{x}) \le \text{lip}\, F(\bar{x}, \bar{y})$.

Step 5: *Metric regularity of* F^{-1} *implies the Lipschitz-like property of* F *with the estimate* $\text{lip}\, F(\bar{x}, \bar{y}) \le \text{reg}\, F^{-1}(\bar{y}, \bar{x})$.

Indeed, denoting $\overline{\mu} := \text{reg}\, F^{-1}(\bar{y}, \bar{x})$ and picking any $\varepsilon > 0$ give us

$$\text{dist}\big(y; F(u)\big) \le (\overline{\mu} + \varepsilon)\, \text{dist}\big(u; F^{-1}(y)\big) \ \text{ for all } \ u \in U \ \text{ and } \ y \in V$$

with some neighborhoods U of $\bar{x}$ and V of $\bar{y}$, which yields in turn that

$$F(x) \cap V \subset F(u) + (\overline{\mu} + 2\varepsilon)\|u - x\|\mathbb{B} \ \text{ for all } \ x, u \in U.$$

This verifies the claimed assertion and completes the proof of (ii). △

3.1.2 Coderivative Characterizations of Well-Posedness

The established equivalences show that any necessary and/or sufficient condition and modulus estimates obtained for one of the three well-posedness properties from Definition 3.1 imply the corresponding assertions for the other ones. The following principal result provides *complete characterizations* of these properties for general *closed-graph* (of our standing assumption) multifunctions with calculating the *exact bounds* of their moduli via the *coderivative* (1.15) precisely *at* the point in question.

Theorem 3.3 (Coderivative Criteria for Well-Posedness of Multifunctions). *Let* $F\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ *with* $(\bar{x}, \bar{y}) \in \text{gph}\, F$. *Then we have the following characterizations of the well-posedness properties:*

(i) *F enjoys the covering property around $(\bar{x}, \bar{y})$ if and only if*

$$\ker D^*F(\bar{x}, \bar{y}) = \{0\}. \tag{3.6}$$

In this case the exact covering bound of F around $(\bar{x}, \bar{y})$ is calculated by

$$\text{cov}\, F(\bar{x}, \bar{y}) = \inf\Big\{\|u\| \ \Big|\ u \in D^*F(\bar{x}, \bar{y})(v),\ \|v\| = 1\Big\}. \tag{3.7}$$

(ii) *F is metrically regular around $(\bar{x}, \bar{y})$ if and only if condition (3.6) holds. In this case the exact regularity bound of F at $(\bar{x}, \bar{y})$ is calculated by*

$$\text{reg}\, F(\bar{x}, \bar{y}) = \|D^*F(\bar{x}, \bar{y})^{-1}\| = \|D^*F^{-1}(\bar{y}, \bar{x})\|, \tag{3.8}$$

where the norm of a positively homogeneous mapping is defined in (1.14).

(iii) *F is Lipschitz-like around $(\bar{x}, \bar{y})$ if and only if*

$$D^*F(\bar{x}, \bar{y})(0) = \{0\}. \tag{3.9}$$

In this case the exact Lipschitzian bound of F around $(\bar{x}, \bar{y})$ is calculated by

$$\text{lip}\, F(\bar{x}, \bar{y}) = \|D^* F(\bar{x}, \bar{y})\|. \tag{3.10}$$

Proof. We split the proof into three major steps of their own interest.

Step 1: *If F is Lipschitz-like around $(\bar{x}, \bar{y})$, then*

$$\|D^* F(\bar{x}, \bar{y})\| \le \text{lip}\, F(\bar{x}, \bar{y}) < \infty \tag{3.11}$$

and therefore the injectivity condition (3.9) holds.

To proceed, observe first that (3.11) yields the validity of (3.9) by

$$\|u\| \le \|D^* F(\bar{x}, \bar{y})\| \cdot \|v\| \;\text{ for all }\; u \in D^* F(\bar{x}, \bar{y})(v), \quad v \in \mathbb{R}^m.$$

We verify (3.11) by showing that the Lipschitz-like property of F around $(\bar{x}, \bar{y})$ with modulus $\ell \ge 0$ implies that

$$\|D^* F(\bar{x}, \bar{y})\| \le \ell. \tag{3.12}$$

Assuming this property, pick any $u \in D^* F(\bar{x}, \bar{y})(v)$ and by using the limiting coderivative representation (1.17) find sequences $(x_k, y_k) \xrightarrow{\text{gph}\, F} (\bar{x}, \bar{y})$ and $(u_k, v_k) \to (u, v)$ such that $(u_k, -v_k) \in \widehat{N}((x_k, y_k); \text{gph}\, F)$ for all $k \in \mathbb{N}$. Fix any k sufficiently large, and observe that, due to the aforementioned robustness of the Lipschitz-like property, F is Lipschitz-like around (x_k, y_k) with the same modulus ℓ, which we assume to be positive by taking into account that the case of $\ell = 0$ is trivial. This means that there exists $\eta > 0$ such that

$$F(x) \cap (y_k + \eta\mathbb{B}) \subset F(z) + \ell\|x - z\|\mathbb{B} \;\text{ for all }\; x, z \in x_k + 2\eta\mathbb{B}.$$

Employing definition (1.5) of regular normals, for any $\varepsilon_k > 0$ we can find a positive number $\nu \le \min\{\eta, \ell\eta\}$ such that

$$\langle u_k, z - x_k\rangle - \langle v_k, w - y_k\rangle \le \varepsilon_k\big(\|z - x_k\| + \|w - y_k\|\big) \tag{3.13}$$

whenever $(z, w) \in \text{gph}\, F$ with $\|z - x_k\| \le \nu$ and $\|w - y_k\| \le \nu$. Choose $z \in x + \min\{\nu, \nu\ell^{-1}\}\mathbb{B}$ and observe that $\|z - x_k\| \le \|z - x\| + \|x - x_k\| \le 2\eta$. Using the above Lipschitz-like relationship with $y \in F(x) \cap (y_k + \eta\mathbb{B})$ and the chosen vector z allows us to find $w \in F(z)$ satisfying

$$\|w - y\| \le \ell\|x - z\| \le \ell \min\big\{\nu, \ell^{-1}\nu\big\} = \min\big\{\ell\nu, \nu\big\} \le \nu.$$

If $0 < \ell < 1$, then we have the estimates

$$\|z - x_k\| \le \nu \;\text{ and }\; \|w - y_k\| \le \ell\nu,$$

which imply by (3.13) that $\nu\|u_k\| \le \ell\nu\|v_k\| + \varepsilon_k(\nu + \ell\nu)$ and hence

$$\|u_k\| \le \ell\|v_k\| + \varepsilon_k(1+\ell).$$

In the remaining case where $\ell > 1$ we have $\|z - x_k\| \le \nu\ell^{-1}$ and $\|w - y_k\| \le \nu$, which imply in turn that $\nu\ell^{-1}\|u_k\| \le \nu\|v_k\| + \varepsilon_k(\nu + \ell^{-1}\nu)$ and hence

$$\|u_k\| \le \ell\|v_k\| + \varepsilon_k(1+\ell).$$

Passing now to the limit as $k \to \infty$ with $\varepsilon_k \downarrow 0$ and taking into account the second representation of the normal cone in (1.7), we arrive at (3.12) and thus verify the assertion claimed in this step.

Step 2: *The kernel condition* (3.6) *ensures the covering property of* F *around* $(\bar{x}, \bar{y})$ *with the exact bound estimate*

$$\operatorname{cov} F(\bar{x},\bar{y}) \ge \inf\Big\{\|u\| \;\Big|\; u \in D^*F(\bar{x},\bar{y})(v),\ \|v\| = 1\Big\} > 0. \tag{3.14}$$

Denote $a(F,\bar{x},\bar{y}) := \inf\big\{\|u\| \,\big|\, u \in D^*F(\bar{x},\bar{y})(v),\ \ \|v\| = 1\big\}$ in (3.14) and observe that condition (3.6) yields $a(F,\bar{y},\bar{x}) > 0$. Indeed, assuming the contrary brings us to a contradiction due to the robustness property of the normal cone. Thus to prove the statements of this step, it suffices to show that every $0 < \kappa < a(F,\bar{x},\bar{y})$ is a covering modulus of F around $(\bar{x},\bar{y})$. Supposing that it is not true for some fixed $0 < \kappa < a(F,\bar{x},\bar{y})$, the negation of (3.1) gives us sequences $x_k \to \bar{x}$, $y_k \to \bar{y}$, and $r_k \downarrow 0$ as well as $z_k \in \mathbb{R}^m$ satisfying

$$y_k \in F(x_k),\ \|z_k - y_k\| \le \kappa r_k,\ z_k \notin F(x) \text{ for all } x \in B_{r_k}(x_k). \tag{3.15}$$

Fix $k \in \mathbb{N}$ and define the set E_k and the function $\theta_k\colon \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}_+$ by

$$E_k := \big(\operatorname{gph} F\big) \cap \big((x_k, y_k) + r_k\mathbb{B}\big) \ \text{ and } \ \theta_k(x,y) := \|x\| + r_k\|y\|,$$

where $\mathbb{B}$ stands for the closed unit ball of $\mathbb{R}^n \times \mathbb{R}^m$. Consider now the l.s.c. function $\varphi_k\colon \mathbb{R}^n \times \mathbb{R}^m \to \overline{\mathbb{R}}$ with the (extended) nonnegative values

$$\varphi_k(x,y) := \|y - z_k\| + \delta\big((x,y); E_k\big), \quad (x,y) \in \mathbb{R}^n \times \mathbb{R}^m,$$

and apply to it Theorem 2.12 with $\varepsilon_k := \kappa r_k$, $\lambda_k := r_k - r_k^2$, the initial point (x_k, y_k), and the function θ_k defined above. Taking into account that $\varphi_k(x_k, y_k) \le \varepsilon_k$ by (3.15) and the structures of φ_k and E_k, we find a pair $(\bar{x}_k, \bar{y}_k) \in \operatorname{gph} F$ with $\|(\bar{x}_k, \bar{y}_k) - (x_k, y_k)\| \le r_k$ such that the function

$$\psi_k(x,y) := \|y - z_k\| + \frac{\kappa}{1 - r_k}\Big(\|x - \bar{x}_k\| + r_k\|y - \bar{y}_k\|\Big) + \delta\big((x,y); \operatorname{gph} F\big)$$

attains its unconditional local minimum on $\mathbb{R}^n \times \mathbb{R}^m$ at $(\bar{x}_k, \bar{y}_k)$. Note that ψ_k can be treated as the sum of two functions, one of which is convex and Lipschitz continuous, while the other is l.s.c. around $(\bar{x}_k, \bar{y}_k)$. Applying now Corollary 2.20 to this sum and using subdifferentiation of the (convex) norm function at zero and nonzero points (see Exercise 1.66) together with the condition $z_k \notin F(\bar{x}_k)$ by (3.15) give us

u_k and v_k satisfying

$$u_k \in D^*F(\bar{x}_k, \bar{y}_k)(v_k) \;\text{ with }\; \|u_k\| \le \frac{\kappa}{1 - r_k} \;\text{ and }\; \|v_k\| \to 1 \;\text{ as }\; k \to \infty.$$

Thus we get (u, v) such that $(u_k, v_k) \to (u, v)$ along a subsequence and that

$$\|u\| \le \kappa \;\text{ with }\; u \in D^*F(\bar{x}, \bar{y})(v),$$

where the latter is due to the robustness property from Proposition 1.3. The obtained contradiction justifies the assertions of Step 2.

Step 3: *Completing the proof of the theorem.*

It follows from the results established in Steps 1 and 2 due to the equivalences of Theorem 3.2 and the relationship

$$\big[D^*F^{-1}F(\bar{y}, \bar{x})(0) = \{0\}\big] \Longleftrightarrow \big[\ker D^*F(\bar{x}, \bar{y}) = \{0\}\big]$$

between the coderivatives of F and its inverse as well as the one

$$1\Big/\|H^{-1}\| = \inf\big\{\|y\| \;\big|\; y \in H(x),\; \|x\| = 1\big\}$$

holding for any positively homogeneous multifunction. △

Before presenting several consequences of Theorem 3.3 in this section (and more later on), let us draw the reader's attention to some principal issues concerning the coderivative criterion for the Lipschitz-like property.

Remark 3.4 (Discussions on the Coderivative Characterization of Lipschitzian Behavior). Observe the following:

(i) The approach of classical analysis is *from continuity to differentiability*, where smooth functions form a subclass of Lipschitz continuous ones. Here we have the opposite direction for nonsmooth and set-valued mappings: *from generalized differentiability to Lipschitz continuity*, where the coderivative allows us to fully characterize Lipschitzian behavior of multifunctions.

(ii) Lipschitz continuity in both classical and set-valued frameworks can be viewed as *continuity with linear rate*, where the rate of linearity is crucial for characterizing such continuity as well as the equivalent notions of linear openness and (first-order) metric regularity in Theorem 3.3.

(iii) Replacing in the coderivative criterion (3.9) the basic normal cone to the graph of F by Clarke's *convexification* (1.60) leads us to the condition

$$\big[(u, 0) \in \overline{N}\big((\bar{x}, \bar{y}); \operatorname{gph} F\big)\big] \Longrightarrow u = 0, \tag{3.16}$$

which is sufficient for the Lipschitz-like property of F around $(\bar{x}, \bar{y})$ but *far removed* from the necessity; see [512, 513] for further details. Amazingly it *never holds* even in the trivial case where the mapping $F = f : \mathbb{R}^n \to \mathbb{R}^n$ is single-valued and locally Lipschitzian while *nonsmooth* around $\bar{x}$. It follows from Rockafellar's theorem on the subspace property of the convexified normal cone; see Exercise 1.46(ii). This phenomenon is also valid for Lipschitzian manifolds (or graphically Lipschitzian set-valued mappings) that are not "strictly smooth" at the reference point; see more discussions in Section 1.5.

3.1.3 Characterizations in Special Cases

This subsection concerns deriving from the coderivative characterizations of Theorem 3.3 for general closed-graph multifunctions some useful consequences in remarkable special cases. We start with characterizing the classical (Hausdorff) *local Lipschitz continuity* of multifunctions meaning that (3.3) holds with $V = \mathbb{R}^m$. Recall that the (local) uniform boundedness of set-valued mappings is defined in Subsection 1.2.1.

Corollary 3.5 (Coderivative Criterion for Local Lipschitz Continuity of Set-Valued Mappings). *Let* $F\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ *be locally bounded around* $\bar{x} \in \operatorname{dom} F$, *for any* $\bar{y} \in F(\bar{x})$. *Then the mapping* F *is locally Lipschitzian around* $\bar{x}$ *if and only if we have the condition*

$$D^*F(\bar{x}, \bar{y})(0) = \{0\} \text{ for all } \bar{y} \in F(\bar{x}).$$

In this case the exact Lipschitzian bound of F *around* $\bar{x}$ *is calculated by*

$$\operatorname{lip} F(\bar{x}) = \max\left\{\|D^*F(\bar{x}, \bar{y})\| \;\middle|\; \bar{y} \in F(\bar{x})\right\}.$$

Proof. Observe that, under the assumptions made, the local Lipschitzian property of F around $\bar{x}$ is equivalent to its Lipschitz-like property around $(\bar{x}, \bar{y})$ for every $\bar{y} \in F(\bar{x})$. This follows from the classical fact that any open covering of a closed and bounded set in finite dimensions can be reduced to a finite subcovering. It implies also that

$$\operatorname{lip} F(\bar{x}) = \max\left\{\operatorname{lip} F(\bar{x}, \bar{y}) \;\middle|\; \bar{y} \in F(\bar{x})\right\},$$

where the maximum is realized due to the upper semicontinuity of $\operatorname{lip} F(\cdot, \cdot)$ on the graph of F. This allows us to deduce the claimed statements from the corresponding ones in Theorem 3.3(iii). △

In the next corollary we present characterizations of metric regularity and covering for set-valued mappings with *convex graphs*, where the coderivative calculation allows us to describe the criteria and exact bound formulas entirely in terms of the range and graph of the given mapping.

Corollary 3.6 (Metric Regularity and Covering of Convex-Graph Multifunctions). *Assume that the graph of* $F\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ *is convex and pick some* $\bar{y} \in \operatorname{rge} F$. *Then the validity of both metric regularity and covering properties of* F *around* $(\bar{x}, \bar{y})$ *for any* $\bar{x} \in F^{-1}(\bar{y})$ *is equivalent to* $\bar{y} \in \operatorname{int}(\operatorname{rge} F)$. *In this case the corresponding exact bounds are calculated by*

$$\operatorname{reg} F(\bar{x}, \bar{y}) = \max_{\|u\| \le 1} \Big\{ \|v\| \Big| \; \langle u, x - \bar{x} \rangle \le \langle v, y - \bar{y} \rangle \;\; \textit{for all} \;\; (x, y) \in \operatorname{gph} F \Big\},$$

$$\operatorname{cov} F(\bar{x}, \bar{y}) = \min_{\|v\| = 1} \Big\{ \|u\| \Big| \; \langle u, x - \bar{x} \rangle \le \langle v, y - \bar{y} \rangle \;\; \textit{for all} \;\; (x, y) \in \operatorname{gph} F \Big\}.$$

Proof. Follows from Theorem 3.3(i,ii) due to the coderivative representation for convex-graph mappings in Proposition 1.13. △

We conclude this section with consequences of Theorem 3.3 applied to two classes of single-valued mappings. The first class contains locally Lipschitzian ones for which the criteria and exact bounds for metric regularity and covering are expressed via basic subgradients of the corresponding scalarization.

Corollary 3.7 (Metric Regularity and Covering of Single-Valued Locally Lipschitzian Mappings). *Let $f: \mathbb{R}^n \to \mathbb{R}^m$ be locally Lipschitz around $\bar{x}$. Then f is metrically regular and enjoys the covering property around this point if and only if we have the implication*

$$\big[0 \in \partial \langle v, f \rangle (\bar{x})\big] \Longrightarrow v = 0.$$

In this case the exact regularity and covering bounds are calculated by

$$\operatorname{reg} f(\bar{x}) = \max \Big\{ \|v\| \;\Big|\; u \in \partial \langle v, f \rangle (\bar{x}), \;\; \|u\| \le 1 \Big\},$$

$$\operatorname{cov} f(\bar{x}) = \min \Big\{ \|u\| \;\Big|\; u \in \partial \langle v, f \rangle (\bar{x}), \;\; \|v\| = 1 \Big\}.$$

Proof. Follows from Theorem 3.3(i,ii) due to the coderivative scalarization of Theorem 1.32 for locally Lipschitzian mappings and the norm definition (1.14). The maximum and minimum in the exact bound formulas are realized due to the robustness property of the basic subdifferential. △

The last corollary of Theorem 3.3 presented here provides complete characterizations of metric regularity and covering for smooth mappings.

Corollary 3.8 (Metric Regularity and Covering of Smooth Mappings). *Let $f: \mathbb{R}^n \to \mathbb{R}^m$ with $m \le n$ be smooth around $\bar{x}$. Then it is metrically regular and enjoys the covering property around this point if and only if its Jacobian matrix $\nabla f(\bar{x})$ has full rank:*

$$\operatorname{rank} \nabla f(\bar{x}) = m, \quad \textit{or equivalently} \;\; \nabla f(\bar{x}) \mathbb{R}^n = \mathbb{R}^m. \tag{3.17}$$

In this case the exact regularity and covering bounds are calculated by

$$\operatorname{reg} f(\bar{x}) = \big\| \big(\nabla f(\bar{x})^* \big)^{-1} \big\|, \;\; \operatorname{cov} f(\bar{x}) = \min \big\{ \|\nabla f(\bar{x})^* v\| \;\big|\; \|v\| = 1 \big\}. \tag{3.18}$$

If furthermore $m = n$, then we have

$$\operatorname{cov} f(\bar{x}) = \|\nabla f(\bar{x})^{-1}\|^{-1}.$$

Proof. Follows from Theorem 3.3, or from Corollary 3.7, due to the coderivative representation for smooth mappings by Proposition 1.12. Let us also present another proof of the *sufficiency* of the surjectivity condition (3.18) for the covering/metric regularity property of f around $\bar{x}$, which works for strictly differentiable mappings between arbitrary Banach spaces. Put $A := \nabla f(\bar{x})$. The open mapping theorem implies by the surjectivity of A (see Exercise 1.53) that for any y from the image space there is $x \in A^{-1}(y)$ satisfying

$$\|x\| \le \mu\|y\| \;\text{ with }\; \mu^{-1} = \inf\Big\{\|A^*v\| \;\Big|\; \|v\| = 1\Big\}. \tag{3.19}$$

Using the strict differentiability of f at $\bar{x}$, for every $\gamma \in (0, \mu^{-1})$, we find a neighborhood U of $\bar{x}$ such that

$$\|f(x_1) - f(x_2) - A(x_1 - x_2)\| \le \gamma\|x_1 - x_2\| \;\text{ for all }\; x_1, x_2 \in U.$$

Our aim now is to justify the inclusion

$$f(\widehat{x}) + (\mu^{-1} - \gamma)r\mathbb{B} \subset f(\widehat{x} + r\mathbb{B}) \;\text{ for }\; \widehat{x} + r\mathbb{B} \subset U,\; r > 0 \tag{3.20}$$

meaning that f enjoys the covering property around $\bar{x}$ with modulus $\kappa = \mu^{-1} - \gamma$. Since $\gamma > 0$ can be taken arbitrarily small, we get from (3.20) that

$$\operatorname{cov} f(\bar{x}) \ge \mu^{-1} = \inf\Big\{\|\nabla f(\bar{x})^* v\| \;\Big|\; \|v\| = 1\Big\},$$

which verifies the covering property of f with the equality in (3.18) by taking into account that the opposite inequality follows directly from Step 1 in the proof of Theorem 3.3 and the equivalences of Theorem 3.2 held in any Banach spaces. Note that we cannot generally claim that the minimum is realized in the exact bound formula of (3.18) in infinite dimensions.

Thus it remains to verify inclusion (3.20), where we can obviously take $\widehat{x} = 0$ and $f(\widehat{x}) = 0$ without loss of generality. The latter means that for every $y \in (\mu^{-1} - \gamma)r\mathbb{B}$ the equation $y = f(x)$ has a solution $x \in r\mathbb{B} \subset U$.

To justify (3.20) via the above claim, fix $y \in Y$ with $\|y\| \le (\mu^{-1} - \gamma)r$ and construct the desired solution x as the limit of a sequence $\{x_k\}$, $k = 0, 1, \ldots$, recurrently defined in the following way. Starting with $x_0 := 0$, we use (3.19) to construct x_k by the iterative procedure of Newton's type, which is known as the *Lyusternik-Graves iterative process* (see Section 3.5):

$$Ax_k = y - f(x_{k-1}) + Ax_{k-1} \;\text{ with }\; \|x_k - x_{k-1}\| \le \mu\,\|y - f(x_{k-1})\|$$

for all $k \in \mathbb{N}$. It follows from the above construction that

$$\|x_{k+1} - x_k\| \le \mu(\mu\gamma)^k\,\|y\| \quad \text{and}$$

$$\|x_k\| \le \sum_{j=1}^{k} \|x_j - x_{j-1}\| \le \mu \, \|y\| \sum_{j=1}^{k} (\mu\gamma)^{j-1}$$

$$\le \mu \, \|y\| \big/ (1 - \mu\gamma) = \|y\| \big/ (\mu^{-1} - \gamma) \le r$$

for every $k \in \mathbb{N}$. Therefore $\{x_k\}$ is a Cauchy sequence that converges to some x with $\|x\| \le r$. Passing to the limit in these iterations as $k \to \infty$, we arrive at $y = f(x)$ and thus complete the alternative proof of the sufficiency. △

3.2 Coderivative Calculus

This section contains basic calculus rules for the coderivative (1.15) of set-valued mappings satisfying our standing closed-graph (i.e., continuity in the single-valued case) assumption. Although the results below are given for mappings between finite-dimensional spaces, it is more convenient here to use the *star-notation* (x^*, y^*, etc.) to signify *dual-space* variables; see also Sections 3.5 and 3.4 for infinite-dimensional extensions.

3.2.1 Coderivative Sum Rules

We start with sum rules, which invokes (in one part) the inner semicontinuity notion for set-valued mappings at graph points defined and discussed in Exercise 2.46. Observe that a multifunction F is inner semicontinuous at $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ if it is Lipschitz-like around this point.

Given two closed-graph multifunctions $F_1, F_2 \colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, consider the auxiliary mapping $S \colon \mathbb{R}^n \times \mathbb{R}^m \rightrightarrows \mathbb{R}^{2m}$ given by

$$S(x, y) := \big\{(y_1, y_2) \in \mathbb{R}^m \times \mathbb{R}^m \big|\ y_1 \in F_1(x),\ y_2 \in F_2(x),\ y = y_1 + y_2\big\}, \quad (3.21)$$

and derive now two related while independent coderivative sum rules.

Theorem 3.9 (General Sum Rules for Coderivatives). *Let $F_i \colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ for $i = 1, 2$, and let $(\bar{x}, \bar{y}) \in \operatorname{gph}(F_1 + F_2)$. The following assertions hold:*

(i) *Fix $(\bar{y}_1, \bar{y}_2) \in S(\bar{x}, \bar{y})$ from (3.21), and suppose that this mapping is inner semicontinuous at $(\bar{x}, \bar{y}, \bar{y}_1, \bar{y}_2)$ and that the qualification condition*

$$D^* F_1(\bar{x}, \bar{y}_1)(0) \cap \big(-D^* F_2(\bar{x}, \bar{y}_2)(0)\big) = \{0\} \quad (3.22)$$

is satisfied. Then for all $y^* \in \mathbb{R}^m$, we have the inclusion

$$D^*(F_1 + F_2)(\bar{x}, \bar{y})(y^*) \subset D^* F_1(\bar{x}, \bar{y}_1)(y^*) + D^* F_2(\bar{x}, \bar{y}_2)(y^*). \quad (3.23)$$

If one of the mappings F_i, say F_1, is single-valued and continuously differentiable around $\bar{x}$, then (3.23) becomes the equality

$$D^*(F_1+F_2)(\bar{x},\bar{y})(y^*)=\nabla F_1(\bar{x})^*y^*+D^*F_2\big(\bar{x},\bar{y}-F_1(\bar{x})\big)(y^*) \tag{3.24}$$

for all $y^*\in\mathbb{R}^m$ without any other assumptions.

(ii) *Suppose that the mapping S in (3.21) is locally bounded around $(\bar{x},\bar{y})$ and that the assumptions in* (i) *ensuring (3.23) hold for every pair $(\bar{y}_1,\bar{y}_2)\in S(\bar{x},\bar{y})$. Then for all $y^*\in\mathbb{R}^m$, we have the inclusion*

$$D^*(F_1+F_2)(\bar{x},\bar{y})(y^*)\subset\bigcup_{(\bar{y}_1,\bar{y}_2)\in S(\bar{x},\bar{y})}\Big[D^*F_1(\bar{x},\bar{y}_1)(y^*)+D^*F_2(\bar{x},\bar{y}_2)(y^*)\Big].$$

Proof. First we justify inclusion (3.23) in assertion (i). Take any (x^*,y^*) with $x^*\in D^*(F_1+F_2)(\bar{x},\bar{y})(y^*)$ and find by definition (1.15) and the first representation in (1.7) sequences $(x_k,y_k)\in\operatorname{gph}(F_1+F_2)$ and $(x_k^*,-y_k^*)\in\widehat{N}_{\varepsilon_k}(x_k,y_k);\operatorname{gph}(F_1+F_2))$ such that $(x_k,y_k)\to(\bar{x},\bar{y})$, $x_k^*\to x^*$, and $y_k^*\to y^*$ as $k\to\infty$. Due to the assumed inner semicontinuity of the mapping S from (3.21) at $(\bar{x},\bar{y},\bar{y}_1,\bar{y}_2)$, we get a sequence $(y_{1k},y_{2k})\to(\bar{y}_1,\bar{y}_2)$ with $(y_{1k},y_{2k})\in S(x_k,y_k)$ for all $k\in\mathbb{N}$. Define further the sets

$$\Omega_i:=\Big\{(x,y_1,y_2)\in\mathbb{R}^n\times\mathbb{R}^m\times\mathbb{R}^m\Big|\;(x,y_i)\in\operatorname{gph}F_i\Big\},\quad i=1,2,$$

that are locally closed around $(\bar{x},\bar{y}_1,\bar{y}_2)$ with $(x_k,y_{1k},y_{2k})\in\Omega_1\cap\Omega_2$. It is easy to check that $(x_k^*,-y_k^*,-y_k^*)\in\widehat{N}((x_k,y_{1k},y_{2k});\Omega_1\cap\Omega_2)$ for all $k\in\mathbb{N}$, which tells us by passing to the limit as $k\to\infty$ that

$$(x^*,-y^*,-y^*)\in N\big((\bar{x},\bar{y}_1,\bar{y}_2);\Omega_1\cap\Omega_2\big).$$

Now we apply Theorem 2.16 to the above set intersection observing that the normal qualification condition (2.26) for these sets reduces to (3.22). The intersection rule (2.27) ensures in this setting the existence of

$$(x_1^*,-y_1^*)\in N\big((\bar{x},\bar{y}_1);\operatorname{gph}F_1\big)\;\text{ and }\;(x_2^*,-y_2^*)\in N\big((\bar{x},\bar{y}_2);\operatorname{gph}F_2\big)$$

such that $(x^*,-y^*,-y^*)=(x_1^*,-y_1^*,0)+(x_2^*,0,-y_2^*)$. This readily shows that $x^*=x_1^*+x_2^*$ with $x_i^*\in D^*F_i(\bar{x},\bar{y}_i)(y^*)$, $i=1,2$ and thus justifies (3.23).

To finish the proof of assertion (i), it remains to verify equality (3.24) if F_1 is single-valued and smooth around $\bar{x}$. In this case we have $D^*F_1(\bar{x})(y^*)=\{\nabla F_1(\bar{x})^*y^*\}$. Hence the qualification condition (3.22) holds automatically and the mapping S in (3.21) is obviously locally bounded around $(\bar{x},\bar{y})$. The inclusion "$\subset$" in (3.24) follows directly from (3.22) and Proposition 1.12. Applying it to the sum $F_2=(F_1+F_2)+(-F_1)$, we arrive at the opposite inclusion in (3.24) and thus justify the claimed sum rule equality.

To verify (ii), observe that the local boundedness of S around $(\bar{x},\bar{y})$ implies the existence of a subsequence of $(y_{1k},y_{2k})\in S(x_k,y_k)$, which converges to some $(\bar{y}_1,\bar{y}_2)$. It follows from the standing closed-graph assumptions imposed on F_i, $i=1,2$, that $(\bar{y}_1,\bar{y}_2)\in S(\bar{x},\bar{y})$. Then we proceed as in the proof of assertion (i) and complete the proof of the theorem. △

The following result reveals an important consequence of Theorem 3.9, which tells us that the qualification condition (3.22) holds automatically if one of the mappings F_i is *Lipschitz-like* around the corresponding point. It is due to the *coderivative characterization* of this property established in Section 3.1.

Corollary 3.10 (Coderivative Sum Rules for Lipschitz-Like Multifunctions). *Suppose in the framework of Theorem 3.9*(i) *that one of the mappings F_i is Lipschitz-like around the corresponding point $(\bar{x}, \bar{y}_i)$, $i = 1, 2$. Then the sum rule inclusion* (3.23) *holds. If in the setting of Theorem 3.9*(ii) *the Lipschitz-like property is imposed on one of F_i around $(\bar{x}, \bar{y}_i)$ for every $(\bar{y}_1, \bar{y}_2) \in S(\bar{x}, \bar{y})$, then we have the sum rule inclusion therein.*

Proof. The validity of the qualification condition (3.22) under the imposed Lipschitz-like assumptions follows from Theorem 3.3(iii). △

3.2.2 Coderivative Chain Rules

Our next theorem unifies several coderivative chain rules providing independent results of the inclusion and equality types in large generality. The *composition* $(F \circ G)\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^q$ of two set-valued mappings $F\colon \mathbb{R}^m \rightrightarrows \mathbb{R}^q$ and $G\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is naturally defined by

$$\big(F \circ G\big)(x) := \bigcup_{y \in G(x)} F(y) = \Big\{z \in \mathbb{R}^q \Big|\ \exists\, y \in G(x) \ \text{ with } \ z \in F(y)\Big\}, \quad x \in \mathbb{R}^n.$$

Theorem 3.11 (General Coderivative Chain Rules). *Given $F\colon \mathbb{R}^m \rightrightarrows \mathbb{R}^q$ and $G\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, let $\bar{x} \in (F \circ G)(\bar{x})$, and consider the mapping*

$$S(x, z) := G(x) \cap F^{-1}(z) = \big\{y \in G(x) \big|\ z \in F(y)\big\} \tag{3.25}$$

for all $(x, z) \in \mathbb{R}^n \times \mathbb{R}^q$. The following assertions hold:

(i) *Fix $\bar{y} \in S(\bar{x}, \bar{z})$ in* (3.25) *and suppose that the mapping S is inner semicontinuous at $(\bar{x}, \bar{z}, \bar{y})$ and that the qualification condition*

$$D^* F(\bar{y}, \bar{z})(0) \cap \ker D^* G(\bar{x}, \bar{y}) = \{0\} \tag{3.26}$$

is satisfied. Then for all $z^* \in \mathbb{R}^q$ we have the inclusion

$$D^*(F \circ G)(\bar{x}, \bar{z})(z^*) \subset D^* G(\bar{x}, \bar{y}) \circ D^* F(\bar{y}, \bar{z})(z^*). \tag{3.27}$$

(ii) *Suppose that the mapping S in* (3.25) *is locally bounded around $(\bar{x}, \bar{z})$ and that the qualification condition* (3.26) *holds for every $\bar{y} \in S(\bar{x}, \bar{z})$. Then for all $z^* \in \mathbb{R}^q$ we have the inclusion*

$$D^*(F \circ G)(\bar{x}, \bar{z})(z^*) \subset \bigcup_{\bar{y} \in S(\bar{x}, \bar{z})} \Big[D^* G(\bar{x}, \bar{y}) \circ D^* F(\bar{y}, \bar{z})(z^*)\Big].$$

(iii) *Let $G = g$ be single-valued and continuously differentiable around $\bar{x}$ with $\bar{y} = g(\bar{x})$. Then we have the equality*

$$D^*(F \circ g)(\bar{x}, \bar{z})(z^*) = \nabla g(\bar{x})^* \circ D^*F(\bar{y}, \bar{z})(z^*), \quad z^* \in \mathbb{R}^q, \tag{3.28}$$

when either the Jacobian matrix $\nabla g(\bar{x})$ has full rank or the qualification condition (3.26) is satisfied and F is graphically regular at $(\bar{y}, \bar{z})$. In the latter case the composition $F \circ g$ is graphically regular at $(\bar{x}, \bar{z})$.

Proof. To verify assertion (i), define $\Phi\colon \mathbb{R}^n \times \mathbb{R}^m \rightrightarrows \mathbb{R}^q$ by

$$\Phi(x, y) := F(y) + \Delta\big((x, y); \operatorname{gph} G\big) \ \text{ for } \ (x, y) \in \mathbb{R}^n \times \mathbb{R}^m, \tag{3.29}$$

where $\Delta(\cdot; \operatorname{gph} G)$ is the *indicator mapping* of the set $\operatorname{gph} G$ relative to $\mathbb{R}^q$ considered in Exercise 1.59. It follows from the sum rule of Theorem 3.9(ii) and the result of the aforementioned exercise applied to the mapping Φ in (3.29) that for any $z^* \in \mathbb{R}^q$ we have the inclusion

$$D^*\Phi(\bar{x}, \bar{y}, \bar{z})(z^*) \subset \big(0, D^*F(\bar{y})(z^*)\big) + N\big((\bar{x}, \bar{y}); \operatorname{gph} G\big) \tag{3.30}$$

under the validity of the qualification condition (3.26). On the other hand, it can be deduced from the construction of Φ in (3.29) and the first representation of the normal cone in Theorem 1.6 that

$$D^*(F \circ G)(\bar{x}, \bar{z})(z^*) \subset \big\{x^* \in \mathbb{R}^n \big|\ (x^*, 0) \in D^*\Phi(\bar{x}, \bar{y}, \bar{z})(z^*)\big\} \tag{3.31}$$

for all $z^* \in \mathbb{R}^q$ under the assumed inner semicontinuity of the mapping S at $(\bar{x}, \bar{z}, \bar{y})$. Combining (3.30) and (3.31) gives us the chain rule inclusion (3.28).

The proof of (ii) is similar to (i). Now we justify assertion (iii), where (3.28) is the equality version of (3.27) for smooth inner mappings. Let us start with showing that inclusion (3.31) holds as *equality* provided that g is locally *Lipschitzian* around $\bar{x}$ with some modulus $\ell \geq 0$. Indeed, take any (x^*, z^*) with $(x^*, 0) \in D^*\Phi(\bar{x}, g(\bar{x}), \bar{z})(z^*)$ and by (1.17) find sequences $(x_k, z_k) \to (\bar{x}, \bar{z})$ and $(x_k^*, y_k^*, z_k^*) \to (x^*, 0, z^*)$ such that $z_k \in F(g(x_k))$ and

$$\limsup_{\substack{x \to x_k,\ z \to z_k \\ z \in F(g(x))}} \frac{\langle (x_k^*, y_k^*, -z_k^*),\ (x, g(x), z) - (x_k, g(x_k), z_k)\rangle}{\|(x, g(x), z) - (x_k, g(x_k), z_k)\|} \leq 0$$

for all $k \in \mathbb{N}$. It implies by the local Lipschitz continuity of g that

$$\limsup_{\substack{x \to x_k,\ z \to z_k \\ z \in F(g(x))}} \frac{\langle x_k^*, x - x_k\rangle - \langle z_k^*, z - z_k\rangle}{\|(x, z) - (x_k, z_k)\|} \leq (\ell + 1)\|y_k^*\| \downarrow 0,$$

and thus $(0, x^*) \in D^*(F \circ g)(\bar{x}, \bar{z})(z^*)$ by (1.15) and the second representation of basic normals in Theorem 1.6. This verifies the equality in (3.31).

It is straightforward to observe from the definitions that we always have

$$\widehat{D}^*\Phi(\bar{x},\bar{y},\bar{z})(z^*)\subset\big(0,\widehat{D}^*F(\bar{y})(z^*)\big)+\widehat{N}\big((\bar{x},\bar{y});\operatorname{gph}G\big),\quad z^*\in\mathbb{R}^q,$$

and so (3.30) becomes an equality if both F and G are graphically regular at the corresponding points. When $G=g$ is single-valued, the graphical regularity for g holds with $D^*g(\bar{x})(y^*)=\{\nabla g(\bar{x})^*y^*\}$ if g is of class $\mathcal{C}^1$ around $\bar{x}$. (It actually reduces to the strict differentiability of g at this point; see [522, Theorem 1.46] and Exercise 1.60(iii) above.) Combining this with the equality in (3.31) justifies the equality and regularity statement in (iii) under the graphical regularity assumption imposed on F.

It remains to verify equality (3.28) when $\nabla g(\bar{y})$ has full rank; note that the graphical regularity of $F\circ g$ is not claimed in this case. Let I be the identity operator on $\mathbb{R}^q$. Then $(g,I)\colon\mathbb{R}^n\times\mathbb{R}^q\to\mathbb{R}^m\times\mathbb{R}^q$ is of class $\mathcal{C}^1$ around $(\bar{x},\bar{z})$ with full rank of $\nabla(g,I)(\bar{x},\bar{z})$. It is easy to see that $(g,I)^{-1}(\operatorname{gph}F)=\operatorname{gph}(F\circ g)$. Thus the chain rule (3.28) follows from representation of normals to inverse images of smooth mappings given in Exercise 1.54(ii). △

There are a great many useful consequences of Theorem 3.11. We present some of them in this and the next sections, while others are formulated as exercises below. Let us start with efficient conditions ensuring the validity of the underlying qualification condition (3.26), and hence the coderivative chain rules, due to the well-posedness characterizations of Section 3.1.

Corollary 3.12 (Coderivative Chain Rules for Lipschitz-Like and Metrically Regular Mappings). *Fix $\bar{z}\in(F\circ G)(\bar{x})$ and $\bar{y}\in G(\bar{x})\cap F^{-1}(\bar{z})$ and suppose that the mapping S in (3.25) is inner semicontinuous at $(\bar{x},\bar{z},\bar{y})$. Then the coderivative chain rule (3.27) holds if either F is Lipschitz-like around $(\bar{y},\bar{z})$ or G is metrically regular around $(\bar{x},\bar{y})$. Alternatively, the local boundedness of S around $(\bar{x},\bar{z})$ and the validity of either the Lipschitz-like property of F around $(\bar{y},\bar{z})$ or metric regularity of G around $(\bar{x},\bar{y})$ for every $\bar{y}\in S(\bar{x},\bar{z})$ ensure the chain rule inclusion in Theorem 3.11*(ii).

Proof. Follows from Theorem 3.11 due to the coderivative characterizations of the well-posedness properties in Theorem 3.3. △

The following corollary of Theorem 3.11 allows us to evaluate normals to inverse images of sets under set-valued mappings. For brevity we consider only the case corresponding to the local boundedness of S in Theorem 3.11.

Corollary 3.13 (Normals to Inverse Images). *Given $G\colon\mathbb{R}^n\rightrightarrows\mathbb{R}^m$ and $\Theta\subset\mathbb{R}^m$ with $\bar{x}\in G^{-1}(\Theta)$, suppose that the mapping $x\mapsto G(x)\cap\Theta$ is locally bounded around $\bar{x}$ and that the qualification condition*

$$N(\bar{y};\Theta)\cap\ker D^*G(\bar{x},\bar{y})=\{0\}\ \textit{for all}\ \ \bar{y}\in G(\bar{x})\cap\Theta \tag{3.32}$$

is satisfied, which is automatic when G is metrically regular around $(\bar{x},\bar{y})$. Then we have the inclusion

$$N\big(\bar{x};G^{-1}(\Theta)\big)\subset\bigcup\Big[D^*G(\bar{x},\bar{y})(y^*)\Big|\ y^*\in N(\bar{y};\Theta),\ \ \bar{y}\in G(\bar{x})\cap\Theta\Big],$$

which becomes an equality when $G = g$ is single-valued and continuously differentiable around $\bar{x}$ and either $\nabla g(\bar{x})$ has full rank or Θ is normally regular at $\bar{y}$. In the latter case, the set $g^{-1}(\Theta)$ is normally regular at $\bar{x}$.

Proof. Observe the composite representation of the inverse image

$$\Delta\big(x; G^{-1}(\Theta)\big) = \big(F \circ G)(x) \ \text{ with } \ F(y) := \Delta(y; \Theta)$$

via the indicator mappings of the sets in question relative to any space $\mathbb{R}^q$ as defined in (1.59). Then the claimed results follow directly from Theorem 3.11 applied to this composition. Note that the case where $\nabla g(\bar{x})$ has full rank in this corollary recovers the calculation formula for normals to inverse images formulated in Exercise 1.54(ii). △

3.2.3 Other Rules of Coderivative Calculus

The next theorem, which is in fact a consequence of Theorems 3.9 and 3.11, applies to general *binary operations* on set-valued mappings including addition, substraction, various kinds of multiplication and division , as well as taking maxima, minima, etc. We formalize this via the following *h-compositions*

$$(F_1 \overset{h}{\diamond} F_2)(x) := \bigcup \big\{h(y_1, y_2)\big|\ y_1 \in F_1(x),\ \ y_2 \in F_2(x)\big\}$$

of two multifunctions $F_1 \colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ and $F_2 \colon \mathbb{R}^n \rightrightarrows \mathbb{R}^l$, where the single-valued mapping $h \colon \mathbb{R}^m \times \mathbb{R}^l \to \mathbb{R}^q$ represents various binary operations. For brevity we present only the inclusion formula for coderivatives corresponding to the case of inner semicontinuity in Theorem 3.11.

Theorem 3.14 (Coderivatives of Compositions with Respect to Binary Operations). *For $F_1 \colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ and $F_2 \colon \mathbb{R}^n \rightrightarrows \mathbb{R}^l$, consider their h-composition $F_1 \overset{h}{\diamond} F_2$ with some $h \colon \mathbb{R}^m \times \mathbb{R}^l \to \mathbb{R}^q$. Pick $\bar{z} \in (F_1 \overset{h}{\diamond} F_2)(\bar{x})$ and suppose that the set-valued mapping $S \colon \mathbb{R}^n \times \mathbb{R}^q \rightrightarrows \mathbb{R}^m \times \mathbb{R}^l$ defined by*

$$S(x, z) := \big\{(y_1, y_2) \in \mathbb{R}^m \times \mathbb{R}^l \big|\ y_i \in F_i(x),\ \ z = h(y_1, y_2)\big\}$$

is inner semicontinuous at the given $(\bar{x}, \bar{z}, \bar{y}) \in \operatorname{gph} S$ with $\bar{y} = (\bar{y}_1, \bar{y}_2)$ and that h is locally Lipschitzian around $\bar{y}$. Then for all $z^* \in \mathbb{R}^q$ we have

$$D^*(F_1 \overset{h}{\diamond} F_2)(\bar{x}, \bar{z})(z^*) \subset \bigcup_{(y_1^*, y_2^*) \in D^*h(\bar{y})(z^*)} \Big[D^*F_1(\bar{x}, \bar{y}_1)(y_1^*) + D^*F_2(\bar{x}, \bar{y}_2)(y_2^*)\Big].$$

Proof. Define $F \colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m \times \mathbb{R}^l$ by $F(x) := (F_1(x), F_2(x))$, and get

$$D^*F(\bar{x}, \bar{y})(y_1^*, y_2^*) \subset D^*F_1(\bar{x}, \bar{y}_1)(y_1^*) + D^*F_2(\bar{x}, \bar{y}_2)(y_2^*). \tag{3.33}$$

Indeed, this follows by applying Theorem 3.9 to the sum $F = \widetilde{F}_1 + \widetilde{F}_2$, where $\widetilde{F}_1(x) := (F_1(x), 0)$ and $\widetilde{F}_2(x) := (0, F_2(x))$. Since

$$(F_1 \overset{h}{\diamond} F_2)(x) = (h \circ F)(x)$$

and h is locally Lipschitzian around $\bar{y}$, we employ the chain rule of Corollary 3.12 to $h \circ F$. Combining it with (3.33) justifies the claimed result. △

We conclude this section by illustrating the application of Theorem 3.14 to calculate the coderivative of *inner product*

$$\langle F_1, F_2\rangle(x) := \big\{\langle y_1, y_2\rangle\big|\ y_i \in F_i(x),\ i = 1, 2\big\}$$

of set-valued mappings $F_1, F_2\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$.

Corollary 3.15 (Coderivatives of Inner Products). *Given $\bar{\nu} \in \langle F_1, F_2\rangle(\bar{x})$ and $\bar{y}_i \in F_i(\bar{x})$ with $\bar{\alpha} = \langle \bar{y}_1, \bar{y}_2\rangle$, suppose that the mapping*

$$(x, \nu) \mapsto \big\{(y_1, y_2) \in \mathbb{R}^{2m}\big|\ y_i \in F_i(x),\ \nu = \langle y_1, y_2\rangle\big\}$$

is inner semicontinuous at $(\bar{x}, \bar{\nu}, \bar{y}_1, \bar{y}_2)$ and that the qualification condition (3.22) is satisfied. Then for all $\lambda \in \mathbb{R}$ we have

$$D^*\langle F_1, F_2\rangle(\bar{x}, \bar{\nu})(\lambda) \subset D^*F_1(\bar{x}, \bar{y}_1)(\lambda\bar{y}_2) + D^*F_2(\bar{x}, \bar{y}_2)(\lambda\bar{y}_1).$$

Proof. Follows from Theorem 3.14 with $h(y_1, y_2) = \langle y_1, y_2\rangle$. △

3.3 Coderivative Analysis of Variational Systems

Now we consider a broad class of *parametric variational systems* (PVS)

$$S(x) := \big\{y \in \mathbb{R}^m\big|\ 0 \in f(x, y) + Q(y)\big\}, \quad x \in \mathbb{R}^n, \tag{3.34}$$

defined by single-valued mappings f and set-valued mappings Q. Employing and further developing appropriate results of *coderivative calculus* allow us to express the coderivative of S via the corresponding constructions for the initial data f and Q. Using these calculations, the *coderivative criteria* for well-posedness, and the subsequent analysis leads us to a rather surprising (at the first glance) conclusion that the naturally desired well-posedness property of *metric regularity fails* for PVS (3.34) in fairly general settings.

3.3.1 Parametric Variational Systems

The parametric formalism of *generalized equations* (GEs) is given by

$$0 \in f(x, y) + Q(y) \quad \text{with} \quad x \in \mathbb{R}^n, \; y \in \mathbb{R}^m, \tag{3.35}$$

where $f\colon \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^q$ is a single-valued *base* mapping dependent on both the *decision* variable y and the *parameter* variable x, while $Q\colon \mathbb{R}^m \rightrightarrows \mathbb{R}^q$ is a parameter-independent set-valued *field* mapping. This formalism and the GE name were coined by Robinson [661] for the case where $Q(y)$ is the *normal cone* mapping to a *convex* set. The GE model (3.35) has been well recognized as a convenient framework to study a variety of qualitative and quantitative/numerical aspects of variational analysis, equilibria, etc. not only in finite-dimensional but also in infinite-dimensional spaces; see Sections 3.4 and 3.5. Note that in the original setting of the normal cone mapping $Q(y) := N(y; \Omega)$ in (3.35) generated by a convex set Ω, the GEs under consideration can be rewritten in the form of parameterized *variational inequalities*:

$$\text{find } \; y \in \mathbb{R}^m \; \text{ such that } \; \langle f(x, y), v - y \rangle \geq 0 \; \text{ for all } \; v \in \Omega, \tag{3.36}$$

which cover various complementarity problems, KKT (Karush-Kuhn-Tucker) systems of first-order conditions in constrained optimization, etc.

The set-valued mapping $x \mapsto S(x)$ of the parameter x defined in (3.34) is known as the *solution map* associated with GE (3.34). Important issues in the theory and applications of parametric GEs revolve around *well-posedness* properties of their solution maps. The three fundamental robust properties of this type have been studied and characterized above via coderivatives in the general framework of set-valued mappings. Natural questions arise about the validity of these properties in the particular framework of solution mappings to parametric GEs. Having in hand the obtained coderivative criteria for well-posedness and the developed rules of coderivative calculus allows us to efficiently resolve these issues for PVS. In fact, a lot has been done in this direction for the Lipschitz-like property of (3.34), a crucial ingredient of robust Lipschitzian stability of parametric GEs; see, e.g., [522, Chapter 4]. The outcome for Lipschitzian stability of (3.34) is generally *positive*: it holds under unrestrictive qualification conditions imposed on the initial data of (3.35). In contrast we show below that this is not the case for the *metric regularity* and the equivalent covering/linear openness properties, which *fail*, in particular, in the case of *subdifferential PVS* where Q stands for subdifferential/normal cone mappings generated by convex and other types of "nice" functions. Observe that the situation is completely different for the general case of *parametric constraint systems* (PCS) given in the form

$$F(x) := \big\{ y \in \mathbb{R}^m \big| \; g(x, y) \in \Theta \big\}, \quad x \in \mathbb{R}^n, \tag{3.37}$$

where both Lipschitz-like and metric regularity properties hold under unrestrictive assumptions; see Section 3.5. The main difference between PVS and PCS is the underlying *subdifferential/normal cone* structure of the multivalued field part $Q(y)$ in (3.34), which accumulates *variational information* on the model (variational inequalities, KKT optimality conditions, etc.).

In the rest of this section, we assume that the base mapping f in (3.35) is *continuously differentiable* around the reference point $(\bar{x}, \bar{y})$ satisfying $\bar{z} := -f(\bar{x}, \bar{y}) \in Q(\bar{y})$ and its partial derivative with respect to the parameter x is *surjective* at this point, i.e., the Jacobian matrix $\nabla_x f(\bar{x}, \bar{y})$ is of *full rank* in finite dimensions; see Sections 3.4 and 3.5 for some relaxations of this assumptions. Recall also that our standing hypotheses include the local *closed-graph* requirement of the field mapping Q.

The following result presents an exact calculation of the coderivative of the solution map (3.34) via the Jacobian of f and the coderivative of Q.

Proposition 3.16 (Coderivative Calculation for General PVS). *Under the imposed full-rank assumption, the coderivative of* (3.34) *is calculated by*

$$D^*S(\bar{x}, \bar{y})(y^*) = \Big\{x^* \in \mathbb{R}^n \Big| \; \exists z^* \in \mathbb{R}^q \;\; \text{with} \;\; x^* = \nabla_x f(\bar{x}, \bar{y})^* z^*, \\ -y^* \in \nabla_y f(\bar{x}, \bar{y})^* z^* + D^*Q(\bar{y}, \bar{z})(z^*)\Big\} \tag{3.38}$$

for any $y^* \in \mathbb{R}^m$*. In particular, we have the relationship*

$$\ker D^*S(\bar{x}, \bar{y}) = -D^*Q(\bar{y}, \bar{z})(0). \tag{3.39}$$

Proof. It is easy to observe the representation

$$\text{gph}\, S = \big\{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m \big| \; g(x, y) \in \Theta\big\} = g^{-1}(\Theta) \\ \text{with} \;\; g(x, y) := \big(y, -f(x, y)\big) \;\; \text{and} \;\; \Theta := \text{gph}\, Q.$$

We deduce from the above structure of g that $\nabla g(\bar{x}, \bar{y})$ is surjective *if and only if* $\nabla_x f(\bar{x}, \bar{y})$ is surjective. Applying the normal cone formula from Exercise 1.54(i) and performing elementary calculations give us representation (3.38). To verify now the relationship in (3.39), take any $y^* \in \ker D^*S(\bar{x}, \bar{y})$ and by the kernel definition and formula (3.38) such $z^* \in \mathbb{R}^q$ that

$$0 = \nabla_x f(\bar{x}, \bar{y})^* z^* \;\; \text{and} \;\; -y^* \in \nabla_y f(\bar{x}, \bar{y})^* z^* + D^*Q(\bar{y}, \bar{z})(z^*). \tag{3.40}$$

Since $\nabla_x f(\bar{x}, \bar{y})$ is surjective, the first equality in (3.40) yields $z^* = 0$. Hence the second equality therein reduces to $-y^* \in D^*Q(\bar{y}, \bar{z})(0)$, which ensures the inclusion "$\subset$" in (3.39). The opposite inclusion in (3.39) follows trivially from (3.38) even without using the surjectivity of $\nabla_x f(\bar{x}, \bar{y})$. △

Now we consider two kinds of *structural PVS*, where the set-valued part Q in (3.34) is represented via some compositions of particular mappings that overwhelmingly arise in theoretical and practical models of optimization, equilibria, economics, mechanics, etc.; see more comments in Section 3.5. The first class of structural PVS is described in the form

$$S(x) = \big\{y \in \mathbb{R}^m \big| \; 0 \in f(x, y) + \partial(\psi \circ g)(y)\big\}, \tag{3.41}$$

where $f: \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ and $g: \mathbb{R}^m \to \mathbb{R}^p$ are single-valued, where $\psi: \mathbb{R}^p \to \overline{\mathbb{R}}$ is extended-real-valued, and where $\partial\varphi: \mathbb{R}^m \rightrightarrows \mathbb{R}^m$ is a subgradient mapping generated by a function $\varphi: \mathbb{R}^m \to \overline{\mathbb{R}}$, which is represented as the composition $\varphi(y) = (\psi \circ g)(y)$. Borrowing the mechanical terminology, we label (3.41) as *subdifferential PVS* with *composite potentials*.

To calculate the coderivative of (3.41) with subsequent applications to metric regularity, we invoke *coderivative calculus* allowing us to deduce from Proposition 3.16 an efficient representation of $D^*S(\bar{x}, \bar{y})$ entirely in terms of the given data of (3.41). Furnishing this requires a new *second-order* subdifferential constructions introduced as follows.

Definition 3.17 (Second-Order Subdifferential). *Let $\varphi: \mathbb{R}^n \to \overline{\mathbb{R}}$ be finite at $\bar{x}$, and let $\bar{x}^* \in \partial\varphi(\bar{x})$. Then the* SECOND-ORDER SUBDIFFERENTIAL *of φ at $\bar{x}$ relative to $\bar{x}^*$ is defined by*

$$\partial^2\varphi(\bar{x}, \bar{x}^*)(u) := \big(D^*\partial\varphi\big)(\bar{x}, \bar{x}^*)(u), \quad u \in \mathbb{R}^n, \tag{3.42}$$

via the coderivative of the first-order subgradient mapping $\partial\varphi: \mathbb{R}^n \rightrightarrows \mathbb{R}^n$, where we drop indicating $\bar{x}^* = \nabla\varphi(\bar{x})$ when $\partial\varphi(\bar{x})$ is a singleton.

It follows from Proposition 1.12 and Corollary 1.24 that the second-order subdifferential mapping (3.17) reduces to the (symmetric) Hessian matrix $\nabla^2\varphi(\bar{x})$ linearly applied to $u \in \mathbb{R}^n$ if φ is $\mathcal{C}^2$-smooth around $\bar{x}$, i.e.,

$$\partial^2\varphi(\bar{x})(u) = \big\{\nabla^2\varphi(\bar{x})^*u\big\} = \big\{\nabla^2\varphi(\bar{x})u\big\}, \quad u \in \mathbb{R}^n. \tag{3.43}$$

This allows us to treat $u \mapsto \partial^2\varphi(\bar{x}, \bar{x}^*)(u)$ as a (positively homogeneous) *generalized Hessian* mapping for extended-real-valued functions.

The next result provides a precise calculation of the coderivative for the subdifferential PVS of type (3.41).

Proposition 3.18 (Coderivative Calculation for Subdifferential PVS with Composite Potentials). *Let $(\bar{x}, \bar{y}) \in \operatorname{gph} S$ for (3.41) with $\bar{q} := -f(\bar{x}, \bar{y}) \in \partial(\psi \circ g)(\bar{y})$. In addition to the full-rank assumption on $\nabla_x f(\bar{x}, \bar{y})$, suppose that $g: \mathbb{R}^m \to \mathbb{R}^p$ is $\mathcal{C}^2$-smooth around $\bar{y}$ with the full -rank derivative $\nabla g(\bar{y})$. Let $\bar{v} \in \mathbb{R}^p$ be a (unique) solution to the system*

$$\bar{q} = \nabla g(\bar{y})^*\bar{v} \;\; \textit{with} \;\; \bar{v} \in \partial\psi(\bar{w}) \;\; \textit{and} \;\; \bar{w} := g(\bar{y}). \tag{3.44}$$

Then the coderivative of S at $(\bar{x}, \bar{y})$ is calculated by

$$\begin{aligned} D^*S(\bar{x}, \bar{y})(y^*) = \Big\{x^* \in \mathbb{R}^n \Big| \; &\exists\, u \in \mathbb{R}^m \;\; \textit{with} \;\; x^* = \nabla_x f(\bar{x}, \bar{y})^*u, \\ -y^* \in \nabla_y f(\bar{x}, \bar{y})^*u + \nabla^2\langle\bar{v}, g\rangle(\bar{y})^*u &+ \nabla g(\bar{y})^*\partial^2\psi(\bar{w}, \bar{v})\big(\nabla g(\bar{y})u\big)\Big\} \end{aligned} \tag{3.45}$$

via the second-order subdifferential (3.42) of ψ. Furthermore, we have

$$\ker D^*S(\bar{x}, \bar{y}) = -\nabla g(\bar{y})^*\partial^2\psi(\bar{w}, \bar{v})(0). \tag{3.46}$$

Proof. Using Theorem 3.16 with $Q = \partial(\psi \circ g)$ in the composite subdifferential model (3.41), we get due to the construction of $\partial^2\varphi$ in (3.42) that

$$D^*S(\bar{x}, \bar{y})(y^*) = \Big\{x^* \in \mathbb{R}^n \Big| \; \exists u \in \mathbb{R}^m \; \text{ with } \; x^* = \nabla_x f(\bar{x}, \bar{y})^* u, \\ -y^* \in \nabla_y f(\bar{x}, \bar{y})^* u + \partial^2\big(\psi \circ g\big)(\bar{y}, \bar{q})(u)\Big\}.$$

The second-order subdifferential chain rule from Exercise 3.78(i) applied to the composition $\psi \circ g$ gives us under the assumptions made that

$$\partial^2\big(\psi \circ g\big)(\bar{y}, \bar{q})(u) = \nabla^2\langle \bar{v}, g\rangle(\bar{y})^* u + \nabla g(\bar{y})^* \partial^2\psi(\bar{w}, \bar{v})\big(\nabla g(\bar{y})u\big). \tag{3.47}$$

Substituting (3.47) into the above expression for $D^*S(\bar{x}, \bar{y})(y^*)$, we arrive at (3.45). The relationship in (3.46) follows from employing the second-order chain rule (3.47) in formula (3.39) with $Q = \partial(\psi \circ g)$. △

Next we consider yet another specification of PVS in (3.34) given by

$$S(x) := \big\{y \in \mathbb{R}^m \big| \; 0 \in f(x, y) + \big(\partial\psi \circ g\big)(y)\big\}, \tag{3.48}$$

where the field Q is a composition of the basic subdifferential of $\psi : \mathbb{R}^p \to \overline{\mathbb{R}}$ and a mapping $g : \mathbb{R}^m \to \mathbb{R}^p$ and where $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p$. Note that such subdifferential PVS with *composite fields* are distinct from those in (3.41) with composite potentials having different ranges of applications. In particular, formalism (3.48) encompasses perturbed *implicit complementarity problems* of the type: find $y \in \mathbb{R}^m$ satisfying the relationships

$$f(x, y) \geq 0, \quad y - g(x, y) \geq 0, \quad \big\langle f(x, y), y - g(x, y)\big\rangle = 0,$$

where the first two inequalities are understood in the vector sense.

The following proposition contains coderivative evaluations for (3.48) with and without the full-rank assumptions on the Jacobian matrix $\nabla g(\bar{y})$.

Proposition 3.19 (Coderivative Evaluations for PVS with Composite Fields). *Consider PVS (3.48) with $(\bar{x}, \bar{y}) \in \operatorname{gph} S$ under the full-rank assumption on $\nabla_x f(\bar{x}, \bar{y})$, where $g : \mathbb{R}^m \to \mathbb{R}^p$ is of class $\mathcal{C}^1$ around $\bar{y}$, while $\psi : \mathbb{R}^p \to \overline{\mathbb{R}}$ is finite at $\bar{w} := g(\bar{y})$. The following assertions hold:*

(i) *If the Jacobian matrix $\nabla g(\bar{y})$ has full rank, then*

$$D^*S(\bar{x}, \bar{y})(y^*) = \Big\{ x^* \in \mathbb{R}^n \Big| \; \exists u \in \mathbb{R}^p \; \text{ with } \; x^* = \nabla_x f(\bar{x}, \bar{y})^* u, \\ -y^* \in \nabla_y f(\bar{x}, \bar{y})^* u + \nabla g(\bar{y})^* \partial^2\psi(\bar{w}, \bar{q})(u)\Big\} \tag{3.49}$$

for all $y^* \in \mathbb{R}^m$, where $\bar{q} := -f(\bar{x}, \bar{y})$. Moreover, we have the relationship

$$\ker D^*S(\bar{x}, \bar{y}) = -\nabla g(\bar{y})^* \partial^2\psi(\bar{w}, \bar{q})(0). \tag{3.50}$$

(ii) *Let the mapping $\partial\psi: \mathbb{R}^p \rightrightarrows \mathbb{R}^p$ be closed-graph around $(\bar{w}, \bar{q})$, and let the full-rank assumption on $\nabla g(\bar{y})$ be replaced by the qualification condition*

$$\partial^2\psi(\bar{w}, \bar{q})(0) \cap \ker \nabla g(\bar{y})^* = \{0\}. \tag{3.51}$$

Then we have the inclusion "$\subset$" in both formulas (3.49) *and* (3.50).

Proof. Applying formula (3.38) of Proposition 3.16 to the composite field $Q = \partial\psi \circ g$ gives us the representation

$$D^*S(\bar{x}, \bar{y})(y^*) = \Big\{ x^* \in \mathbb{R}^n \Big| \; \exists\, u \in \mathbb{R}^p \;\text{ with }\; x^* = \nabla_x f(\bar{x}, \bar{y})^* u, \\ -y^* \in \nabla_y f(\bar{x}, \bar{y})^* u + D^*\big(\partial\psi \circ g\big)(\bar{y}, \bar{q})(u) \Big\} \tag{3.52}$$

for the mapping S from (3.48). To proceed further, we need to use an appropriate chain rule for evaluating the coderivative of the composition $\partial\psi \circ g$. In case (i) it follows from Theorem 3.11(iii) under the full-rank assumption that

$$D^*\big(\partial\psi \circ g\big)(\bar{y}, \bar{q})(u) = \nabla g(\bar{y})^* \partial^2\psi(\bar{w}, \bar{q})(u), \quad u \in \mathbb{R}^p, \tag{3.53}$$

where the closed-graph property of the subgradient mapping $F = \partial\psi$ is not needed in this case; see [522, Theorem 1.66]. Substituting this chain rule into (3.52), we get (3.49) and similarly derive (3.50) from (3.39).

In case (ii) we apply the coderivative chain rule held as the inclusion "$\subset$" from Theorem 3.11(i) with $D^*g(\bar{y})(y^*) = \{\nabla g(\bar{y})^* y^*\}$, where the closed-graph property of $\partial\psi$ is required in the proof; cf. [522, Theorem 3.16]. The qualification condition (3.26) reduces in this case to (3.51), while the chain rule inclusion (3.27) yields "$\subset$" in both formulas (3.49) and (3.50). △

3.3.2 Coderivative Conditions for Metric Regularity of PVS

In this subsection, based on the coderivative characterization of metric regularity for general closed-graph multifunction from Theorem 3.3(ii) and the exact coderivative calculation for PVS (3.34) given in Proposition 3.16, we establish conditions ensuring metric regularity of general PVS and their important specifications. The latter requires applying coderivative calculus.

The first theorem concerns general PVS (3.34) and contains, in particular, the *equivalence* statement regarding the well-posedness properties of metric regularity for solution maps to GEs (3.35) and Lipschitzian behavior of their fields at the corresponding points.

Theorem 3.20 (Metric Regularity of General PVS). *Under the standing assumptions made, we have that the solution map S in* (3.34) *is metrically regular around $(\bar{x}, \bar{y}) \in \operatorname{gph} S$ if and only if*

$$D^*Q(\bar{y}, \bar{z})(0) = \{0\} \;\textit{ with }\; \bar{z} := -f(\bar{x}, \bar{y}) \in Q(\bar{y}), \tag{3.54}$$

i.e., it is equivalent to the Lipschitz-like property of the field Q around $(\bar{y}, \bar{z})$. Furthermore, the exact regularity bound of S around $(\bar{x}, \bar{y})$ is calculated by

$$\operatorname{reg} S(\bar{x}, \bar{y}) = \max \Big\{ \|y^*\| \;\Big|\; \exists\, z^* \in \mathbb{R}^q \;\; \text{with} \;\; \|\nabla_x f(\bar{x}, \bar{y})^* z^*\| \le 1, \quad -y^* \in \nabla_y f(\bar{x}, \bar{y})^* z^* + D^* Q(\bar{y}, \bar{z})(z^*) \Big\}. \tag{3.55}$$

Proof. Since the solution map $S(\cdot)$ is clearly closed-graph around $(\bar{x}, \bar{y})$ in the setting of the theorem, characterization (3.54) of its metric regularity around this point follows from Theorem 3.3(ii) and formula (3.39) in Proposition 3.16. The Lipschitz-like property of Q around $(\bar{y}, \bar{z})$ is the result of Theorem 3.3(iii). The exact bound representation (3.55) is a consequence of the general formula (3.8) and the coderivative calculation for PVS in (3.38). The maximum is attained in (3.55) due to the assumed surjectivity of $\nabla_x f(\bar{x}, \bar{y})$ in the finite-dimensional setting under consideration. △

Now we derive two consequences of Theorem 3.20 for the subdifferential PVS considered above based on calculating the coderivative of Q as in the proofs of Propositions 3.18 and 3.19. Note that, besides the coderivative calculation in (3.54), we need also checking the closed-graph property of the fields in these systems, which is the standing assumption of Theorem 3.20.

Corollary 3.21 (Metric Regularity of Subdifferential PVS with Composite Potentials). *In addition to the assumptions of Proposition 3.18, suppose that the subgradient mapping $\partial\psi\colon \mathbb{R}^p \rightrightarrows \mathbb{R}^p$ is closed-graph around $(\bar{w}, \bar{v})$ in the notation therein. Then S in (3.41) is metrically regular around $(\bar{x}, \bar{y})$ if and only if $\partial\psi$ is Lipschitz-like around $(\bar{w}, \bar{v})$.*

Proof. The first-order subdifferential chain rule in Exercise 1.72 clearly implies that the closed-graph assumption on $\partial\psi$ ensures this property of the mapping $Q = \partial(g \circ \psi)$ and hence of S in (3.41). It follows from Theorem 3.20 and Definition 3.17 of the second-order subdifferential that S from (3.41) is metrically regular around $(\bar{x}, \bar{y})$ if and only if we have

$$D^* Q(\bar{y}, \bar{q})(0) := \partial^2 \big(\psi \circ g\big)(\bar{y}, \bar{q})(0) = \{0\}. \tag{3.56}$$

Applying now the second-order subdifferential chain rule (3.47) tells us that (3.56) is equivalent to the condition

$$\nabla g(\bar{y})^* \partial\psi(\bar{w}, \bar{v})(0) = \{0\},$$

which is equivalent in turn to $\partial^2\psi(\bar{w}, \bar{v})(0) = \{0\}$ due to the surjectivity of $\nabla g(\bar{y})$; see Exercise 1.53. The latter is a characterization of the Lipschitz-like property of the subgradient mapping $\partial\psi$ around $(\bar{w}, \bar{v})$ by Theorem 3.3(iii). Note that we can arrive at the same conclusion by using the kernel formula (3.46) due to Theorem 3.3(ii). △

The next corollary concerns metric regularity of the second type (3.48) of the subdifferential PVS under consideration.

Corollary 3.22 (Metric Regularity of Subdifferential PVS with Composite Fields). *In the setting of Proposition 3.19*(i)*, suppose in addition that the subgradient mapping $\partial\psi\colon \mathbb{R}^p \rightrightarrows \mathbb{R}^p$ is closed-graph around $(\bar{w}, \bar{q})$. Then the solution map S from* (3.48) *is metrically regular around $(\bar{x}, \bar{y})$ if and only if $\partial\psi$ is Lipschitz-like around $(\bar{w}, \bar{q})$.*

Proof. Since both mappings $Q = \partial\psi \circ g$ and S are obviously closed-graph under the imposed assumptions, the claimed metric regularity assertion reduces by Theorem 3.20 and the coderivative criteria of Theorem 3.3 to check that

$$\text{either } \; D^*(\partial\psi \circ g)(\bar{y}, \bar{q})(0) = \{0\} \;\text{ or }\; \ker S(\bar{x}, \bar{y}) = \{0\}.$$

Both conditions above are equivalent to

$$\nabla g(\bar{y})^* \partial\psi^2(\bar{w}, \bar{q})(0) = \{0\}$$

by using the chain rule equality from Theorem 3.11(iii) in the first case and by formula (3.50) from Proposition 3.19 in the second one. The latter condition can be equivalently rewritten as $\partial\psi^2(\bar{w}, \bar{q})(0) = \{0\}$ by the injectivity of $\nabla g(\bar{y})^*$ as in the proof of Corollary 3.21, and thus we characterize the Lipschitz-like property of the subgradient mapping $\partial\psi$ around $(\bar{w}, \bar{q})$. △

Note that the local *closedness* assumption imposed of the subdifferential graph $\operatorname{gph} \partial\psi$ in Corollaries 3.21 and 3.22 surely holds if ψ is *continuous* around the corresponding points. This immediately follows from the *robustness* of our basic subdifferential; see graphs Proposition 1.20. On the other hand, the subdifferential closed-graph property holds also for some remarkable classes of extended-real-valued functions. In particular, it happens for every (locally) l.s.c. *convex* function $\psi\colon \mathbb{R}^p \to \overline{\mathbb{R}}$; this can be deduced directly from Proposition 1.25 by the classical subdifferential definition of convex analysis. In fact, the *closed-graph property* is satisfied for subgradients of a significantly broader class of extended-real-valued *amenable* functions defined as follows.

Definition 3.23 (Amenable and Strongly Amenable Functions). *A function $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ is* AMENABLE *at $\bar{x} \in \operatorname{dom}\varphi$ if there is a neighborhood U of $\bar{x}$ on which φ is represented as a composition $\theta \circ h$ of a $\mathcal{C}^1$-smooth mapping $h\colon U \to \mathbb{R}^m$ and a convex l.s.c. function $\theta\colon \mathbb{R}^m \to \overline{\mathbb{R}}$ with*

$$\partial^\infty\theta(\bar{y}) \cap \ker \nabla h(\bar{x})^* = \{0\} \;\; \textit{for} \;\; \bar{y} := \theta(\bar{x}).$$

The function φ is STRONGLY AMENABLE *at $\bar{x}$ if the inner mapping $h\colon U \to \mathbb{R}^m$ above can be selected as $\mathcal{C}^2$-smooth on U.*

Besides convex and smooth functions, amenability encompasses various compositions that naturally appear in numerous settings of variational analysis and constrained optimization, in particular, those written in the unconstrainedextended-real-valued framework; see Exercise 3.88 for some properties of amenable functions and Section 3.5 for more discussions.

We'll employ strongly amenable functions in the next subsection.

3.3.3 Failure of Metric Regularity for Major Classes of PVS

Here we use the coderivative characterizations of metric regularity for PVS obtained above to reveal that this property fails for major classes of such systems, particularly those having a subdifferential/normal cone descriptions that is typical in optimization and equilibria. This strictly distinguishes metric regularity of PVS from the well-posedness property of their robust Lipschitzian stability, in contrast to the case of general PCS (3.37).

An important fact, which eventually rules out the validity of metric regularity for subdifferential PVS generated by nonsmooth convex functions and the like, is the following specification of the fundamental Kenderov's theorem on monotone lower semicontinuous operators; see [408]. The standard *local monotonicity* property of a set-valued operator $T\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ around $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ means that there are neighborhoods U of $\bar{x}$ and V of $\bar{y}$ with

$$\langle v_1 - v_2, u_1 - u_2\rangle \geq 0 \ \text{ for all } \ (u_1, v_1), (u_2, v_2) \in \operatorname{gph} T \cap (U \times V).$$

Proposition 3.24 (Single-Valuedness of Lipschitz-Like Monotone Operators). *Let $T\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be locally monotone and Lipschitz-like around $(\bar{x}, \bar{y}) \in \operatorname{gph} T$. Then it is single-valued around $(\bar{x}, \bar{y})$.*

Proof. Arguing by contradiction, suppose that T is multivalued in any neighborhood of $(\bar{x}, \bar{y})$. Then there exist sequences $x_k \to \bar{x}$ and $y_k, u_k \in T(x_k)$ with $(y_k, u_k) \to (\bar{y}, \bar{y})$ such that $u_k \neq y_k$ for all $k \in \mathbb{N}$. Denote $a_k := \|u_k - y_k\| > 0$ and $z_k := (u_k - y_k)/a_k$ for which we have

$$\langle u_k, z_k\rangle = a_k + \langle y_k, z_k\rangle, \quad k \in \mathbb{N}. \tag{3.57}$$

The assumed Lipschitz-like property of T around $(\bar{x}, \bar{y})$ gives us the existence of positive numbers ℓ and γ such that

$$T(x) \cap B_\gamma(\bar{x}) \subset T(u) + \ell\|x - u\|\mathbb{B} \ \text{ for all } \ x, u \in B_\gamma(\bar{y}).$$

Choose now a sequence of $\nu_k > 0$ satisfying the conditions

$$\nu_k \downarrow 0 \ \text{ with } \ \nu_k < a_k/2\ell \ \text{ as } \ k \to \infty. \tag{3.58}$$

Since $x_k, x_k + \nu_k z_k \in B_\gamma(\bar{x})$ for large k, the Lipschitz-like property of T yields

$$\|v_k - y_k\| \leq \ell\nu_k \ \text{ for some } \ v_k \in T(x_k) \cap B_\gamma(\bar{y}). \tag{3.59}$$

Employing the local monotonicity property of T around $(\bar{x}, \bar{y})$ tells us that

$$\langle v_k - u_k, x_k + \nu_k - x_k \rangle \geq 0,$$

which implies by (3.57) the inequalities

$$\langle v_k, z_k \rangle \geq \langle u_k, z_k \rangle \geq a_k + \langle y_k, z_k \rangle.$$

It follows from here the choice of ν_k in (3.58) and the estimate in (3.59) that

$$a_k + \langle y_k, z_k \rangle \leq \langle v_k, z_k \rangle \leq \langle y_k, z_k \rangle + \ell \nu_k < \langle y_k, z_k \rangle + a_k/2,$$

a contradiction, which verifies the single-valuedness of T around $\bar{x}$. △

The following result, utilizing this proposition and the equivalence relationship of Theorem 3.20, reveals the failure of metric regularity for a general class of PVS (3.34) with monotone fields. Recall again that the lower semicontinuity of extended-real-valued functions is our standing assumption.

Theorem 3.25 (Failure of Metric Regularity for PVS with Monotone Fields). *In addition to the standing assumption of Theorem 3.20, suppose that the field mapping Q is monotone around $(\bar{y}, \bar{z})$ and that there is no neighborhood of $\bar{y}$ on which Q is entirely single-valued. Then PVS (3.34) is not metrically regular around the reference point $(\bar{x}, \bar{y}) \in \operatorname{gph} S$.*

Proof. It follows from Theorem 3.20 in the general setting under consideration that the metric regularity of the solution map S in (3.34) around $(\bar{x}, \bar{y})$ is equivalent to the Lipschitz-like property of the field Q around $(\bar{y}, \bar{z})$. The imposed local monotonicity of Q around this point yields the single-valuedness of Q around $\bar{y}$ by Proposition 3.24. This contradicts the assumption of the theorem and thus completes its proof. △

Since the *set-valuedness* of field mappings is a *characteristic* feature of *generalized* equations as a satisfactory model to describe *variational systems* (otherwise they reduce just to standard equations, which are not of particular interest in the variational framework under consideration), the conclusion of Theorem 3.25 reads that variational systems with monotone fields are not metrically regular under the Jacobian full-rank assumption on base mappings that doesn't seem to be restrictive in the GE setting. A major consequence of Theorem 3.25 is the following corollary concerning *subdifferential* systems with *convex* potentials, which encompass the classical cases of variational inequalities and complementarity problems in (3.36).

Recall that a function $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ finite at $\bar{x}$ is *Gâteaux differentiable* at this point with the Gâteaux derivative $d\varphi(\bar{x})$ if

$$\lim_{t \to 0} \frac{\varphi(\bar{x} + tw) - \varphi(\bar{x}) - t\langle d\varphi(\bar{x}), w \rangle}{t} = 0$$

for any direction $w \in \mathbb{R}^n$; similarly in infinite dimensions. It is obvious that the Fréchet differentiability of φ at $\bar{x}$ implies the Gâteaux one with the same derivative $d\varphi(\bar{x}) = \nabla\varphi(\bar{x})$; see also Exercise 3.90 for other properties.

Corollary 3.26 (Failure of Metric Regularity for Subdifferential PVS with Convex Potentials). *Let $Q(y) = \partial\varphi(y)$ in the setting of Theorem 3.25, where $\varphi\colon Y \to \overline{\mathbb{R}}$ is a convex function finite at $\bar{y}$ but not Gâteaux differentiable around this point. Then S is not metrically regular around $(\bar{x}, \bar{y})$.*

Proof. Observe first that the assumptions imposed on φ ensure that the subgradient mapping $Q(y) = \partial\varphi(y)$ is closed-graph. Furthermore, the fundamental result on monotone operators (due to Moreau and Rockafellar) establishes the maximal monotonicity of the convex subgradient mapping $x \mapsto \partial\varphi(x)$. Thus the conclusion of the corollary follows from the well-known fact of convex analysis that the subdifferential of such a function is a singleton at the reference point if and only if the function is Gâteaux differentiable at it; see, e.g., [638, 667] and the references therein on these classical results. △

Note that the classical settings of variational inequalities and complementarity problems in (3.36) correspond to the *highly nonsmooth* (extended-real-valued) case of the convex *indicator functions* $\varphi(y) = \delta(y; \Omega)$ in (3.34). In fact, essentially more general *nonconvex* subdifferential structures of parametric variational systems prevent the fulfillment of metric regularity for PVS (3.34) *without* reducing them to the case of field monotonicity while by using appropriate *calculus rules* for coderivatives and second-order subdifferentials.

The next major result provides a significant extension of Corollary 3.26 to nonconvex subdifferential structures of fields with *composite potentials* (3.41), being however fully independent of Theorem 3.25 that imposes field monotonicity. We now deal with the case of $Q(y) = \partial\varphi(y)$, where the nonconvex potential φ admits a composite representation $\varphi = \psi \circ g$ via a $\mathcal{C}^2$-smooth mapping $g\colon \mathbb{R}^m \to \mathbb{R}^p$ and an extended-real-valued function $\psi\colon \mathbb{R}^p \to \overline{\mathbb{R}}$ belonging to a broad class of functions well-recognized in variational analysis.

Definition 3.27 (Prox-Regularity and Subdifferential Continuity).

(i) *A function $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ is* PROX-REGULAR *at $\bar{x} \in \operatorname{dom}\varphi$* FOR SOME $\bar{v} \in \partial\varphi(\bar{x})$ *if it is l.s.c. around $\bar{x}$ and there are $\gamma > 0$, $\eta \geq 0$ such that*

$$\varphi(u) \geq \varphi(x) + \langle v, u - x\rangle - \frac{\eta}{2}\|u - x\|^2 \quad \text{whenever } v \in \partial\varphi(x)$$
$$\text{with } \|v - \bar{v}\| \leq \gamma,\ \|u - \bar{x}\| \leq \gamma,\ \|x - \bar{x}\| \leq \gamma,\ \varphi(x) \leq \varphi(\bar{x}) + \gamma.$$

If this holds FOR ANY $\bar{v} \in \partial\varphi(\bar{x})$, *$\varphi$ is said to be* PROX-REGULAR AT $\bar{x}$.

(ii) *A function $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ is* SUBDIFFERENTIALLY CONTINUOUS *at $\bar{x}$* FOR SOME $v \in \partial\varphi(\bar{x})$ *if $\varphi(x_k) \to \varphi(\bar{x})$ whenever $x_k \to \bar{x}$, $v_k \to v$ as $k \to \infty$ with $v_k \in \partial\varphi(x_k)$. When this property holds* FOR ANY $\bar{v} \in \partial\varphi(\bar{x})$, *$\varphi$ is said to be* SUBDIFFERENTIALLY CONTINUOUS *at $\bar{x}$.*

For brevity we label as *continuously prox-regular* any extended-real-valued function satisfying both properties in Definition 3.27. Such functions are overwhelmingly involved in many areas of variational analysis and optimization, especially those related for second-order aspects and applications; see more discussions in Section 3.5. In particular, this class includes every l.s.c. and *convex* and—more generally—*strongly amenable* function as well as functions of *class* $\mathcal{C}^{1,1}$ around $\bar{x}$, i.e., such that there is a neighborhood U of $\bar{x}$ on which φ is smooth and its derivative is Lipschitz continuous; see Exercise 3.92.

The following lemma of its own interest allows us to establish the failure of metric regularity for subdifferential PVS with composite potentials given by continuously prox-regular functions.

Lemma 3.28 (Continuously Prox-Regular Functions with Lipschitz-Like Subdifferentials). *Let* $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ *be continuously prox-regular at* $\bar{x} \in \operatorname{int}(\operatorname{dom}\varphi)$ *for some* $\bar{v} \in \partial\varphi(\bar{x})$*, and let the subgradient mapping* $\partial\varphi\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *be Lipschitz-like around* $(\bar{x}, \bar{v})$*. Then there is a neighborhood* U *of* $\bar{x}$ *such that* φ *is of class* $\mathcal{C}^{1,1}$ *on* U*.*

Proof. See the hints and discussions in Exercise 3.93. △

Now we are ready to justify the failure of metric regularity for subdifferential PVS (3.41) involving continuously prox-regular functions.

Theorem 3.29 (Failure of Metric Regularity for Subdifferential PVS with Continuously Prox-Regular Potentials). *In addition to the assumptions of Corollary 3.21, suppose that* ψ *is continuously prox-regular at* $\bar{w} = g(\bar{y})$ *for the subgradient* $\bar{v} \in \partial\psi(\bar{w})$*, which is uniquely determined by* $\nabla g(\bar{y})^*\bar{v} = -f(\bar{x}, \bar{y})$*. Then PVS* (3.41) *is not metrically regular around* $(\bar{x}, \bar{y})$ *provided that* ψ *is not Gâteaux differentiable around* $\bar{w}$*.*

Proof. It follows from Corollary 3.21 that the metric regularity of S from (3.41) around $(\bar{x}, \bar{y})$ is equivalent to the Lipschitz-like property of the subgradient mapping $\partial\psi$ around $(\bar{w}, \bar{v})$. Further, the imposed continuous prox-regularity of ψ at $\bar{w}$ allows us to conclude by Lemma 3.28 that the latter property of $\partial\psi$ implies that $\psi \in \mathcal{C}^{1,1}$ around $\bar{x}$. This yields by Exercise 3.90(ii) the Gâteaux differentiability of ψ around $\bar{w}$, which shows that S cannot be metrically regular around $(\bar{x}, \bar{y})$ by the last assumption of the theorem. △

The following result is a clear consequence of Theorem 3.29. However, we present its direct proof independent of Lemma 3.28.

Corollary 3.30 (Failure of Metric Regularity for Composite Subdifferential PVS with Strongly Amenable Potentials). *In addition to the assumptions of Proposition 3.18, suppose that* ψ *is convex and finite at* $\bar{w} = g(\bar{y})$ *while not Gâteaux differentiable around this point. Then the parametric variational system* S *from* (3.34) *is not metrically regular around* $(\bar{x}, \bar{y})$*.*

Proof. Observe that according to Definition 3.23, the potential $\varphi = \psi \circ g$ is strongly amenable at $\bar{y}$ and that the subgradient mapping $\partial\psi\colon \mathbb{R}^p \rightrightarrows \mathbb{R}^p$ is locally closed-graph due to the assumptions imposed on ψ; see Exercise 3.88. Since all the requirements of Corollary 3.21 are met, we conclude that the metric regularity of S around $(\bar{x}, \bar{y})$ is equivalent to the Lipschitz-like property of $\partial\psi$ around $(\bar{w}, \bar{v})$, where $\bar{v} \in \partial\psi(\bar{w})$ is uniquely determined by $\nabla g(\bar{y})^*\bar{v} = -f(\bar{x}, \bar{y})$. Arguing finally as in the proof of Corollary 3.26 shows that $\partial\psi$ is not Lipschitz-like $(\bar{w}, \bar{v})$ and thus completes the proof. △

Next we obtain conditions ensuring the failure of metric regularity for subdifferential PVS with *composite fields* (3.48) involving continuously prox-regular functions in their subdifferential components.

Theorem 3.31 (Failure of Metric Regularity for PVS with Composite Fields Containing Subdifferentials of Prox-Regular Functions). *In addition to the assumptions of Corollary 3.22, suppose that ψ is continuously prox-regular at $\bar{w} = g(\bar{y})$ for the subgradient $\bar{q} := -f(\bar{x}, \bar{y}) \in \partial\psi(\bar{w})$ and that ψ is not Gâteaux differentiable around $\bar{w}$. Then the solution map S from (3.48) fails to be metrically regular around $(\bar{x}, \bar{y})$.*

Proof. It follows from Corollary 3.22 that the metric regularity of S around $(\bar{x}, \bar{y})$ is equivalent to the Lipschitz-like property of ψ around $(\bar{w}, \bar{q})$ under the assumptions therein. Employing now Lemma 3.28 tells us that ψ must be of class $\mathcal{C}^{1,1}$ around $\bar{w}$, which shows that S cannot be metrically regular around $(\bar{x}, \bar{y})$ due to the last assumption of the theorem. △

3.4 Exercises for Chapter 3

Exercise 3.32 (Relations for Openness and Covering Properties).

(i) Show that the function $f(x) = x^m$ on $\mathbb{R}$ possesses the conventional openness property (3.4) at $\bar{x} = 0$ for any odd number $1 \neq m \in \mathbb{N}$, but for $m \geq 3$ it doesn't satisfies the covering property (3.1), i.e., openness with linear rate.

(ii) Formulate extensions of these properties to set-valued mappings between metric spaces in terms of the distance functions.

Exercise 3.33 (Lipschitz-Like Property via Distance Functions).

(i) Formulate an equivalent description of the Lipschitz-like property for set-valued mappings $F\colon X \rightrightarrows Y$ between metric spaces via their distance functions.

(ii) Show that F is Lipschitz-like around $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ if and only if the function $(x, y) \mapsto \operatorname{dist}(y; F(x))$ is Lipschitz continuous around this point. *Hint:* Proceed by the definitions and compare it with [674, Theorem 2.3] and [522, Theorem 1.41].

Exercise 3.34 (Lipschitz Continuity of Locally Compact Multifunctions) Let $F\colon X \rightrightarrows Y$ be a set-valued mapping between Banach spaces.

(i) Let F be compact-valued on a given subset $U \subset X$. Show that the Lipschitz continuity of F on U with modulus $\ell \geq 0$ (i.e., the validity of (3.3) with $V = Y$) is equivalent to the Lipschitz continuity

$$\operatorname{haus}\big(F(u), F(x)\big) \leq \ell\|x - u\| \quad \text{for all } \; x, u \in U$$

of the single-valued mapping $x \mapsto F(x)$ from U to the collections of the compact subset of Y equipped with the *Pompeiu-Hausdorff metric*

$$\text{haus}(\Omega_1, \Omega_2) := \inf\left\{\eta \ge 0 \middle|\ \Omega_1 \subset \Omega_2 + \eta\mathbb{B},\ \ \Omega_2 \subset \Omega_1 + \eta\mathbb{B}\right\}.$$

(ii) Let F be locally compact around $\bar{x} \in \text{dom}\, F$, i.e., the values $F(x)$ for all x near x are enclosed into a compact set. Check that F is locally Lipschitzian around $\bar{x}$ if and only if it is Lipschitz-like around $(\bar{x}, \bar{y})$ for every $\bar{y} \in F(\bar{x})$. In this case the exact Lipschitzian bound $\text{lip}\, F(\bar{x})$ of F around $\bar{x}$ is calculated by

$$\text{lip}\, F(\bar{x}) = \max\left\{\text{lip}\, F(\bar{x}, \bar{y}) \middle|\ \bar{y} \in F(\bar{x})\right\} < \infty.$$

(iii) Do (i) and (ii) hold for mappings between general metric spaces?

Hint: To verify (ii), proceed with selecting of a finite covering of a compact set by a collection of neighborhoods; compare it with the proof in [522, Theorem 1.42].

Exercise 3.35 (Coderivatives of Lipschitzian Mappings Between Banach Spaces). Let $F\colon X \rightrightarrows Y$ be a mapping between Banach spaces, and let $\varepsilon \ge 0$.

(i) Assume that F is Lipschitz-like around some $(\bar{x}, \bar{y}) \in \text{gph}\, F$ with modulus $\ell \ge 0$ and show that there exists a number $\eta > 0$ such that

$$\sup\left\{\|x^*\| \ \middle|\ x^* \in \widehat{D}^*_\varepsilon F(x, y)(y^*)\right\} \le \ell\|y^*\| + \varepsilon(1 + \ell), \quad y^* \in Y^*, \tag{3.60}$$

whenever $x \in \bar{x} + \eta\mathbb{B}$ and $y \in F(x) \cap (\bar{y} + \eta\mathbb{B})$. Furthermore, we have

$$D^*_M F(\bar{x}, \bar{y})(0) = \{0\} \ \text{ and } \ \|D^*_M F(\bar{x}, \bar{y})\| \le \text{lip}\, F(\bar{x}, \bar{y}) < \infty. \tag{3.61}$$

(ii) If F is locally Lipschitzian around some $\bar{x} \in \text{dom}\, F$ with modulus $\ell \ge 0$, then there exists a number $\eta > 0$ such that (3.60) holds for all $x \in \bar{x} + \eta\mathbb{B}$ and $y \in F(x)$. Furthermore, the conditions in (3.61) are satisfied for any $\bar{y} \in F(\bar{x})$.

Hint: To verify (3.60), proceed similarly to Step 1 in the proof of Theorem 3.3 and then pass to the limit as $(x, y) \to (\bar{x}, \bar{y})$ and $\varepsilon \downarrow 0$ by the mixed coderivative construction (1.65); compare it with the proofs of [522, Theorems 1.43 and 1.44].

Exercise 3.36 (Semilocal Metric Regularity). Following [510], we say that a set-valued mapping $F\colon X \rightrightarrows Y$ between Banach spaces is *semilocally metrically regular* around $\bar{x} \in \text{dom}\, F$ (resp. around $\bar{y} \in \text{rge}\, F$) with modulus $\mu > 0$ if estimate (3.2) holds with a neighborhood U of $\bar{x}$ and $V = Y$ (resp. with a neighborhood V of $\bar{y}$ and $U = X$) subject to the condition $\text{dist}(y; F(x)) \le \gamma$ for some $\gamma > 0$. The infimum of such moduli is denoted by $\text{reg}\, F(\bar{x})$ (resp. by $\text{reg}\, F(\bar{y})$).

(i) Verify that F is locally Lipschitzian around $\bar{x} \in \text{dom}\, F$ if and only if F^{-1} is semilocally metrically regular around $\bar{x} \in \text{rge}\, F^{-1}$ with $\text{lip}\, F(\bar{x}) = \text{reg}\, F^{-1}(\bar{x})$.

(ii) Assume that F is locally compact around $\bar{x} \in \text{dom}\, F$, and show that F is semilocally metrically regular around this point if and only if it is (locally) metrically regular around $(\bar{x}, \bar{y})$ in the sense of Definition 3.1(b) for every $\bar{y} \in F(\bar{x})$.

(iii) Assume that F^{-1} is locally compact around $\bar{y} \in \text{rge}\, F$ and show that F is semilocally metrically regular around this point if and only if it is (locally) metrically regular around $(\bar{x}, \bar{y})$ for every $\bar{x} \in F^{-1}(\bar{y})$.

Hint: Proceed similarly to the proof in the local case of Theorem 3.2 with taking into account the results of Exercise 3.34(ii).

Exercise 3.37 (Equivalences Between Local Well-Posedness Properties in Banach Spaces). Let $F\colon X \rightrightarrows Y$ be a set-valued mapping between Banach spaces, and let $(\bar{x}, \bar{y}) \in \text{gph}\, F$. Check that Theorem 3.2 and its proof hold in this setting.

Exercise 3.38 (Semilocal Covering). A set-valued mapping $F\colon X \rightrightarrows Y$ between Banach spaces has the *semilocal covering property* around $\bar{x} \in \text{dom}\, F$ with modulus $\kappa > 0$ if there is a neighborhood U of $\bar{x}$ such that inclusion (3.1) holds with $V = Y$. The supremum of all such moduli is denoted by $\text{cov}\, F(\bar{x})$.

(i) Verify that F has the semilocal covering property around $\bar{x} \in \operatorname{dom} F$ if and only if it is semilocally metrically regular around this point. In this case we have the modulus relationship $\operatorname{cov} F(\bar{x}) = 1/\operatorname{reg} F(\bar{x})$.

(ii) Assume that F is locally compact around $\bar{x}$, and show that in this case the semilocal covering property of F around $\bar{x}$ is equivalent to the (local) covering property of F around $(\bar{x}, \bar{y})$ from Definition 3.1(a) for every $\bar{y} \in F(\bar{x})$.

Hint: To justify (i), proceed by the definitions and compare it with the proof of [522, Theorem 1.52]. This yields (ii) by taking into account Exercise 3.36(iii).

Exercise 3.39 (Global Well-Posedness Properties and Their Comparisons).

• The *global* counterpart of the covering property from Definition 3.1(i) is clearly formulated for mappings between metric spaces as follows. Let $F\colon X \rightrightarrows Y$, let $B_X(x, r)$ be the closed ball of X centered at x with radius $r \geq 0$, and let

$$\vartheta(A) := \sup\{r \geq 0 \big| \; B_X(x, r) \subset A\}$$

for some $A \subset X$. Given $\Omega \subset X$ and $\Theta \subset Y$, we say (cf. [505, 507] and [522, Definition 1.51(i)]) that F has the *κ-covering property relative to* Ω *and* Θ if

$$B_X(x, r) \subset \Omega \Longrightarrow \big[B_Y\big(F(x) \cap \Theta, \kappa r\big) \subset F\big(B_X(x, r)\big)\big]. \tag{3.62}$$

• Another global κ-covering property of the mapping $F\colon X \rightrightarrows Y$ relative to the sets $\widetilde{\Omega} \subset X$ and $\widetilde{\Theta} \subset Y$ was introduced in [26] via the following implication:

$$B_X(x, r) \subset \widetilde{\Omega} \Longrightarrow \big[B_Y\big(F(x), \kappa r\big) \cap \widetilde{\Theta} \subset F\big(B_X(x, r)\big)\big]. \tag{3.63}$$

• More recently yet another version of the κ-covering property of $F\colon X \rightrightarrows Y$ relative to $\widehat{\Omega} \subset X$ and $\widehat{\Theta} \subset Y$ with $(\bar{x}, \bar{y}) \in (\widehat{\Omega} \times \widehat{\Theta}) \cap \operatorname{gph} F$ has been considered in [375] and labeled there as $\operatorname{sur}(F, \widehat{\Omega}, \widehat{\Theta}, \gamma, \kappa, \bar{x}, \bar{y})$. This property means that, given a modulus $\kappa \geq 0$, the following implication holds:

$$\big[x \in B_X(x_0, \gamma),\; r \in [0, \gamma]\big] \Longrightarrow \big[B_Y\big(F(x) \cap B_Y(y_0, \kappa\gamma), \kappa r\big) \cap \widehat{U} \subset F\big(B_X(x, r)\big)\big].$$

(i) Prove that $\operatorname{sur}(F, \widehat{\Omega}, \widehat{\Theta}, \gamma, \kappa, \bar{x}, \bar{y}) \Longrightarrow$(3.62) provided that $B_Y(\Theta, \kappa\gamma) \subset \widehat{\Theta}$, $\Theta \subset B_Y(\bar{y}, \kappa\gamma)$, $\Omega \subset \widehat{\Omega} \cap B_X(\bar{x}, \gamma)$, and $\vartheta(\Omega) \leq \gamma$. Show also that the converse implication (3.62)$\Longrightarrow$ $\operatorname{sur}(F, \widehat{\Omega}, \widehat{\Theta}, \gamma, \kappa, \bar{x}, \bar{y})$ holds whenever $B_X(\widehat{\Omega} \cap B_X(\bar{x}, \gamma), \gamma) \subset \Omega$, $B_X(\widehat{x}, 2\gamma) \subset \widehat{\Omega}$, and $B_Y(\bar{y}, \kappa\gamma) \subset \Theta$.

(ii) Show that (3.62)$\Longrightarrow$(3.63) provided that $\widetilde{\Theta} \subset \Theta$ and $B_Y\big(\widetilde{\Theta}, \kappa\vartheta(\widetilde{\Omega})\big) \subset \Theta$. Conversely, verify that (3.63)$\Longrightarrow$(3.62) if $\Omega \subset \widetilde{\Omega}$ and $B_Y\big(\Theta, \kappa\vartheta(\Omega)\big) \subset \widetilde{\Theta}$.

(iii) Formulate metric regularity and Lipschitzian counterparts of the above properties, and establish relationships between them.

Hint: To verify (i) and (ii), proceed by the definitions and compare it with the proof given in [789, Theorem 1].

Exercise 3.40 (Metric Regularity of Differentiable Mappings in Banach Spaces). Let $f\colon X \to Y$ be a single-valued mapping between Banach spaces.

(i) Assume that f is Fréchet differentiable at $\bar{x}$, and show that the space $\nabla f(\bar{x})X$ is closed in Y provided that f is metrically regular at $\bar{x}$. *Hint:* Use the iterative procedure as in the proof of [522, Lemma 1.56].

(ii) Assume that f is strictly differentiable at $\bar{x}$ and show that the surjectivity of $\nabla f(\bar{x})\colon X \to Y$ is necessary and sufficient for the metric regularity of f around $\bar{x}$ with the validity of the exact bound formulas (3.18), where "min" is replaced by "inf" in the second one. *Hint:* Deduce the necessity of the surjectivity condition $\nabla f(\bar{x})X = Y$ from $\ker \nabla f(\bar{x})^* = \{0\}$ due to (i), (3.61), and Exercises 3.37, 1.57(ii). To justify the sufficiency, proceed as in the alternative proof of Corollary 3.8.

Exercise 3.41 (Neighborhood Characterizations of Lipschitz-Like Multifunctions in Asplund Spaces). Let $F\colon X \rightrightarrows Y$ be a (locally closed-graph) mapping between Asplund spaces. The following assertions are equivalent:

(a) F is Lipschitz-like around $(\bar{x}, \bar{y})$.

(b) There are positive numbers ℓ and η such that

$$\sup\Big\{\|x^*\|\;\Big|\;x^* \in \widehat{D}^*F(x,y)(y^*)\Big\} \le \ell\|y^*\|$$

whenever $x \in B_\eta(\bar{x})$, $y \in F(x) \cap B_\eta(\bar{y})$, and $y^* \in Y^*$.

Furthermore, the exact Lipschitzian bound of F around $(\bar{x}, \bar{y})$ is calculated by

$$\operatorname{lip} F(\bar{x},\bar{y}) = \inf_{\eta>0}\sup\Big\{\|\widehat{D}^*F(x,y)\|\;\Big|\;x \in B_\eta(\bar{x}),\;\; y \in F(x) \cap B_\eta(\bar{y})\Big\}.$$

Hint: Consider first the finite-dimensional case and derive this from Theorem 3.3(iii) and the coderivative representation (1.17). In the Asplund space case, proceed similarly to the proof of [522, Theorem 4.1] for the covering property.

Exercise 3.42 (Sequential and Partial Sequential Normal Compactness of Mappings). A set-valued mapping $F\colon X \rightrightarrows Y$ between Banach spaces is said to be *sequentially normally compact* (SNC) at $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ if for any sequence $(\varepsilon_k, x_k, y_k, x_k^*, y_k^*) \in [0,\infty) \times (\operatorname{gph} F) \times X^* \times Y^*$ we have the implication

$$\begin{aligned}&\Big[\varepsilon_k \downarrow 0,\;(x_k, y_k) \to (\bar{x},\bar{y}),\;(x_k^*, y_k^*) \xrightarrow{w^*} (0,0),\\ &(x_k^*, y_k^*) \in \widehat{N}_{\varepsilon_k}\big((x_k, y_k); \operatorname{gph} F\big)\Big] \Longrightarrow \|(x_k^*, y_k^*)\| \to 0\;\text{ as }\;k \to \infty.\end{aligned} \tag{3.64}$$

The mapping F is *partially sequentially normally compact* (PSNC) at $(\bar{x}, \bar{y})$ for any sequence $(\varepsilon_k, x_k, y_k, x_k^*, y_k^*) \in [0,\infty) \times (\operatorname{gph} F) \times X^* \times Y^*$ we have

$$\begin{aligned}&\Big[\varepsilon_k \downarrow 0,\;(x_k, y_k) \to (\bar{x},\bar{y}),\;x_k^* \xrightarrow{w^*} 0,\;\|y_k^*\| \to 0,\\ &(x_k^*, y_k^*) \in \widehat{N}_{\varepsilon_k}\big((x_k, y_k); \operatorname{gph} F\big)\Big] \Longrightarrow \|x_k^*\| \to 0\;\text{ as }\;k \to \infty.\end{aligned} \tag{3.65}$$

If F is single-valued at $\bar{x}$, the indication of $\bar{y} = F(\bar{x})$ above is omitted.

(i) Check that the SNC property of the mapping F at $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ is equivalent to the SNC property of its graph at the same point.

(ii) Show that we can equivalently put $\varepsilon_k \equiv 0$ in (3.65) if X and Y are Asplund.

(iii) Verify that, besides the obvious cases where $\dim X < \infty$ and where F is SNC at $(\bar{x}, \bar{y})$, the PSNC property of F at $(\bar{x}, \bar{y})$ holds for any mapping $F\colon X \rightrightarrows Y$ which is Lipschitz-like around $(\bar{x}, \bar{y})$. *Hint:* Use Exercise 3.35(i).

Exercise 3.43 (Coderivative Normality). A mapping $F\colon X \rightrightarrows Y$ between Banach spaces is called *coderivatively normal* at $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ if

$$\|D_M^*F(\bar{x},\bar{y})\| = \|D_N^*F(\bar{x},\bar{y})\|. \tag{3.66}$$

(i) We obviously have (3.66) if $D_M^*F(\bar{x},\bar{y})(y^*) = D_N^*F(\bar{x},\bar{y})(y^*)$ for all $y^* \in Y^*$. Does the converse implication hold when $\dim Y = \infty$?

(ii) Give an example showing that (3.66) may fail for a Lipschitzian mapping $f\colon \mathbb{R} \to H$ with values in any separable infinite-dimensional Hilbert space Y.

(iii) Derive sufficient conditions for coderivative normality of set-valued mappings with values in infinite-dimensional spaces. *Hint:* Distill this from the results of the previous exercises and consult also with [522, Proposition 4.9].

Exercise 3.44 (Pointbased Characterizations of Lipschitz-Like Property in Asplund Spaces). Let $F\colon X \rightrightarrows Y$ be a set-valued mapping between Asplund spaces that is closed-graph around $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ by our standing assumption.

(i) Prove that F is Lipschitz-like around $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ if and only if it is simultaneously PSNC at this point and satisfies the condition $D^*_M F(\bar{x}, \bar{y})(0) = \{0\}$.

(ii) Verify the exact bound estimates

$$\|D^*_M F(\bar{x}, \bar{y})\| \le \operatorname{lip} F(\bar{x}, \bar{y}) \le \|D^*_N F(\bar{x}, \bar{y})\| \tag{3.67}$$

in the following cases: (a) arbitrary Banach spaces X and Y for the lower estimate and (b) $\dim X < \infty$ and Y is Asplund for the upper one.

(iii) Is the condition $\dim X < \infty$ essential for the upper estimate in (3.67)?

Hint: Proceed by passing to the limit in the corresponding conditions of Exercise 3.41 and compare it with the proof of [522, Theorem 4.10].

Exercise 3.45 (Local Lipschitz Continuity of Extended-Real-Valued Functions). Let $\varphi\colon X \to \overline{\mathbb{R}}$ be a (l.s.c.) function on an Asplund space X, and let $\bar{x} \in \imath(\operatorname{dom}\varphi)$. Prove that φ is locally Lipschitzian around $\bar{x}$ if and only if $\partial^\infty\varphi(\bar{x}) = \{0\}$ and φ is SNEC at $\bar{x}$.

Hint: Apply the coderivative criterion for Lipschitz-like property of set-valued mappings from Exercise 3.44(i) to the epigraphical multifunction $x \mapsto \operatorname{epi}\varphi$.

Exercise 3.46 (Lipschitzian Properties of Convex-Graph Multifunctions). Let $F\colon X \rightrightarrows Y$ be a convex-graph multifunction between Asplund spaces, and let $\bar{x} \in \operatorname{dom} F$. The following assertions are equivalent:

(a) There is $\bar{y} \in F(\bar{x})$ such that F is Lipschitz-like around $(\bar{x}, \bar{y})$.
(b) The range of F^{-1} is SNC at $\bar{x}$ and $N(\bar{x}; \operatorname{rge} F^{-1}) = \{0\}$.
(c) $\bar{x}$ is an interior point of the range of F^{-1}.
(d) F is Lipschitz-like at $(\bar{x}, \bar{y})$ for every $\bar{y} \in F(\bar{x})$.

If in addition $\dim X < \infty$, then whenever $\bar{y} \in F(\bar{x})$ we have the exact bound formula

$$\operatorname{lip} F(\bar{x}, \bar{y}) = \sup_{\|y^*\|\le 1} \Big\{\|x^*\| \;\Big|\; \langle x^*, x - \bar{x}\rangle \le \langle y^*, y - \bar{y}\rangle \;\text{ for all }\; (x, y) \in \operatorname{gph} F\Big\}.$$

Hint: Derive this from Exercise 3.44 by using the particular coderivative form for convex-graph multifunctions; compare this with the proof of [522, Theorem 4.12].

Exercise 3.47 (Neighborhood Characterizations of Metric Regularity and Covering). Let $F\colon X \rightrightarrows Y$ be a set-valued mapping between Asplund spaces.

(i) Given $(\bar{x}, \bar{y}) \in \operatorname{gph} F$, show that the following assertions are equivalent:

(a) F is metrically regular around $(\bar{x}, \bar{y})$.

(b) We have $\widehat{b}(F, \bar{x}, \bar{y}) < \infty$, where

$$\begin{aligned}\widehat{b}(F, \bar{x}, \bar{y}) := \inf_{\eta>0}\inf\Big\{\mu > 0 \;\Big|\; &\|y^*\| \le \mu\|x^*\|, \quad x^* \in \widehat{D}^* F(x, y)(y^*),\\ &x \in B_\eta(\bar{x}), \quad y \in F(x) \cap B_\eta(\bar{y})\Big\}.\end{aligned}$$

Furthermore, the exact regularity bound of F around $(\bar{x}, \bar{y})$ is calculated by

$$\begin{aligned}\operatorname{reg} F(\bar{x}, \bar{y}) &= \widehat{b}(F, \bar{x}, \bar{y})\\ &= \inf_{\eta>0}\sup\Big\{\|\widehat{D}^* F(x, y)^{-1}\| \;\Big|\; x \in B_\eta(\bar{x}),\; y \in F(x) \cap B_\eta(\bar{y})\Big\}.\end{aligned}$$

(ii) Given $\bar{x} \in \operatorname{dom} F$, obtain versions of the assertions in (i) for the semilocal metric regularity of the mapping F around $\bar{x}$.

(iii) Derive the corresponding counterparts of the assertions in (i) and (ii) for the (local) covering and semilocal covering properties of the mapping F.

Hint: Deduce (i) from the results on the Lipschitz-like property in Exercise 3.41 and the equivalence in Exercise 3.37. Proceed similarly (ii) and (iii).

Exercise 3.48 (Pointbased Characterizations of Metric Regularity in Infinite Dimensions). Given a set-valued mapping $F\colon X \rightrightarrows Y$ between Banach spaces, the *reversed mixed coderivative* of F at $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ is defined by

$$\widetilde{D}^*_M F(\bar{x}, \bar{y})(y^*) := \left\{x^* \in X^* \middle| \; y^* \in -D^*_M F^{-1}(\bar{y}, \bar{x})(-x^*)\right\}, \quad y^* \in Y^*, \tag{3.68}$$

via the mixed coderivative (1.65) of the inverse mapping $F^{-1}\colon Y \rightrightarrows X$.

(i) Provided that X and Y are Asplund, verify that F is metrically regular (or has the covering property) around $(\bar{x}, \bar{y})$ if and only if F^{-1} is PSNC at $(\bar{y}, \bar{x})$ and the kernel condition $\ker \widetilde{D}^*_M F(\bar{x}, \bar{y}) = \{0\}$ is satisfied.

(ii) Show that the "only if" part of (i) holds in the general Banach space setting.

(iii) Derive estimates and precise coderivative formulas for calculating the exact bounds $\operatorname{reg} F(\bar{x}, \bar{y})$ and $\operatorname{cov} F(\bar{x}, \bar{y})$.

(iv) Show that for any separable Banach space X, there is a convex-valued mapping $F\colon X \rightrightarrows X$ which doesn't have the covering and metric regularity properties around $(0, 0) \in \operatorname{gph} F$ while $\ker D^*_N F(0, 0) = \{0\}$.

Hint: To get (i)–(iii), apply the results of Exercise 3.44 to the inverse mapping F^{-1} due to the equivalence relationships from Exercise 3.37. To verify (iv), construct a mapping F by using a countable basis in X such that the PSNC property fails for F^{-1}; compare this with [522, Example 4.19].

Exercise 3.49 (Metric Regularity and Covering Properties for Convex-Graph Multifunctions). Derive characterizations of metric regularity and covering properties of convex-graph multifunctions $F\colon X \rightrightarrows Y$ in the framework of Exercise 3.46. *Hint:* Combine the results of Exercises 3.37 and 3.46, and compare it with the classical Robinson-Ursescu theorem in Banach spaces; see Section 3.5.

Exercise 3.50 (Covering Relative to Mappings and Sets). Given mappings $F\colon X \rightrightarrows Y$ and $\Omega\colon X \rightrightarrows X$ between Banach spaces, $\kappa > 0$, and $\bar{x} \in \Omega(\bar{x}) \cap \operatorname{dom} F$, we say [505] that F has the *covering* property around $\bar{x}$ *relative to* the mapping Ω (in particular, relative to the set Ω when $\Omega(x) \equiv \Omega$) with some modulus $\kappa > 0$ if there is a neighborhood U of $\bar{x}$ such that

$$F(x) + \kappa r \mathbb{B} \subset F\big((x + r\mathbb{B}) \cap \Omega(x)\big) \;\text{ whenever }\; x + r\mathbb{B} \subset U,\; r > 0. \tag{3.69}$$

(i) In the finite-dimensional setting, introduce the *relative covering constant*

$$\kappa(F, \Omega, \bar{x}) := \inf\left\{\|u_1 + u_2\| \middle| u_1 \in D^*F(\bar{x}, \bar{y})(v), u_2 \in N\big(\bar{x}; \Omega(\bar{x})\big),\; \bar{y} \in F(\bar{x}), \|v\| = 1\right\}$$

and show that the condition $\kappa(F, \Omega, \bar{x}) > 0$ is *necessary and sufficient* for the validity of the relative covering property of F with respect to Ω around $\bar{x}$ with some modulus $\kappa > 0$ provided that F is locally Lipschitzian around $\bar{x}$ and that Ω is *normally semicontinuous* at $\bar{x}$ in the following sense:

$$\Big[x_k \xrightarrow{\operatorname{dom} F} \bar{x},\; u_k \xrightarrow{\Omega(x_k)} \bar{x},\; v_k \to v,\; v_k \in N\big(u_k; \Omega(x_k)\big)\Big] \Longrightarrow v \in N\big(\bar{x}; \Omega(\bar{x})\big).$$

Hint: Proceed similarly to the proof of Theorem 3.3, and compare it with the corresponding arguments in [507, Theorem 5.3].

(ii) Show that the mapping $\Omega\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is normally semicontinuous at $\bar{x}$ in the following two cases: (a) $\Omega(x) \equiv \Omega$ around $\bar{x}$ and (b) $\Omega(\cdot)$ is convex-valued around $\bar{x}$ and inner semicontinuous at this point. Any other sufficient conditions?

(iii) Derive extensions of the results in (i) and (ii) to infinite-dimensional spaces.

Exercise 3.51 (Metric Subregularity and Calmness of Multifunctions). Let $F\colon X \rightrightarrows Y$ be a set-valued mapping between Banach spaces, and let $(\bar{x}, \bar{y}) \in \operatorname{gph} F$. The mapping F is *metrically subregular* at $(\bar{x}, \bar{y})$ with modulus $\mu > 0$ if estimate (3.2) holds while $y = \bar{y}$ therein. The corresponding *semilocal* versions of metrical subregularity of F at $\bar{x} \in \operatorname{dom} F$ and $\bar{y} \in \operatorname{rge} F$ are defined similarly to the case of semilocal metric regularity in Exercise 3.36. The mapping F is *calm* at $(\bar{x}, \bar{y})$ with modulus $\ell \geq 0$ if inclusion (3.3) holds while $u = \bar{x}$. If in the latter case $V = Y$, the mapping F is called *upper* (or *outer*) *Lipschitzian* at $\bar{x} \in \operatorname{dom} F$.

(i) Construct examples of mappings between finite-dimensional spaces that are metrically subregular (resp. calm) at some point $(\bar{x}, \bar{y})$ while not metrically regular (resp. Lipschitz-like) around this point.

(ii) Establish two-sided relationships between the metric subregularity (resp. semilocal metric regularity) and calmness (resp. upper Lipschitzian) properties of F and its inverse at the corresponding points similarly to those given in Theorem 3.2 and Exercise 3.36. *Hint:* Proceed as in the proof of Theorem 3.2.

(iii) Formulate an appropriate "subcovering/subopenness" property of multifunctions $F\colon X \rightrightarrows Y$ together with its semilocal version and establish the corresponding relationships with metric subregularity and calmness/upper Lipschitzian properties of F and F^{-1} defined above in this exercise.

Exercise 3.52 (Second-Order Growth Conditions for Metric Regularity and Metric Subregularity of Subdifferential Mappings).

(i) Let $\varphi\colon X \to \overline{\mathbb{R}}$ be convex and l.s.c. function on a Hilbert space X with $\bar{x} \in \operatorname{dom}\varphi$ and $\bar{v} \in \partial\varphi(\bar{x})$. Verify that the subgradient mapping $\partial\varphi\colon X \rightrightarrows X$ is metrically regular around $(\bar{x}, \bar{v})$ if and only if there exist neighborhoods U of $\bar{x}$ and V of $\bar{v}$ along with some $\gamma > 0$ such that

$$(\partial\varphi)^{-1}(v) \neq \emptyset \ \text{ for all } \ v \in V \ \text{ and}$$
$$\varphi(x) \geq \varphi(\bar{x}) - \langle v, u - x\rangle + \gamma \operatorname{dist}^2\big(x; (\partial\varphi)^{-1}(v)\big) \text{ for all } x \in U, u \in (\partial\varphi)^{-1}(v), v \in V.$$

Hint: Use the construction of the subdifferential in convex analysis together with Ekeland's variational principle, and compare it with the proof of [20, Theorem 3.6]. Does this proof hold in any Banach space X?

(ii) Let $\varphi\colon X \to \overline{\mathbb{R}}$ be convex and l.s.c. function on a Banach space X with $\bar{x} \in \operatorname{dom}\varphi$ and $\bar{v} \in \partial\varphi(\bar{x})$. Show that φ is metrically subregular at $(\bar{x}, \bar{v})$ if and only if there is a neighborhood U of $\bar{x}$ and a constant $\gamma > 0$ such that the following second-order/quadratic growth condition is satisfied:

$$\varphi(x) \geq \varphi(\bar{x}) - \langle \bar{v}, \bar{x} - x\rangle + \gamma \operatorname{dist}^2\big(x; (\partial\varphi)^{-1}(\bar{v})\big) \ \text{ whenever } \ x \in U.$$

Hint: Proceed similarly to (i) and compare it with the proofs of [20, Theorem 3.3] for Hilbert spaces and [21, Theorem 2.1] in the general Banach space setting.

(iii) Establish extensions of the results in (i) and (ii) to the basic subdifferential of l.s.c. functions defined on Asplund spaces, with quantitative interrelations between constants in quadratic growth and metric regularity/subregurity, provided that $\bar{x}$ is a local minimizer of φ. *Hint:* Proceed as in the proofs of [232, Theorem 3.1 and Corollary 3.2] in the case of metric subregularity.

(iv) Clarify interconnections between the above second-order growth conditions for metric subregularity of the subgradient mappings $\partial\varphi$ and those for the upper Lipschitzian property of the $(\partial\varphi)^{-1}$ discussed in Exercise 3.55(ii–iv).

Exercise 3.53 (Preservation of Calmness and Metric Subregularity Under Intersections). Let $F_1\colon X_1 \rightrightarrows Y$ and $F_2\colon X_2 \rightrightarrows Y$ be set-valued mappings between metric spaces. Define the intersection mapping $(F_1 \cap F)\colon (X_1 \times X_2) \rightrightarrows Y$ by

$$(F_1 \cap F_2)(x_1, x_2) := F_1(x_1) \cap F_2(x_2), \quad x_1 \in X_1, \ x_2 \in X_2.$$

(i) Assume that F_1 and F_2 are calm at $(\bar{x}_1, \bar{x}) \in \operatorname{gph} F_1$ and $(\bar{x}_2, \bar{x}) \in \operatorname{gph} F_2$, respectively, that F_2^{-1} is Lipschitz-like around $(\bar{x}, \bar{x}_2)$, and that $x_2 \mapsto F_1(\bar{x}_1) \cap F_2(x_2)$ is calm at $(\bar{x}_2, \bar{x})$. Then show that the intersection mapping $F_1 \cap F_2$ is calm at $(\bar{x}_1, \bar{x}_2, \bar{x})$. *Hint:* Compare with the proof of [420, Theorem 2.5].

(ii) Obtain relationships between the exact calmness bound of $F_1 \cap F_2$ at $(\bar{x}_1, \bar{x}_2, \bar{x})$ and the exact bounds of the other properties involved in (i).

(iii) Establish counterparts of (i) and (ii) for the metric subregularity of $F_1 \cap F_2$.

Exercise 3.54 (Outer Derivative of Multifunctions). The *outer derivative* of $F\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ at $\bar{x} \in \operatorname{dom} F$ in the direction $\bar{u} \in \mathbb{R}^n$ is defined by

$$\widehat{D}F(\bar{x})(\bar{u}) := \mathop{\operatorname{Lim\,sup}}_{\substack{t\downarrow 0\\ u\to\bar{u}}} \frac{1 - \Pi_{F(\bar{x})}\big(F(\bar{x}+tu)\big)}{t}, \tag{3.70}$$

where $(1-\Pi_\Omega)(\Theta) := \{z - w \in \mathbb{R}^m \,|\, z \in \Theta,\ w \in \Pi_\Omega(z)\}$ with the Euclidean projector $\Pi_\Omega(z)$ of z to the (locally closed) set Ω taken from (1.3).

(i) Show that $\widehat{D}F(\bar{x})(\bar{u})$ reduces to the contingent derivative $DF(\bar{x})(\bar{u})$ from (1.87) provided that $F(\bar{x})$ is a singleton.

(ii) Assume that $\bar{x}$ is a local minimizer of $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ and verify that

$$\widehat{D}E_\varphi(\bar{x})(u) = \{0\} \text{ for all } u \in \mathbb{R}^n,$$

where $E_\varphi\colon \mathbb{R}^n \rightrightarrows \mathbb{R}$ is an epigraphical multifunction associated with φ.

(iii) Supposing that the set $F(\bar{x})$ is bounded, show that for any $v \in \widehat{D}F(\bar{x})(0)$ there is $z \in F(\bar{x})$ such that $v \in N(z; F(\bar{x}))$.

Hint: Proceed directly by using the corresponding definitions.

Exercise 3.55 (Upper Lipschitzian Mappings and Inverse Subdifferentials).

(i) Prove that $F\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is upper Lipschitzian at $\bar{x} \in \operatorname{dom} F$ if and only if the graph of the outer coderivative $(x,u) \mapsto \widehat{D}F(x)(u)$ is (locally) closed and we have $\widehat{D}F(\bar{x})(0) = \{0\}$. *Hint:* Combine the construction in (3.70) with the definition of the upper Lipschitzian property in finite dimensions; cf. [771, Theorem 3.2].

(ii) Let $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ be l.s.c. on $\mathbb{R}^n$, and let the inverse $(\partial\varphi)^{-1}\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ to the basic subdifferential mapping be upper Lipschitzian at the origin. Show that for any set $\Omega \subset (\partial\varphi)^{-1}(0)$ there are positive constants γ and ν such that

$$\varphi(x) \ge \inf\varphi + \gamma \operatorname{dist}^2\big(x; (\Omega + 2\nu\mathbb{B}) \cap (\partial\varphi)^{-1}(0)\big) \text{ if } x \in \Omega + \nu\mathbb{B}. \tag{3.71}$$

Hint: Use the finite-dimensional variational principle from Theorem 2.12 with a Lipschitzian subadditive function $\theta\colon \mathbb{R}^n \to \mathbb{R}_+$ satisfying $\partial\theta(0) \subset \mathbb{B}$ therein, and then apply the semi-Lipschitzian subdifferential sum rule from Corollary 2.20. Compare this with the proof of [771, Theorem 4.2].

(iii) Assume that φ in (ii) is convex and that $(\partial\varphi)^{-1}(0) \neq \emptyset$. Verify that the quadratic growth condition (3.71) is necessary and sufficient for the upper Lipschitzian property of $(\partial\varphi)^{-1}\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ at the origin, and that (3.71) can be equivalently rewritten in the simplified form

$$\varphi(x) \le \inf\varphi + \gamma \operatorname{dist}^2\big(x; (\partial\varphi)^{-1}(0)\big) \text{ whenever } x \in (\partial\varphi)^{-1}(0) + \nu\mathbb{B}.$$

Hint: Use the subdifferential expression for convex functions; cf. [771, Theorem 4.3].

(iv) Employing the infinite-dimensional versions of the results mentioned in the hint to (ii), extend this statement to the case of Asplund spaces and then show that characterization (iii) holds in any Banach space. What about an infinite-dimensional extension of the outer derivative characterization in (i)?

(v) Clarify the possibility to characterize the calmness and metric subregularity properties of multifunctions between finite-dimensional and infinite-dimensional spaces via an appropriate derivative construction of type (3.70) .

Exercise 3.56 (Semimetric Regularity of Multifunctions). Let $F\colon X \rightrightarrows Y$ be a mapping between Banach spaces, $\Omega \subset X$, and $\bar{x} \in \Omega \cap \operatorname{dom} F$. Consider the set

$$S := \big\{x \in \Omega \big|\ F(x) \cap F(\bar{x}) \neq \emptyset\big\}$$

and say [505] that F is *semimetrically regular* at $\bar{x}$ *relative* to Ω if

$$\operatorname{dist}(x; S) \le \mu\, \beta\big(F(\bar{x}), F(x)\big) \text{ for all } x \in \Omega,\ \|x - \bar{x}\| \le \gamma \tag{3.72}$$

for some $\mu, \gamma > 0$ via the *Hausdorff semidistance* from Θ_1 to Θ_2 defined by

$$\beta(\Theta_1, \Theta_2) := \sup_{x\in\Theta_1} \inf_{u\in\Theta_2} \|x - u\|.$$

(i) Compare this notion with metric subregularity and its semilocal version defined in the corresponding setting of Exercise 3.51.

(ii) In the finite-dimensional setting, suppose that there are $\gamma, b > 0$ such that for any $x \in \Omega \setminus S$ with $\|x - \bar{x}\| \le \gamma$, the mapping F is outer semicontinuous, the function $x \mapsto \operatorname{dist}(z; F(x))$ is locally Lipschitzian when $z \in F(\bar{x})$, and the condition

$$\sup_{z\in F(\bar{x})} \inf\Big\{ \|u_1 + u_2\| \ \Big|\ u_1 \in D^*F(x, y)(v),\ \ y \in \Pi\big(z; F(x)\big),$$
$$\langle v, y - z\rangle = \|y - z\|,\ \|v\| = 1,\ u_2 \in N(x; \Omega)\Big\} \ge b$$

is satisfied. Then the mapping F is semimetrically regular at $\bar{x}$ relative to the set Ω, and we have the modulus estimate $\mu \ge b^{-1}$ in (3.72).

(iii) Derive an extension of (ii) to the case of Asplund spaces.

Hint: Consider the function $\varphi_z(x) := \operatorname{dist}(z; F(x)) + \delta(x; \Omega)$ on $(\Omega \setminus S) \cap B_\gamma(\bar{x})$, and proceed similarly to Step 2 in the proof of Theorem 3.3 with the usage of the subdifferential sum rule and subdifferentiation of the distance function at out-of-set points; compare it with the proof of [507, Theorem 5.4] in finite dimensions.

Exercise 3.57 (Interconnections Between Semimetric Regularity and Covering of Mappings Relative to Sets). In the setting of Exercise 3.50 with $\Omega(\cdot) \equiv \Omega$, suppose that F is locally Lipschitzian around $\bar{x}$. Verify the following assertions (i) and (ii) in the case of $X = \mathbb{R}^n$ and $Y = \mathbb{R}^m$:

(i) If $\kappa(F, \Omega, \bar{x}) > 0$ for the relative covering constant of F with respect to the set Ω, then the mapping F is semimetrically regular at $\bar{x}$ relative to Ω.

(ii) If F is not semimetrically regular at $\bar{x}$ relative to Ω, then there are elements $\bar{y} \in F(\bar{x})$, $v \in \mathbb{R}^m$ with $\|v\| = 1$, and $u \in D^*F(\bar{x}, \bar{y})(v)$ such that $-u \in N(\bar{x}; \Omega)$.

(iii) Extend the results of (i) and (ii) to infinite-dimensional spaces X and Y.

Hint: Proceed by the definitions with using the results from Exercises 3.50and 3.56 and the above coderivative properties of Lipschitzian multifunctions under appropriate sequential normal compactness in infinite dimensions. Compare it with the proof of [507, Corollary 5.4.1] in the case of finite-dimensional spaces.

Exercise 3.58 (Metric Hemiregularity of Multifunctions). A set-valued mapping $F\colon X \rightrightarrows Y$ between Banach spaces is said to be *metrically hemiregular* at $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ with modulus $\mu > 0$ if there is a neighborhood $V \subset Y$ of $\bar{y}$ such that

$$\operatorname{dist}\big(\bar{x}, F^{-1}(y)\big) \le \mu\|y - \bar{y}\| \text{ for all } y \in V. \tag{3.73}$$

The infimum of $\{\mu\}$ over all the combinations (μ, V) for which (3.73) holds is called the *exact hemiregularity bound* of F at $(\bar{x}, \bar{y})$ and is denoted by $\operatorname{hemireg} F(\bar{x}, \bar{y})$. Furthermore, F is *strongly* *metrically hemiregular* at $(\bar{x}, \bar{y})$ with modulus $\mu > 0$ if there are neighborhoods $U \subset X$ of $\bar{x}$ and $V \subset Y$ of $\bar{y}$ such that (3.73) holds and that F^{-1} admits a single-valued localization on $U \times V$ meaning that the mapping $y \mapsto F^{-1}(y) \cap U$ is single-valued on V.

(i) Show that a linear bounded operator $A\colon X \to Y$ is metrically hemiregular at every point $\bar{x} \in X$ if and only if it is surjective. In this case we have the relationships

$$\text{hemreg}\, A = \text{reg}\, A = \|(A^*)^{-1}\|,$$

where hemreg A stands for the common exact hemiregularity bound of A at all the points $\bar{x} \in X$. *Hint:* Proceed by the definitions.

(ii) Prove that $F\colon X \rightrightarrows Y$ is strongly hemiregular at $(\bar{x}, \bar{y})$ if and only if the inverse mapping $F^{-1}\colon Y \rightrightarrows X$ admits a calm single-valued localization $s(\cdot)$ around $(\bar{y}, \bar{x})$ with the equality between the corresponding exact bounds

$$\text{hemreg}\, F(\bar{x}, \bar{y}) = \text{clm}\, s(\bar{y}).$$

Hint: Proceed by the definitions and compare with the proof of [23, Proposition 5.8].

(iii) Given an example of a function $f\colon \mathbb{R}^2 \to \mathbb{R}$, which is metrically hemiregular at the origin while not being metrically regular around this point.

Exercise 3.59 (Coderivative Sum Rules in Infinite Dimensions). Let $F_i\colon X \rightrightarrows Y$, $i = 1, 2$, be set-valued mappings between Banach spaces with $(\bar{x}, \bar{y}) \in \text{gph}\,(F_1 + F_2)$.

(i) Suppose that the mapping F_1 is single-valued and Fréchet differentiable at $\bar{x}$. Then for all $y^* \in Y^*$ we have the equality

$$\widehat{D}^*(F_1 + F_2)(\bar{x}, \bar{y})(y^*) = \nabla F_1(\bar{x})^* y^* + \widehat{D}^* F_2\big(\bar{y} - F_1(\bar{x})\big)(y^*).$$

If furthermore F_1 is strictly differentiable at $\bar{x}$, then equality (3.24) holds for both limiting coderivatives $D^* = D^*_N, D^*_M$. *Hint:* To justify the inclusions "$\subset$" in the sum rules above for each case $\widehat{D}^*$, D^*_N, and D^*_M, proceed similarly to the proof of [522, Theorem 1.38] by using the corresponding definitions. To verify the opposite inclusions therein, apply the established ones "$\subset$" to the sum $(F_1 + F_2) + (-F_1)$.

(ii) Let X and Y be Asplund while F_1, F_2 be arbitrary (closed-graph) multifunctions. Fix $(\bar{y}_1, \bar{y}_2) \in S(\bar{x}, \bar{y})$ from (3.21) and suppose that this mapping is inner semicontinuous at $(\bar{x}, \bar{y}, \bar{y}_1, \bar{y}_2)$, that either F_1 is PSNC at $(\bar{x}, \bar{y}_1)$ or F_2 is PSNC at $(\bar{x}, \bar{y}_2)$ and that the qualification condition (3.22) is valid in terms of the mixed coderivative $D^* = D^*_M$. Then show that the sum rule (3.23) holds for *both* coderivatives $D^* = D^*_N, D^*_M$. Check that all the assumptions above are satisfied if either F_1 or F_2 is Lipschitz-like at the corresponding point $(\bar{x}, \bar{y}_i)$, $i = 1, 2$.

Hint: Proceed similarly to the proof of Theorem 3.9(i) with using the result of Exercise 2.42(i) and the subsequent limiting procedure under the imposed PSNC conditions. Then use (3.61) and Exercise 3.42(iii) for Lipschitz-like multifunctions. Compare this with the proof of [522, Theorem 3.8].

(iii) Clarify whether the N-regularity (resp. M-regularity) assumptions on both multifunction F_i at $(\bar{x}, \bar{y}_i)$ ensures the equality and the corresponding regularity property of $F_1 + F_2$ at $(\bar{x}, \bar{y})$.

(iv) Derive an infinite-dimensional counterpart of Theorem 3.9(ii).

Exercise 3.60 (Coderivative Intersection Rules). Let $F_1, F_2\colon X \rightrightarrows Y$ be set-valued mappings between Asplund spaces, let $(\bar{x}, \bar{y}) \in \text{gph}\, F_1 \cap \text{gph}\, F_2$, and let the graphical normal qualification condition

$$N\big((\bar{x}, \bar{y}); \text{gph}\, F_1\big) \cap \big[-N\big((\bar{x}, \bar{y}); \text{gph}\, F_2\big)\big] = \{0\}$$

be satisfied. Assume also that one of the mappings F_i, $i = 1, 2$, is SNC at $(\bar{x}, \bar{y})$. Then for all $y^* \in Y^*$, we have the inclusion

$$D^*_N(F_1 \cap F_2)(\bar{x}, \bar{y})(y^*) \subset \bigcup_{y_1^* + y_2^* = y^*} \Big[D^*_N F_1(\bar{x}, \bar{y})(y_1^*) + D^*_N F_2(\bar{x}, \bar{y})(y_2^*)\Big],$$

which holds as equality when both F_i are N-regular at $(\bar{x}, \bar{y})$. *Hint*: Apply the normal intersection rule from Theorem 2.16 to the sets $\Omega_i = \text{gph}\, F_i$, $i = 1, 2$, and its infinite-dimensional extension from Exercise 2.43(iv).

Exercise 3.61 (Chain Rules for Coderivatives in Infinite Dimensions). Let $G\colon X \rightrightarrows Y$ and $F\colon Y \rightrightarrows Z$ be (closed-graph) mappings between Asplund spaces, and let $\bar{z} \in (F \circ G)(\bar{x})$. Consider the set-valued mapping $S\colon X \times Z \rightrightarrows Y$ defined as in (3.25) and verify the following chain rule assertions:

(i) Given $\bar{y} \in S(\bar{x}, \bar{z})$, assume that S is inner semicontinuous at $(\bar{x}, \bar{z}, \bar{y})$, that either F is PSNC at $(\bar{y}, \bar{z})$ or G^{-1} is PSNC at $(\bar{y}, \bar{x})$, and that

$$D^*_M F(\bar{y}, \bar{z})(0) \cap \big(-D^*_M G^{-1}(\bar{y}, \bar{x})(0)\big) = \{0\},$$

which all hold if either F is Lipschitz-like around $(\bar{y}, \bar{z})$ or G is metrically regular around $(\bar{x}, \bar{y})$. Then for both coderivatives $D^* = D^*_N, D^*_M$ we have the inclusion

$$D^*(F \circ G)(\bar{x}, \bar{z})(z^*) \subset D^*_N G(\bar{x}, \bar{y}) \circ D^* F(\bar{y}, \bar{z})(z^*), \quad z^* \in Z^*.$$

Hint: Apply the corresponding coderivative sum rule of Exercise 3.59(ii) to the mapping $\Phi\colon X \times Y \rightrightarrows Z$ from (3.29). The validity of the imposed assumptions for the mentioned classes of F and G follows from Exercises 3.42(iii), 3.44(i), and 3.37.

(ii) Derive the Asplund space counterparts of Theorem 3.11(ii,iii) with the corresponding equality and regularity statements, and compare them with the results and proofs of [522, Theorem 3.13(ii,iii)].

(iii) Given $\bar{y} \in S(\bar{x}, \bar{z})$, assume that S is inner semicontinuous at $(\bar{x}, \bar{z}, \bar{y})$ and that F is Lipschitz-like around $(\bar{y}, \bar{z})$. Verify that

$$D^*_M(F \circ G)(\bar{x}, \bar{z})(0) \subset \big\{x^* \in X^* \big|\, x^* \in D^*_M G(\bar{x}, \bar{y})(0)\big\}.$$

Hint: Proceed in the way of proving the mixed coderivative sum rule in Exercise 3.59(ii) with employing Exercise 2.42(ii) and applying it Φ from (3.29). Then use the coderivative condition (3.60) for Lipschitz-like mappings before passing to the limit; see the proof of [522, Theorem 3.14] for more details.

Exercise 3.62 (Product Rules for Coderivatives in Finite and Infinite Dimensions). Let $F(x) := F_1(x) \times F_2(x)$ for all $x \in X$ with $F_i\colon X \rightrightarrows Y$, and let $\bar{y} := (\bar{y}_1, \bar{y}_2)$ with $\bar{y}_i \in F_i(\bar{x})$ as $i = 1, 2$.

(i) Assume that both spaces X and Y are finite-dimensional and that the qualification condition (3.22) is satisfied. Show that

$$D^* F(\bar{x}, \bar{y})(y^*) \subset D^* F_1(\bar{x}, \bar{y}_1)(y_1^*) + D^* F_2(\bar{x}, \bar{y}_2)(y_2^*) \text{ for all } y^* = (y_1^*, y_2^*) \in Y^* \times Y^*,$$

where the equality holds if each F_i is graphically regular at $(\bar{x}, \bar{y}_i)$, $i = 1, 2$. *Hint:* Following the proof of [202, Proposition 3.2], observe that $\text{gph}\, F = f^{-1}(\Theta)$ for

$$f(x, y) := f_1(x, y) \times f_2(x, y), \ \ f_i(x, y_i), \quad \text{and} \ \ \Theta := \text{gph}\, F_1 \times \text{gph}\, F_2$$

and apply the representations of the normals to inverse images from Corollary 3.13.

(ii) Extend the results in (i) to products of finitely many multifunctions.

(iii) Derive counterparts of (i) and (ii) for the case of Asplund spaces X and Y.

Exercise 3.63 (Partial Coderivatives). Consider a mapping $F\colon X \times Y \rightrightarrows Z$ between Asplund spaces with $(\bar{x}, \bar{y}, \bar{z}) \in \text{gph}\, F$. The *partial coderivative* $D^*_x F(\bar{x}, \bar{y}, \bar{z})$ of F with respect to x at $(\bar{x}, \bar{y}, \bar{z})$ is the coderivative of $F(\cdot, \bar{y})$ at $(\bar{x}, \bar{z})$. Impose the assumptions that F is PSNC at $(\bar{x}, \bar{y}, \bar{z})$ and that

$$(0, y^*) \in D^*_M F(\bar{x}, \bar{y}, \bar{z})(0) \Longrightarrow y^* = 0,$$

which are automatic when F is Lipschitz-like around $(\bar{x}, \bar{y}, \bar{z})$. Then prove that

$$D_x^* F(\bar{x}, \bar{y}, \bar{z})(z^*) \subset \text{proj}_x D^* F(\bar{x}, \bar{y}, \bar{z})(z^*), \quad z^* \in Z^*,$$

for both coderivatives $D^* = D_N^*, D_M^*$, where the symbol "proj_x" signifies the projection of the set $D^* F(\bar{x}, \bar{y}, \bar{z})(z^*) \subset X^* \times Y^*$ on X^*. Show furthermore that this inclusion holds as equality if F is N-regular (resp. M-regular) at $(\bar{x}, \bar{y}, \bar{z})$, which ensures the corresponding regularity property of $x \mapsto F(\bar{x}, \bar{y})$ at $(\bar{x}, \bar{z})$.

Hint: Apply Theorem 3.11(iii) and its Asplund space extensions from Exercise 3.61(ii) to the composition $F(\cdot, \bar{y}) = F \circ g$ with $g(x) := (x, \bar{y})$.

Exercise 3.64 (Basic Normals to Inverse Images in Infinite Dimensions). Let $\bar{x} \in G^{-1}(\Theta)$, where $G: X \rightrightarrows Y$ is a multifunction between Asplund spaces and where Θ is a nonempty subset of Y. Assume that the set-valued mapping $x \mapsto G(x) \cap \Theta$ is inner semicompact at $\bar{x}$ and that for every $\bar{y} \in G(\bar{x}) \cap \Theta$ the following hold: **(a)** Either G^{-1} is PSNC at $(\bar{y}, \bar{x})$, or Θ is SNC at $\bar{y}$. **(b)** The pair $\{G, \Theta\}$ satisfies the qualification condition

$$N(\bar{y}; \Theta) \cap \ker \widetilde{D}_M^* G(\bar{x}, \bar{y}) = \{0\}.$$

Prove that under these assumptions we have the inclusion

$$N(\bar{x}; G^{-1}(\Theta)) \subset \bigcup \Big[D_N^* G(\bar{x}, \bar{y})(y^*) \Big| \; y^* \in N(\bar{y}; \Theta), \; \bar{y} \in G(\bar{x}) \cap \Theta \Big],$$

which holds as equality if $G = g$ is single-valued and strictly differentiable at $\bar{x}$ and either the derivative operator $\nabla g(\bar{x}): X \to Y$ is surjective or Θ is normally regular at $\bar{x}$. Show also that in the latter case the set $g^{-1}(\Theta)$ is normally regular at $\bar{x}$.

Hint: Proceed as in the proof of Corollary 3.13 with employing the coderivative chain rule and regularity assertion from Exercise 3.61(ii).

Exercise 3.65 (Coderivatives of Special Compositions of Mappings Between Asplund Spaces). Derive infinite-dimensional versions of Theorem 3.14 and Corollary 3.15 proceeding in the same way as in the proofs therein while using the above coderivative calculus rules in the Asplund space settings. *Hint:* Compare this with [522, Theorem 3.18 and Corollary 3.19].

Exercise 3.66 (PSNC and SNC Properties of Mappings Under Summation). Let F_1, F_2 be closed-graph set-valued mappings between Asplund spaces X and Y, and let $(\bar{x}, \bar{y}) \in \text{gph}\,(F_1 + F_2)$. Assume that the mapping $S: X \times Y \rightrightarrows Y^2$ defined by (3.21) is inner semicompact at $(\bar{x}, \bar{y})$. Prove the following statements:

(i) If for every $(\bar{y}_1, \bar{y}_2) \in S(\bar{x}, \bar{y})$ each F_i is PSNC at $(\bar{x}, \bar{y}_i)$, respectively, and if the mixed qualification condition (3.82) is satisfied, then $F_1 + F_2$ is PSNC at $(\bar{x}, \bar{y})$.

(ii) If in the setting of (i) each F_i is SNC at $(\bar{x}, \bar{y}_i)$ and if the (normal) qualification condition (3.22) with $D^* = D_N^*$ is satisfied, then $F_1 + F_2$ is SNC at $(\bar{x}, \bar{y})$.

Hint: Proceed according to the definitions by applying the normal intersection rules from Exercises 2.42 and 2.43 for (i) and (ii), respectively. Compare this with the proofs of [522, Theorems 3.88 and 3.90] based on the extremal principle.

Exercise 3.67 (SNC Properties of Inverse Images of Sets Under Set-Valued Mappings Between Asplund Spaces). Consider the inverse image $G^{-1}(\Theta)$ of $\Theta \subset Y$ under a mapping $G: X \rightrightarrows Y$ between Asplund spaces. Which SNC/PSNC requirements do we need to impose on G and Θ to ensure the SNC property of $G^{-1}(\Theta)$ at $\bar{x}$ under an appropriate qualification condition of type (3.32)? *Hint*: Apply the result formulated in Exercise 2.42 and compare it with [522, Theorem 3.84].

Exercise 3.68 (PSNC and SNC Properties of Mappings Under Compositions). Consider the composition $F \circ G$ of set-valued mappings $G: X \rightrightarrows Y$ and $F: Y \rightrightarrows Z$ between Asplund spaces with $\bar{z} \in (F \circ G)(\bar{x})$, and assume that the mapping S from (3.25) is inner semicompact at $(\bar{x}, \bar{z})$. Prove the following statements:

(i) If for all $\bar{y} \in S(\bar{x}, \bar{z})$ both G and F are PSNC at $(\bar{x}, \bar{y})$ and $(\bar{y}, \bar{z})$, respectively, and if the qualification condition

$$D_M^* F(\bar{y}, \bar{z})(0) \cap \ker D_N^* G(\bar{x}, \bar{y}) = \{0\}$$

is satisfied, then the composition $F \circ G$ is PSNC at $(\bar{x}, \bar{z})$.

(ii) If for all $\bar{y} \in S(\bar{x}, \bar{z})$ both G and F are SNC at $(\bar{x}, \bar{y})$ and $(\bar{y}, \bar{z})$, respectively, and if the (normal) qualification condition (3.26) with $D^* = D_N^*$ is satisfied, then the composition $F \circ G$ is SNC at $(\bar{x}, \bar{z})$.

Hint: Apply intersection rules from Exercises 2.42 and 2.43 to the sets $\Omega_1 := \operatorname{gph} G \times Z$ and $\Omega_2 := X \times \operatorname{gph} F$. Compare this with [522, Theorems 3.95 and 3.98].

Exercise 3.69 (PSNC Property for Sets in Products of Two Spaces). Given a set $\Omega \subset X \times Y$ in the product of Banach spaces, we say that it is PSNC at $(\bar{x}, \bar{y}) \in X \times Y$ with respect to X if for any sequences $\varepsilon_k \downarrow 0$, $(x_k, y_k) \stackrel{\Omega}{\to} (\bar{x}, \bar{y})$ and $(x_k^*, y_k^*) \in \widehat{N}_{\varepsilon_k}((x_k, y_k); \Omega)$ we have the implication

$$\left[\|y_k^*\| \to 0,\ x_k^* \stackrel{w^*}{\to} 0\right] \Longrightarrow \|x_k^*\| \to 0 \ \text{ as } \ k \to \infty.$$

(i) Show that it is possible to equivalently put $\varepsilon_k \equiv 0$ if both spaces X, Y are Asplund. *Hint:* Use Exercise 1.42.

(ii) Prove in the Asplund space setting that for any (locally closed) sets $\Omega_1, \Omega_2 \subset X \times Y$ such that Ω_1 is SNC at $(\bar{x}, \bar{y}) \subset \Omega_1 \times \Omega_2$ and Ω_2 is PSNC at $(\bar{x}, \bar{y})$ with respect to X, we have the PSNC property of $\Omega_1 \cap \Omega_2$ at $(\bar{x}, \bar{y})$ with respect to X if

$$N(\bar{x}, \bar{y}); \Omega_1) \cap \big(- N(\bar{x}, \bar{y}); \Omega_2)\big) = \{(0, 0)\}.$$

Hint: Simplify the proof of [522, Theorem 3.79] based on the extremal principle.

Exercise 3.70 (Preservation of the Lipschitz-Like Property Under Various Operations). Obtain conditions ensuring the preservation of the Lipschitz-like property with exact bound relationships for set-valued mappings between Asplund spaces under the following operations:

(i) For compositions $F \circ G$ of $G \colon X \rightrightarrows Y$ and $F \colon Y \rightrightarrows Z$. *Hint:* Use the coderivative criterion for the Lipschitz-like property and the chain rules for coderivatives together with the corresponding PSNC calculus; cf. [522, Theorem 4.14].

(ii) For sums of mappings $F_1, F_2 \colon X \rightrightarrows Y$. *Hint:* Use the coderivative criterion for the Lipschitz-like property and the sum rules for the mixed coderivative together with the corresponding PSNC calculus presented above; cf. [522, Theorem 4.16].

Exercise 3.71 (Metric Regularity and Covering Under Compositions).

(i) In the framework of Exercise 3.70(i), obtain conditions ensuring the preservation of the metric regularity and covering properties with exact bound relationships. *Hint:* Apply Exercise 3.70(i) to the composition $(F \circ G)^{-1} = G^{-1} \circ F^{-1}$.

(ii) Could we proceed in the same way with sums $F_1 + F_2$?

Exercise 3.72 (Coderivatives of General Parametric Constraint Systems). Consider the class of PCS in form (3.37), where $g \colon X \times Y \to Z$ is a mapping between Banach spaces that is strictly differentiable at $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ with the surjective derivative $\nabla g(\bar{x}, \bar{y})$. Denoting $\bar{z} := g(\bar{x}, \bar{y}) \in \Theta$, show that:

(i) The normal coderivative of F is calculated by

$$D_N^* F(\bar{x}, \bar{y})(y^*) = \big\{x^* \in X^* \big|\ (x^*, -y^*) \in \nabla g(\bar{x}, \bar{y})^* N(\bar{z}; \Theta)\big\}, \tag{3.74}$$

where the above representation also holds true for the mixed coderivative $D_M^* F(\bar{x}, \bar{y})$ if $\dim Z < \infty$ (and obviously if $\dim Y < \infty$). *Hint:* Use the normal cone formula for inverse images from Exercise 1.54(ii), and compare with [522, Theorem 4.31(i)].

(ii) Formula (3.74) can be also for calculating the reversed mixed coderivative $\widetilde{D}^*_M F(\bar{x}, \bar{y})$ provided that, besides the trivial case of $\dim X < \infty$, Θ is DUALLY NORM-STABLE at $\bar{z}$ in the sense that $N(\bar{z}; \Theta) = N_{\|\cdot\|}(\bar{z}; \Theta)$, where

$$N_{\|\cdot\|}(\bar{z}; \Theta) := \Big\{z^* \in X^* \Big|\ \exists\, \varepsilon_k \downarrow 0,\, z_k \overset{\Theta}{\to} \bar{z},\, z_k^* \overset{\|\cdot\|}{\to} z^* \text{ with } z_k^* \in \widehat{N}_{\varepsilon_k}(z_k; \Theta),\ k \to \infty\Big\}.$$

Note that, besides the obvious case of $\dim Z < \infty$, every set Θ that is *normally regular* at $\bar{z}$ is dually norm-stable at this point. *Hint:* Compare with [277, Theorem 3.2].

(iii) Derive counterparts of formula (3.74) for all the three coderivatives under consideration in the case of Asplund spaces X, Y, and Z, where the surjectivity condition on $\nabla g(\bar{x}, \bar{y})$ is replaced by the constraint qualification

$$N(\bar{z}; \Theta) \cap \ker \nabla g(\bar{x}, \bar{y})^* = \{0\} \tag{3.75}$$

provided that Θ is normally regular at $\bar{z}$ and either it is SNC at $\bar{z}$ or g^{-1} is PSNC at $(\bar{z}, \bar{x}, \bar{y})$. *Hint:* Use representations of basic normals to inverse images from Exercise 3.64 together with SNC/PSNC preservation rules under composition from Exercise 3.68, and compare this with [522, Theorem 4.31(ii)] and [277, Theorem 3.2(ii)].

Exercise 3.73 (Coderivatives of Constraint Systems in Nonlinear Programming). Parametric constraint systems in nonlinear programming are given by

$$F(x) := \Big\{y \in Y \Big|\ \varphi_i(x, y) \le 0,\ i = 1, \ldots, m;\ \varphi_i(x, y) = 0,\ i = m+1, \ldots, m+r\Big\},$$

where all the functions φ_i, $i = 1, \ldots, m+r$, are strictly differentiable at the feasible point $(\bar{x}, \bar{y}) \in \operatorname{gph} F$. Denoting by

$$I(\bar{x}, \bar{y}) := \Big\{i \in \{1, \ldots, m\} \Big|\ \varphi_i(\bar{x}, \bar{y}) = 0\Big\}$$

the collection of active constraint indices, verify that all the three coderivatives $D^* = D^*_N, D^*_M, \widetilde{D}^*_M$ of F at $(\bar{x}, \bar{y})$ admit the representation

$$\begin{aligned} D^*F(\bar{x}, \bar{y})(y^*) = \Big\{x^* \in X^* \Big|\ & (x^*, -y^*) \in \sum_{i \in I(\bar{x}, \bar{y})} \lambda_i \nabla \varphi_i(\bar{x}, \bar{y}), \\ & \lambda_i \ge 0,\ i \in \{1, \ldots, m\} \cap I(\bar{x}, \bar{y})\Big\} \end{aligned}$$

with $y^* = (\lambda_1, \ldots, \lambda_{m+r}) \in \mathbb{R}^{m+r}$ in each of the following cases:

(i) Both spaces X and Y are Banach, and the *linear independence constraint qualification* (LICQ) holds at $(\bar{x}, \bar{y})$, i.e., the active constraint gradients $\nabla \varphi_i(\bar{x}, \bar{y})$, $i \in I(\bar{x}, \bar{y})$, are linearly independent in $X^* \times Y^*$.

(ii) Both spaces X and Y are Asplund and the MFCQ condition from Exercise 2.53 while with respect to (x, y) holds at $(\bar{x}, \bar{y})$.

Hint: Deduce this from Exercise 3.72(i,ii), respectively.

Exercise 3.74 (Coderivatives of Constraint Systems in Nondifferentiable Programming). Let F and $I(\bar{x}, \bar{y})$ with $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ be defined as in Exercise 3.73, and let the spaces X and Y be Asplund. Assume that all the functions φ_i, $i = 1, \ldots, m+r$, are locally Lipschitzian around $(\bar{x}, \bar{y})$ and that

$$\Big[\sum_{i \in I(\bar{x}, \bar{y})} \lambda_i (x_i^*, y_i^*) = 0\Big] \Longrightarrow \Big[\lambda_i = 0,\ i \in I(\bar{x}, \bar{y})\Big]$$

whenever $\lambda_i \ge 0$ for $i \in I(\bar{x}, \bar{y})$, $(x_i^*, y_i^*) \in \partial \varphi_i(\bar{x}, \bar{y})$ for $i \in \{1, \ldots, m\} \cap I(\bar{x}, \bar{y})$, and $(x_i^*, y_i^*) \in \partial \varphi_i(\bar{x}, \bar{y}) \cup \partial(-\varphi_i)(\bar{x}, \bar{y})$ for $i = m+1, \ldots, m+r$. Then we have

$$D_N^*F(\bar{x},\bar{y})(y^*)\subset\Big\{x^*\in X^*\Big|\;(x^*,-y^*)\in\sum_{i\in\{1,\ldots,m\}\cap I(\bar{x},\bar{y})}\lambda_i\partial\varphi_i(\bar{x},\bar{y})$$
$$+\sum_{i=m+1}^{m+r}\lambda_i\Big(\partial\varphi_i(\bar{x},\bar{y})\cup\partial(-\varphi_i)(\bar{x},\bar{y})\Big),\;\lambda_i\ge 0\;\text{ as }\;i\in I(\bar{x},\bar{y})\Big\}.$$

Hint: Deduce it from Exercise 3.64, where $\Theta\subset\mathbb{R}^{m+r}$ and $G\colon X\to\mathbb{R}^{m+r}$ are clearly defined by the constraint system under consideration, and use the subdifferential sum rule for Lipschitzian functions; compare it with [522, Corollary 4.36].

Exercise 3.75 (Coderivatives of Implicit Multifunctions). Consider the *implicit multifunction* defined by

$$F(x):=\big\{y\in Y\big|\;g(x,y)=0\big\},$$

where $g\colon X\times Y\to Z$ be a mapping between Banach spaces that is strictly differentiable at some point $(\bar{x},\bar{y})$ satisfying $g(\bar{x},\bar{y})=0$ with the derivative $\nabla g(\bar{x},\bar{y})$. Verify the coderivative representations

$$\widetilde{D}_M^*F(\bar{x},\bar{y})(y^*)=D_N^*F(\bar{x},\bar{y})(y^*)=\big\{x^*\in X^*\big|\;(x^*,-y^*)\in\nabla g(\bar{x},\bar{y})^*Z^*\big\}.$$

and show that the same representation holds for $D_M^*F(\bar{x},\bar{y})$ provided that either Y or Z is finite-dimensional. *Hint:* Deduce it from Exercise 3.72 with $\Theta=\{0\}$.

Exercise 3.76 (Coderivatives of Parametric Variational Systems). Consider the setting of Proposition 3.16 for PVS (3.34) and justify the following extensions of the results therein in finite and infinite dimensions:

(i) Formulas (3.38) and (3.39) hold true if the mapping $f\colon\mathbb{R}^n\to\mathbb{R}^m$ is merely strictly differentiable at $(\bar{x},\bar{y})$ with the full rank of $\nabla_x f(\bar{x},\bar{y})$.

(ii) Let $f\colon X\to Y$ be a mapping between arbitrary Banach spaces that is strictly differentiable at $(\bar{x},\bar{y})$ with the surjective derivative $\nabla_x f(\bar{x},\bar{y})$. Then the reversed mixed coderivative of PVS is calculated by

$$\widetilde{D}_M^*S(\bar{x},\bar{y})(y^*)=\Big\{x^*\in X^*\Big|\;\exists z^*\in Z^*\;\text{ with }\;x^*=\nabla_x f(\bar{x},\bar{y})^*z^*,$$
$$-y^*\in\nabla_y f(\bar{x},\bar{y})^*z^*+D_M^*Q(\bar{y},\bar{z})(z^*)\Big\}.$$

Furthermore, we have the relationship

$$\ker\widetilde{D}_M^*S(\bar{x},\bar{y})=-D_M^*Q(\bar{y},\bar{z})(0).$$

Hint: Proceeding as in the proof of Proposition 3.16, both assertions can be deduced from Exercise 3.72(ii) if either $\dim X<\infty$ or the set $\Theta=\operatorname{gph}Q$ is dually norm-stable at $(\bar{z},\bar{y})$. To avoid these assumptions, conduct a more delicate analysis involving [522, Lemma 1.16] and compare with [277, Theorem 4.1].

Exercise 3.77 (Second-Order Subdifferentials of Smooth Functions).

(i) Show that representations (3.43) with $u\in X^{**}$ are valid in any Banach space X if φ is continuously differentiable around $\bar{x}$ and its derivative mapping $x\mapsto\nabla\varphi(x)$ is strictly differentiable at this point.

(ii) Verify whether the results of (i) hold for the *normal* (resp. *mixed*) *second-order subdifferential* of φ at $\bar{x}$ relative to $\bar{x}^*\in\partial\varphi(\bar{x})$ defined, respectively, by

$$\partial_N^2\varphi(\bar{x},\bar{x}^*)(u):=\big(D_N^*\partial\varphi\big)(\bar{x},\bar{x}^*)(u),\quad u\in X^{**},\tag{3.76}$$

$$\partial_M^2\varphi(\bar{x},\bar{x}^*)(u):=\big(D_M^*\partial\varphi\big)(\bar{x},\bar{x}^*)(u),\quad u\in X^{**}.\tag{3.77}$$

Exercise 3.78 (Second-Order Subdifferential Chain Rules).

(i) Prove the second-order subdifferential chain rule (3.47) under the assumptions imposed in Proposition 3.18.

(ii) Justify the validity of the mixed second-order subdifferential counterpart of (3.47) formulated via (3.77) in arbitrary Banach spaces:

$$\partial^2_M\big(\psi\circ g\big)(\bar y,\bar q)(u)=\nabla^2\langle\bar v,g\rangle(\bar y)^*u+\nabla g(\bar y)^*\partial^2_M\psi(\bar w,\bar v)\big(\nabla g(\bar y)^{**}u\big),\quad u\in X^{**},$$

provided that g is $\mathcal{C}^2$-smooth around $\bar x$ with the surjective derivative.

(iii) Under which conditions the second-order chain rule above holds for the normal second-order subdifferential (3.76)?

Hint: Compare with the proof of [522, Theorem 1.127] with its simplification in the case of finite-dimensional spaces.

Exercise 3.79 (Reversed Mixed Coderivative of Subdifferential PVS with Composite Potentials). Consider the parametric variational system $S\colon X\rightrightarrows Y$ defined in form (3.41) by using the mappings $f\colon X\times Y\to Y^*$, $g\colon Y\to W$, and $\varphi\colon W\to\overline{\mathbb{R}}$ with arbitrary Banach spaces X, Y, and W. Assume that f is strictly differentiable at $(\bar x,\bar y)$ with the surjective partial derivative $\nabla_x f(\bar x,\bar y)$ and that g is $\mathcal{C}^2$-smooth around $\bar y$ with the surjective derivative $\nabla g(\bar y)$. Let $\bar v\in W^*$ be uniquely determined by (3.44) with $\bar q:=-f(\bar x,\bar y)\in\partial(\psi\circ g)(\bar y)$. Show that the reversed mixed coderivative of S at $(\bar x,\bar y)$ is calculated by

$$\begin{aligned}\widetilde{D}^*_M S(\bar x,\bar y)(y^*)=\Big\{x^*\in X^*\Big|\ &\exists u\in Y^{**}\ \text{ with }\ x^*=\nabla_x f(\bar x,\bar y)^*u,\\ -y^*\in\nabla_y f(\bar x,\bar y)^*u+\nabla^2\langle\bar v,g\rangle(\bar y)^*u&+\nabla g(\bar y)^*\partial^2_M\psi(\bar w,\bar v)\big(\nabla g(\bar y)^{**}u\big)\Big\},\end{aligned}$$

and furthermore we have the relationship

$$\ker\widetilde{D}^*_M S(\bar x,\bar y)=-\nabla g(\bar y)^*\partial^2_M\psi(\bar w,\bar v)(0).$$

Hint: Proceed as in the proof of Proposition 3.18 with the usage of the second-order chain rule from Exercise 3.78(ii).

Exercise 3.80 (Reversed Mixed Coderivative of Subdifferential PVS with Composite Fields). Consider the setting of Proposition 3.19 with the mappings $g\colon Y\to W$, $f\colon X\times Y\to W^*$, and $\psi\colon W\to\overline{\mathbb{R}}$ between Banach spaces under the surjectivity assumption on the partial derivative $\nabla_x f(\bar x,\bar y)$.

(i) Assume that $\nabla g(\bar y)$ is surjective, and show that we have

$$\begin{aligned}\widetilde{D}^*_M S(\bar x,\bar y)(y^*)=\Big\{x^*\in X^*\Big|\ &\exists u\in W^{**}\ \text{ with }\ x^*=\nabla_x f(\bar x,\bar y)^*u,\\ &-y^*\in\nabla_y f(\bar x,\bar y)^*u+\nabla g(\bar y)^*\partial^2_M\psi(\bar w,\bar q)(u)\Big\}\end{aligned}$$

for all $y^*\in Y^*$ with the additional relationship

$$\ker\widetilde{D}^*_M S(\bar x,\bar y)=-\nabla g(\bar y)^*\partial^2_M\varphi(\bar w,\bar q)(0).$$

(ii) Let the spaces X, Y and both spaces W and W^* be Asplund. Assume that the subgradient mapping $\partial\psi\colon W\rightrightarrows W^*$ is graph-closed around $(\bar w,\bar q)$, that PSNC is valid at this point, and that the second-order qualification condition is satisfied:

$$\partial^2_M\psi(\bar w,\bar q)(0)\cap\ker\nabla g(\bar y)^*=\big\{0\big\}.$$

Show that the inclusion "$\subset$" holds in both formulas presented in (i).

Hint: Combine the results of Exercise 3.76(ii) with the coderivative chain rule in the inclusion form from [522, Theorem 3.16].

Exercise 3.81 (Metric Regularity of General PCS). Let $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ for the parametric constraint system F from (3.37), where $g\colon X \times Y \to Z$ is a mapping between Banach spaces that is strictly differentiable at $(\bar{x}, \bar{y})$ with $\bar{z} := g(\bar{x}, \bar{y}) \in \Theta$.

(i) Assume that the derivative operator $\nabla g(\bar{x}, \bar{y})$ is surjective and that either Θ is dually normstable at $\bar{z}$ or $\dim X < \infty$. Show that the condition

$$(0, y^*) \in \nabla g(\bar{x}, \bar{y})^* N(\bar{z}; \Theta) \Longrightarrow y^* = 0 \tag{3.78}$$

is necessary for the metric regularity of F around $(\bar{x}, \bar{y})$ while being also sufficient for this property if either Y is Asplund and Θ is SNC at $\bar{z}$ or $\dim Y < \infty$. In the latter case we have the exact bound formula

$$\operatorname{reg} F(\bar{x}, \bar{y}) = \sup\Big\{\|y^*\| \;\Big|\; (x^*, -y^*) \in \nabla g(\bar{x}, \bar{y})^* N(\bar{z}; \Theta),\ \ \|x^*\| \le 1\Big\}. \tag{3.79}$$

(ii) Assume that all the spaces X, Y, and Z are Asplund, that the constraint qualification (3.75) is satisfied, and that Θ is SNC at $\bar{z}$. Then condition (3.78) is sufficient for the metric regularity of F around $(\bar{x}, \bar{y})$ while being also necessary for this property if Θ is normally regular at $\bar{z}$. If in addition $\dim Y < \infty$, we have (3.79).

Hint: Combine the results of Exercises 3.48 and 3.72.

Exercise 3.82 (Metric Regularity of Constraint Systems in Nonlinear Programming). Show that in the setting of Exercise 3.73(ii) the implication

$$\Big[\sum_{i \in I(\bar{x}, \bar{y})} \lambda_i \nabla_x \varphi_i(\bar{x}, \bar{y}) = 0\Big] \Longrightarrow \Big[\sum_{i \in I(\bar{x}, \bar{y})} \lambda_i \nabla_y \varphi_i(\bar{x}, \bar{y}) = 0\Big]$$

for $\lambda_i \ge 0$ if $i \in \{1, \ldots, m\} \cap I(\bar{x}, \bar{y})$ and $\lambda_i \in \mathbb{R}$ otherwise is necessary and sufficient for the metric regularity of F around $(\bar{x}, \bar{y})$. Furthermore, the exact bound formula

$$\operatorname{reg} F(\bar{x}, \bar{y}) = \max\Big\{\Big\|\sum_{i \in I(\bar{x}, \bar{y})} \lambda_i \nabla_y \varphi_i(\bar{x}, \bar{y})\Big\| \text{ subject to } \Big\|\sum_{i \in I(\bar{x}, \bar{y})} \lambda_i \nabla_x \varphi_i(\bar{x}, \bar{y})\Big\| \le 1\Big\}$$

holds, where λ_i satisfy the sign and complementary slackness conditions as above.

Hint: Combine the results presented in Exercises 3.48 and 3.73(ii). Verify that the maximum is realized in the exact bound formula due to the imposed Mangasarian-Fromovitz constraint qualification; cf. the proof of [522, Corollary 4.39].

Exercise 3.83 (Metric Regularity of Implicit Multifunctions). Consider the implicit multifunction F in the setting of Exercise 3.75, and show that the condition

$$\big[\nabla_x g(\bar{x}, \bar{y})^* z^* = 0\big] \Longrightarrow \big[\nabla_y g(\bar{x}, \bar{y})^* z^* = 0\big] \text{ whenever } z^* \in Z^*$$

is necessary and sufficient for the metric regularity of F around $(\bar{x}, \bar{y})$ provided that X is Asplund and that either Y is Asplund and $\dim Z < \infty$ or $\dim Y < \infty$. Verify that in the latter case we have the exact bound formula

$$\operatorname{reg} F(\bar{x}, \bar{y}) = \max\Big\{\big\|\nabla_y g(\bar{x}, \bar{y})^* z^*\big\| \;\Big|\; \big\|\nabla_x g(\bar{x}, \bar{y})^* z^*\big\| \le 1,\ z^* \in Z^*\Big\}.$$

Hint: Combine the results presented in Exercises 3.48 and 3.75.

Exercise 3.84 (Metric Regularity of General PVS in Asplund Spaces). Consider the general PVS in form (3.34), where $f\colon X \times Y \rightrightarrows Z$ is a mapping strictly differentiable at $(\bar{x}, \bar{y}) \in \operatorname{gph} S$ with the surjective partial derivative $\nabla_x f(\bar{x}, \bar{y})$ and where $Q\colon Y \rightrightarrows Z$ is closed-graph around $(\bar{y}, \bar{z})$ with $\bar{z} := -f(\bar{x}, \bar{y})$. Assume that both spaces X and Y are Asplund while Z is arbitrarily Banach.

(i) Show that S is metrically regular around $(\bar{x}, \bar{y})$ if and only if Q is PSNC at $(\bar{y}, \bar{z})$ and condition (3.54) holds with $D^* = D^*_M$, and therefore the metric regularity of S around $(\bar{x}, \bar{y})$ is equivalent to the Lipschitz-like property of Q around $(\bar{y}, \bar{z})$.

(ii) Verify the exact bound formula (3.55), with replacing "max" therein by "sup," provided that $\dim Y < \infty$ and Q is coderivatively normal at $(\bar{y}, \bar{z})$.

(iii) Find sufficient conditions ensuring that the maximum is attained in the exact bound formula for $\operatorname{reg} S(\bar{x}, \bar{y})$.

Hint: Use the Asplund space extensions of the coderivative criterion for metric regularity from Exercise 3.48 and the calculation of the reversed mixed coderivative in Exercise 3.76 together with the equivalence between the PSNC properties of Q around $(\bar{y}, \bar{z})$ and S^{-1} around $(\bar{y}, \bar{x})$. Compare it with [277, Theorem 5.6].

Exercise 3.85 (Metric Regularity and Subregularity of PVS in Banach Spaces). Consider the PVS setting of Exercise 3.84 in the case of arbitrary Banach spaces X and Y under the surjectivity assumption on $\nabla_x f(\bar{x}, \bar{y})$.

(i) Verify the equivalence between the metric regularity of S around $(\bar{x}, \bar{y})$ and the Lipschitz-like property of Q around $(\bar{y}, \bar{z})$.

(ii) Show that the equivalence holds true for the case of metric subregularity of S at $(\bar{x}, \bar{y})$ and the calmness property of Q at $(\bar{y}, \bar{z})$.

Hint: To justify both (i) and (ii), use the Lyusternik-Graves iterative process as in the alternative proof of Corollary 3.8; cf. [22, Theorem 3.3].

Exercise 3.86 (Metric Regularity of PVS with Composite Potentials in Infinite Dimensions). Based on the characterization of metric regularity of general PVS from Exercise 3.84(i) and the second-order subdifferential chain rule from Exercise 3.78(ii), derive an infinite-dimensional counterpart of Corollary 3.21.

Hint: Verify by using the second-order subdifferential chain rule from Exercise 3.78(ii) that the PSNC property of $Q = \partial(\psi \circ g)$ at $(\bar{y}, \bar{q})$ is equivalent to the PSNC property of $\partial\psi$ around $(\bar{w}, \bar{v})$.

Exercise 3.87 (Metric Regularity of PVS with Composite Fields in Infinite Dimension). Consider the setting of Exercise 3.80(i) under the additional assumptions that the spaces X and Y are Asplund and that the graph of $\partial\psi\colon W \rightrightarrows W^*$ is closed around $(\bar{w}, \bar{q})$. Show that S from (3.48) is metrically regular around $(\bar{x}, \bar{y})$ if and only if the subdifferential mapping $\partial\psi$ is Lipschitz-like around $(\bar{w}, \bar{q})$.

Hint: Verify that the PSNC property of $\partial\psi \circ g$ at $(\bar{y}, \bar{w})$ is equivalent to this property of $\partial\psi$ at $(\bar{w}, \bar{q})$ under the imposed surjectivity assumption on $\nabla g(\bar{y})$, and then use the coderivative criterion from Exercise 3.48(i) together with the expression of $\ker \widetilde{D}^*_M S(\bar{x}, \bar{y})$ presented in Exercise 3.80.

Exercise 3.88 (Some Properties of Amenable Functions). Let $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$, and let $\bar{x} \in \operatorname{dom} \varphi$. Show that the following hold:

(i) If φ is amenable or strongly amenable at $\bar{x}$, it maintains the corresponding property around this point.

(ii) If φ is amenable at $\bar{x}$, it is subdifferentially regular at this point.

(iii) The maximum of finitely many functions $\varphi_i \in \mathcal{C}^1$ is amenable at the corresponding point.
Hint: See [678, Section 10F].

Exercise 3.89 (Failure of Metric Regularity for PVS with Monotone Fields in Infinite-Dimensional Spaces).

(i) Based on the proof of Theorem 3.25 and the conclusions of Exercises 3.84(i) and 3.85(i), show that the result of Theorem 3.25 holds for PVS in Asplund and Banach spaces, respectively.

(ii) Give an example showing that a nonrobust counterpart of Theorem 3.25 for the metric subregularity of PVS as in Exercise 3.51 doesn't hold even for one-dimensional monotone mappings $Q\colon \mathbb{R} \rightrightarrows \mathbb{R}$.

Exercise 3.90 (Gâteaux Differentiability). Let $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ be finite at $\bar{x}$.

(i) Does the Gâteaux differentiability of φ imply its continuity at $\bar{x}$?

(ii) Show that the Gâteaux differentiability of φ at $\bar{x} \in \text{int}(\text{dom}\,\varphi)$ is equivalent to the Fréchet differentiability at this point provided that φ is locally Lipschitzian around $\bar{x}$. *Hint:* Compare with the proof in [537, Proposition 3.2].

(iii) Does the assertion in (ii) holds in infinite dimensions?

(iv) Assuming the convexity of φ, show that the Gâteaux differentiability of φ at $\bar{x} \in \text{int}(\text{dom}\,\varphi)$ is equivalent to its Fréchet differentiability at $\bar{x}$ and holds if and only if the subdifferential $\partial\varphi(\bar{x})$ is a singleton. *Hint:* Compare with [537, Theorem 3.3].

Exercise 3.91 (Metric Regularity and Subregularity of PVS with Convex Subdifferential Fields in Finite and Infinite Dimensions).

(i) Based on the proof of Corollary 3.26 and the conclusions of Exercises 3.84(i) and 3.85(i), establish extensions of the obtained result on the failure of metric regularity for subdifferential PVS with convex potentials without Gâteaux differentiability to the cases of Asplund and Banach spaces, respectively.

(ii) Give an example showing that the result of Corollary 3.26 doesn't hold if metric regularity is replaced by metric subregularity. *Hint:* Let $f(x, y) := x$, and let the field $Q\colon \mathbb{R} \rightrightarrows \mathbb{R}$ be given by

$$Q(y) := \begin{cases} \big[2^{-(k+1)}, 2^{-k}\big] & \text{for } \; y = 2^{-(k/3)}, \\ 2^{-(k+1)} & \text{for } \; y \in \big(2^{(-(k+1)/3)}, 2^{-(k/3)}\big), \\ 0 & \text{for } \; y = 0, \\ \big[-2^{-k}, -2^{-(k+1)}\big] & \text{for } \; y = -2^{-(k/3)}, \\ -2^{-(k+1)} & \text{for } \; y \in \big(-2^{-(k/3)}, -2^{(-(k+1)/3)}\big) \end{cases}$$

as depicted in Fig. 3.3. Verify that PVS (3.34) is *not metrically regular* around (0, 0) while it is *strongly q-subregular* of any order $q \in (0, 2]$ at this point; compare it with [564] and see Section 3.5 and also Chapter 5 for more discussions.

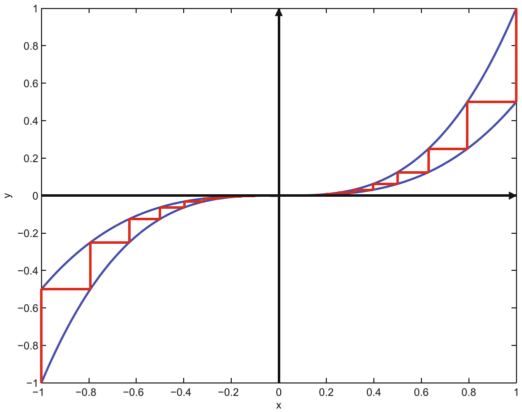


Fig. 3.3 Metric regularity vs. subregularity

Exercise 3.92 (Classes of Continuously Prox-Regular Functions). Prove that the following classes of functions $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ are continuously prox-regular:

(i) If φ is l.s.c. and convex, then this holds at any $\bar{x} \in \text{dom}\,\varphi$.

(ii) If φ is strongly amenable at $\bar{x}$, then this holds on a neighborhood of $\bar{x}$.

(iii) Let $\varphi \in \mathcal{C}^{1,1}$ on an open set U, i.e., it is continuously differentiable with the Lipschitz continuous gradient $\nabla\varphi$ on U. Then φ is continuously prox-regular on U.

Hint: Compare it with [678, Section 13.F].

Exercise 3.93 (Continuously Prox-Regular Functions with Lipschitz-Like Subdifferentials in Finite and Infinite Dimensions).

(i) Prove Lemma 3.28 by using the fact that the local single-valuedness of any mapping $F: \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is equivalent to the simultaneous validity of its *maximal hypomonotonicity* (see Chapter 5 below) and Lipschitz-like property around $(\bar{x}, \bar{y})$ with some $\bar{y} \in F(\bar{x})$. *Hint:* Compare it with [455].

(ii) Prove a Hilbert space extension of Lemma 3.28 by using properties of the *Moreau envelope* of $\varphi: X \to \overline{\mathbb{R}}$ defined, given a rate $\lambda > 0$, by

$$\varphi_\lambda(x) := \inf_{u \in X}\Big(\varphi(u) + \frac{1}{2\lambda}\|x - u\|^2\Big), \quad x \in X.$$

Hint: Compare it with the proof of [45, Theorem 5.3].

Exercise 3.94 (Failure of Metric Regularity for Subdifferential PVS with Composite Prox-Regular Potentials in Infinite Dimensions). Consider the class of subdifferential PVS with composite potentials (3.41), where $f: X \times Y \to Y^*$ is a mapping between Asplund spaces that is strictly differentiable at $(\bar{x}, \bar{y})$ with $-f(\bar{x}, \bar{y}) \in \partial(\psi \circ g)(\bar{y})$ and with the surjective partial derivative $\nabla_x f(\bar{x}, \bar{y})$, where $g: Y \to W$ is a $\mathcal{C}^2$-smooth mapping around $\bar{y}$ with the surjective derivative at $\bar{w} := \nabla g(\bar{y}) \in \operatorname{dom}\psi$, and where $\psi: W \to \overline{\mathbb{R}}$ is not Gâteaux differentiable at this point. Then the metric regularity of (3.41) around $(\bar{x}, \bar{y})$ fails in the following cases:

(i) W is Hilbert and ψ is continuously prox-regular at $\bar{w}$ for the subgradient $\bar{v} \in \partial\psi(\bar{w})$, which is uniquely determined by $\nabla g(\bar{y})^*\bar{v} = -f(\bar{x}, \bar{y})$. *Hint:* Proceed as in the proof of Theorem 3.29 with the usage of Exercise 3.93(ii).

(ii) W is Banach and ψ is convex and l.s.c. around $\bar{w}$. *Hint:* Using the equivalence from Exercise 3.85(i) together with the second-order chain from Exercise 3.78(ii), reduce the metric regularity of S in (3.41) around $(\bar{x}, \bar{y})$ to the Lipschitz-like property of the subdifferentiable mapping $\partial\psi$ around $(\bar{w}, \bar{v})$, which fails for the function ψ under the imposed assumptions; cf. Exercise 3.91(i).

Exercise 3.95 (Failure of Metric Regularity for Subdifferential PVS with Composite Fields in Infinite Dimensions). Let S be given by (3.48) with $(\bar{x}, \bar{y}) \in \operatorname{gph} S$, where $f: X \times Y \to W^*$ is strictly differentiable at $(\bar{x}, \bar{y})$ with the surjective partial derivative $\nabla_x f(\bar{x}, \bar{y})$, where $g: Y \to W$ is strictly differentiable $\bar{y}$ with the surjective derivative $g(\bar{y})$, and where $\psi: W \to \overline{\mathbb{R}}$ is not Gâteaux differentiable at $\bar{w} := g(\bar{w})$. Assume also that X and Y are Asplund. Then S which is not metrically regular around $(\bar{x}, \bar{y})$ fails in the following two cases:

(i) W is Hilbert and ψ is continuously prox-regular at $\bar{w}$ for the subgradient $\bar{q} := -f(\bar{x}, \bar{y}) \in \partial\psi(\bar{w})$. *Hint:* Proceed as in the proof of Theorem 3.31 with the usage of assertion (ii) from Exercise 3.93 instead of (i) therein.

(ii) W is Banach and ψ convex and l.s.c. around $\bar{w}$. *Hint:* Using the result of Exercise 3.87, reduce the metric regularity of S in question to the Lipschitz-like property of the convex subgradient mapping $\partial\psi$ around $(\bar{w}, \bar{q})$. Show that the latter fails by combining the results of Exercises 3.89(i) and 3.91(i).

3.5 Commentaries to Chapter 3

Section 3.1. The well-posedness properties discussed in Section 3.1 are fundamental in many areas of nonlinear analysis and its applications, particularly in those involving variational issues. In the commentaries to the author's book [522], the reader can find detailed discussions on the history of these notions, the genesis of ideas in their developments, and relationships between them that are reflected in Theorem 3.2. Some additional material, alternative terminology, and related properties can be found in the monographs [227, 375, 420, 678, 685]. The most recent one [375] contains a systematic study of regularity notions, their various aspects, and many applications in metric, Banach, and finite-dimensional spaces with a broad involvement of Ekeland's variational principle. However, some discussions presented therein are clearly and unfortunately biased, incomplete, and misleading; see, e.g., the corresponding commentaries to this and two previous chapters.

In what follows we mainly comment on some results in finite-dimensional and Banach spaces that are related to the contents of this book. The most impressive developments on well-posedness, which are not discussed in this book, include—from the author's viewpoint—the usage of *slopes* introduced in analysis by De Giorgi, Marino, and Tosques [191] and first brought to the theory of metric regularity and related topics by Azé, Corvellec, and Lucchetti [44] (preprint of 1998) and the research on various aspects of *directional metric regularity* that were initiated by Arutyunov and his collaborators (see, e.g., [24, 25, 27, 29]) and then were further developed in numerous publications as in [28, 282, 375, 613] and the references therein.

All the three equivalent well-posedness properties from Definition 3.1 have numerous applications in variational analysis and optimization including those presented in this book. As mentioned, we place this topic in one chapter with coderivative calculus due to the underlying coderivative criteria for well-posedness established in Theorem 3.3. The given proof of this theorem, which mainly follows the original proof in [507, Theorem 5.2] (precisely formulated in [505]) for the case of covering, is based on *variational arguments*, although there is no optimization involved in its statement. Note that the proof of the Lipschitzian part of this result given in the book by Rockafellar and Wets [678, Theorem 9.40] under the name of "Mordukhovich criterion" is much different from our proof while it is also based on optimization ideas married to finite-dimensional geometry. In the other direction, the *necessity* of the obtained coderivative characterization of well-posedness is crucial for the *coderivative calculus* developed in Section 3.2, since it allows us to reveal broad classes of mappings for which, e.g., the major sum and chain rules hold.

The coderivative criterion (3.6) of Theorem 3.3 with the *precise formula* (3.7) for the exact covering bound first appeared in the author's paper [505, Theorem 8] even in a more general form, although it was announced and discussed much earlier in seminar talks and private communications. In the beginning, this criterion came as a big surprise, to the degree of not accepting its correctness. It was probably related to the fact that the author's result concerned the covering property *around* the reference point but not *at* the point in question (what is now called "metric subregularity"—see below) as, e.g., in the book [378] and the subsequent papers [363, 366, 368], where sufficient conditions for the latter property and related nonrobust ones were obtained under certain assumptions on smooth and nonsmooth operators. Note that the *robustness* (*"around"*) *requirements* for covering and metric regularity properties as in Definition 3.1 were pioneered and strongly emphasized by A. A. Milyutin; see, e.g., [217] where a sufficient condition for the "covering in a neighborhood" property was obtained for single-valued Lipschitzian mappings in terms of Clarke's generalized gradient with mentioning the nonadequateness of the result obtained in such terms even in simple finite-dimensional cases of Lipschitzian operators.

Observe to this end that the crucial advantage of the coderivative criteria for well-posedness, in contrast to other known conditions in this direction formulated in terms of nonrobust constructions in primal and dual spaces, slopes, etc., is the presence of comprehensive *pointbased coderivatives calculus*, which is not available for the latter objects. This robust calculus allows us to deal with various composite models of optimization, variational analysis, and their applications as demonstrated, e.g., in [522, 523, 678] among numerous publications including this book.

In [505, 507] we also addressed a more general notion of *relative κ-covering* for a set-valued mapping $F: \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ with respect to another one $G: \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ as $x \in G(x)$ around $\bar{x}$ (in particular, with respect to a set Ω when $G(x) := x + \Omega$) formulated as follows: there is a neighborhood U of $\bar{x}$ such that inclusion (3.69) holds with some modulus $\kappa > 0$. While (3.69) can be immediately reformulated for set-valued mappings between metric spaces, we obtained in [505, Theorem 8] and [507, Theorem 5.3] in the case of finite dimensions (with the proof in vein of Theorem 3.3) its *coderivative characterization* in the form $\kappa(F, G, \bar{x}) > 0$ via the *rela-*

tive covering constant $\kappa(F, G, \bar{x})$ defined in Exercise 3.50. Furthermore, it has been established in [505, 507] that $\kappa(F, G, \bar{x})$ gives the *exact bound* of covering moduli $\kappa > 0$ in (3.69). Various modified versions of relative covering of set-valued mappings with respect to sets as well as interconnected notions of relative metric regularity have been recently studied in the literature; see, e.g., [25, 26, 27, 30, 86, 238, 372, 375, 522, 591, 613, 721, 789] for more details and references. Detailed comparisons between major notions of local and global covering properties have been recently done in [789, 790]; see Exercise 3.39 for the global covering and related versions.

As mentioned in Corollary 3.8, for *smooth* single-valued mappings $F = f$, the coderivative condition (3.6) reduces to the *surjectivity* (full rank in finite dimensions) of the derivative operator $\nabla f(\bar{x})$, which is the classical *Lyusternik-Graves regularity condition* discovered independently by Lyusternik [483] and Graves [305] for properties related, respectively, to metric regularity and covering/openness in the modern terminology. It follows from Theorem 3.3 that this condition is not only *sufficient* but also *necessary* for the properties under consideration. Moreover, we have the *exact bound formulas* for the corresponding moduli that had never been an issue in classical nonlinear analysis. Note that the *necessity statement* in Theorem 3.3 as well as in the "smooth" Corollary 3.8 is due (besides of the aforementioned robustness) to a *linear rate* of the covering and metric regularity properties that has revealed only in the modern framework of analysis. While in finite dimensions the necessity in Corollary 3.8 follows directly from Theorem 3.3 and the coderivative representation for smooth mappings, the Banach space version of this implication requires nontrivial considerations; see [522, Lemma 1.56 and Theorem 1.57].

Corollary 3.6 is a finite-dimensional version of the fundamental *Robinson-Ursescu theorem* for convex-graph mappings/convex processes between Banach spaces; see [658, 659, 726]. Similarly to the case of smooth mappings, the original contributions addressed to *sufficient* conditions for metric regularity and covering *without* paying attention to their *necessity* and *exact bound formulas*; see more comments in [522]. Further extensions of the Lyusternik-Graves and Robinson-Ursescu theorems can be found, e.g., in [161, 178, 222, 227, 239, 375, 448, 719, 721]. Remarkable applications of covering and metric regularity properties to *fixed* and *coincidence points* were given in [25, 26, 28, 87, 222, 239, 372, 501] and the references therein.

Let us mention to this end brand new applications of the covering property and machinery of variational analysis to *feedback stabilization* of dynamical (continuous-time and discrete-time) control systems obtained in the author's joint paper with Gupta, Jafari, and Kipka [316]. A seminal result in this direction is due to Brockett [129] who proved, via degree theoretic topological techniques, that the *openness* property (3.4) of a smooth mapping $f\colon \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ below is *necessary* for local asymptotic stabilization of the nonlinear *ODE control system*

$$\dot{x} = f(x, u), \;\; t \geq 0, \tag{3.80}$$

by means of continuous stationary feedback laws. As well recognized, the openness property of f fails to be sufficient for such a stabilization. It is shown in [316], via variational techniques, that replacing openness by *linear openness*/covering/metric regularity of f allows us to obtain efficient conditions on the system data and linear openness moduli supporting the *sufficiency* in Brockett's theorem and providing local *exponential* stabilization of (3.80) by means of continuous stationary feedback regulators. In this way new conditions ensuring the *necessity* of linear openness for both local exponential and asymptotic stabilization of (3.80) by means of stationary continuous as well as smooth feedback laws are derived in [316]. Some counterparts of these results are established in [316] by the developed variational approach for asymptotic feedback stabilization of nonlinear *discrete-time* control systems.

The coderivative criterion of the *Lipschitz-like* property in Theorem 3.3(iii) first appeared in [508] with the proof essentially different from [505, 507] for covering/metric regularity characterizations; see also [513]. Note that the obtained coderivative characterization strongly departs from the previous sufficient conditions for this "pseudo-Lipschitzian" property given by Aubin [35] and Rockafellar [674] in terms of Clarke's normal cone to the graph. As discussed in Remark 3.4(iii), for single-valued and major classes of set-valued mappings the latter conditions hold in fact *only in smooth settings*. Comprehensive treatments of the well-posedness properties, their nonlocal versions, and their coderivative characterizations in finite dimensions were given in [510]. As demonstrated in [522, 523, 678], such *dual-space* characterizations play a fundamental role in many aspects of variational analysis due to their *robustness* and full *coderivative calculus*. We refer to [227, 375, 420] to different characterizations of well-posedness via generalized derivatives in *primal spaces*, which may lack robustness and calculus rules but still can be useful in certain settings. Some combined *primal-dual* characterizations of well-posedness were given in [280, 281, 282, 630, 631]. Applications of well-posedness criteria to *inverse* and *implicit* (*multi*)*functions* can be found, e.g., in [23, 30, 114, 227, 238, 287, 382, 420, 444, 445, 450, 453, 454, 653, 664, 725, 751].

The most recognized and useful *nonrobust* Lipschitzian behavior of multifunctions is the *upper Lipschitzian* property of F *at* $\bar{x}$ introduced by Robinson [661] in form (3.3) with $u = \bar{x}$ and $V = \mathbb{R}^m$. It is often called nowadays the *calmness* of F at $\bar{x}$. The same name is associated with its graphical version at $(\bar{x}, \bar{y}) \in \operatorname{gph} F$, which is written as (3.3) with $u = \bar{x}$ being equivalent to the *metric subregularity* (the term coined by Dontchev and Rockafellar [678]) of the inverse mapping. The latter property, known also as "regularity at a point," goes back to Ioffe and Tikhomirov [378] in the case of single-valued mappings, while its full set-valued version is due to Ye and Ye [746] called there "pseudo-upper Lipschitz continuity." The fundamental result by Robinson [663] justifies the validity of the upper Lipschitzian property for *piecewise polyhedral* mappings between finite-dimensional spaces. Crucial contributions to the study and applications of these properties to broad classes of optimization and equilibrium problems have been made by Henrion, Outrata, and their collaborators; see, e.g., [287, 335, 337, 338, 340, 341, 622]. In particular, Henrion and Outrata were the first [337] to obtain efficient coderivative/subdifferential conditions for calmness of multifunctions in terms of our basic constructions. Observe close relationships of calmness and metric subregularity properties with *error bounds* in optimization, which go back to Hoffman [353] for linear inequality systems; see [42, 256, 363, 375, 434, 466, 467, 607, 612] for more recent developments. We refer the reader to, e.g., [18, 22, 23, 86, 138, 139, 209, 216, 227, 240, 260, 280, 290, 299, 327, 375, 376, 384, 414, 420, 454, 485, 594, 595, 693, 719, 720, 723, 745, 770, 776, 779, 782] for numerous results in these directions and applications to variational problems. An interesting survey of recent results on metric subregularity utilizing normal cones and coderivatives is given by Zheng [772].

The aforementioned nonrobustness of calmness and metric subregularity doesn't allow to develop adequate calculus/preservation results for them, their stability with respect to perturbation, and restricts therefore the scope of their applications. In the joint paper with Gfrerer [285], we introduced a certain *uniform* metric subregularity property for solution maps to parametric constraint systems

$$g(x, p) \in C \subset \mathbb{R}^m \quad \text{with} \quad x \in \mathbb{R}^n \quad \text{and} \quad p \in P, \tag{3.81}$$

where the set C is closed and the perturbation parameter p belongs to a topological space P. Such a stability property has been actually considered by Robinson [660] in the case where C is a convex cone, and so we labeled this property in [285] as the *Robinson stability* of (3.81). The paper [285] contains verifiable first-order and second-order conditions ensuring the Robinson regularity of (3.81) and its robustness in the classes of perturbations under consideration.

Note further that *q-versions* of both metric regularity and subregularity properties (as well as their other well-posedness equivalents) have been also considered in the literature [158, 274, 419, 276, 283, 465, 435, 752, 765, 780], where the main attention and valuable applications were given for the Hölder case of $0 < q \le 1$. It is easy to see that there is no sense to consider the case of $q > 1$ for metric q-regularity, since only constant mappings satisfy estimate (3.2) with replacing $\operatorname{dist}(y, F(x))$ by $\operatorname{dist}^q(y, F(x))$ when $q > 1$. But it is different for *q-subregularity* with $y = \bar{y}$ therein, where the case of $q > 1$ is nontrivial and occurs to be important for both variational theory and applications. The general case of metric q-subregularity and its *strong q-subregularity* counterpart whenever $q > 0$ has been recently studied by the author and Ouyang [564] with characterizing these properties for subdifferential variational systems. Furthermore, using strong q-subregularity with $q > 1$ in [564] allowed us to obtain *higher convergent rates* in *quasi-Newton methods* of solving generalized equations in comparison with the corresponding results by Dontchev [221] who extended the celebrated Dennis-Moré theorem [192] for nonlinear equations; see also the recent paper [160] for further extensions and various applications. Note also that similar conclusions on better convergence rates were deduced in the joint paper with Li [465] for the *proximal point method* to find zero s of maximal monotone operators in Hilbert spaces under the Hölder metric q-subregularity with $0 < q < 1$.

It is worth mentioning yet another nonrobust *metric hemiregularity* property of set-valued mappings defined in Exercise 3.58, together with its *strong* counterpart, following the paper by Aragón and Mordukhovich [23], where it was studied and applied to deriving enhanced versions of implicit multifunction theorems and stability of generalized equations. Hemiregularity can be viewed as a symmetric counterpart of subregularity with fixing the *domain* point $\bar{x}$ instead of the *range* one $\bar{y}$. As mentioned in the final version of [23], hemiregularity was independently examined by Kruger, in his extended study [433] of various well-posedness properties of set-valued mappings, under the name of "metric semiregularity." The latter name was earlier used by Pühl and Schirotzek [650] for a completely different regularity property; see also the book [685, Section 10.6]. To avoid confusions, we coined the "hemiregularity" terminology in [23]. The inverse property to metric hemiregularity was defined (but not investigated) by Klatte and Kummer [420, p. 10] as "Lipschitz lower semicontinuity," while the inverse one to strong metric hemiregularity was designated in [23] as the "calm single-valued localization" discussed in Exercise 3.58(ii). Note the quite recent study by Uderzo [725] containing, in particular, a new implicit multifunction theorem under hemiregularity (complemented to the one in [23]) and applications to exact penalization in constrained optimization.

Next let us comment on *infinite-dimensional extensions* of the *coderivative characterizations* of well-posedness properties in Theorem 3.3. As in [522], in infinite dimensions we distinguish between two types of characterizations of well-posedness: *neighborhood* and *pointbased* (sometimes called "pointwise" or "point") ones. The former criteria involve not only the point in question but a neighborhood of it, while the latter ones have the pointbased form of Theorem 3.3 but under additional assumptions that automatically hold in finite dimensions.

Neighborhood characterizations of the covering property were initiated by Kruger for mappings between *Fréchet smooth* spaces; see [432]. His results of the *dual* nature (the first one was announced in [429] for locally Lipschitzian functions by using ε-subdifferentials of type (1.34) as $\varepsilon > 0$) were formulated in terms of several neighborhood constants defined via two-parametric constructions depending on ε and the neighborhood size. The author and Shao [578] essentially improved such characterizations by using merely the regular/Fréchet coderivative $\widehat{D}^*F$ (i.e., with $\varepsilon = 0$) in *Asplund spaces* and also established their counterparts for nonlocal well-posedness properties. Sufficient neighborhood conditions with corresponding modulus estimates in terms of other subdifferential constructions in the suitable "trustworthy" Banach spaces were derived by Ioffe; see [375] and the references therein. We also mention *primal-space* neighborhood developments by Kummer [444, 445] (presented in his book with Klatte [420]) via the so-called Ekeland's points.

Comprehensive *pointbased* extensions of the coderivative characterizations of well-posedness in Theorem 3.3 were obtained by the author [514] for closed-graph mappings $F\colon X \rightrightarrows Y$ between *Asplund* spaces (with necessary conditions holding in any *Banach* spaces) in terms of the *mixed coderivative* (1.65) and the *PSNC property* of F at $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ defined in (3.65). For the case of Lipschitzian behavior the obtained characterizations are presented in Exercise 3.44; see also [522, Theorem 4.10] for more information. Observe that the exact bound formula (3.10) is split now into the two inequalities in (3.67), and thus the full infinite-dimensional counterpart of the coderivative criterion of Theorem 3.3 requires imposing the coderivative normality from Exercise 3.43 introduced and studied in [522]. We refer the reader to [581, 582] for previous pointbased *sufficient* conditions of well-posedness for mappings between Asplund spaces in terms of the normal coderivative D^*_N as well as to [371, 375, 397, 398, 402, 635, 637] for related sufficiency results based on other coderivatives in suitable Banach spaces and the corresponding counterparts of the PSNC property discussed in more detail in [522]. Observe also that the papers by Jourani and Thibault [402] and by Ioffe [371] contained *necessary* pointbased conditions for well-posedness expressed in terms of the "approximate coderivative" (cf. Section 1.5) and addressed, in the case of the Lipschitz-like property, to set-valued mappings $F\colon X \rightrightarrows Y$ with a *Banach* domain space X and a *finite-dimensional* image space Y. However, the exact pointbased bound estimates of type (3.67) were not obtained in the aforementioned publications.

A very interesting recent development has been done by Clason and Valkonen [169] on applications of the author's coderivative criterion [510] to the study of *stability of saddle points* for a broad class of constrained optimization problems in Hilbert spaces, including inverse problems with PDE constraints. A main ingredient of their approach was to reduce the infinite-dimensional setting under consideration to a finite-dimensional one by using *pointwise subdifferentiation* of *integral functionals* and thus avoiding any SNC-type assumptions. In this way they obtained explicit stability conditions for various infinite-dimensional problems arising, in particular, in parameter identification, image processing, and PDE-constrained optimization.

We conclude the commentaries to this section by mentioning some applications of the discussed well-posedness properties of mappings and their generalized derivative characterizations to *numerical aspects* of variational analysis and the convergence of algorithms, which can be found [19, 71, 72, 73, 206, 227, 260, 344, 354, 384, 420, 421, 458, 480, 566, 665, 727] and the references therein. Note also the pioneering paper by Dontchev, Lewis, and Rockafellar [224] who established relationships between *well-posedness* and *ill-posedness* properties via calculating the *radius of metric regularity* by using the author's coderivative characterizations and the exact bound formulas from Theorem 3.3. In this way they justified connections between the radius of metric regularity and Renegar's "distance to infeasiblity" [657] arising in complexity theory for linear and conic programming. We also refer the reader to [227, 375, 522] for more comments and references on infinite-dimensional extensions.

Section 3.2. The coderivative calculus results presented in this section were first established by the author [511] for general multifunctions between finite-dimensional spaces based on the extremal principle. The sum rule of Theorem 3.9 was derived earlier in the author's paper [509] by direct applying the method of metric approximations. This sum rule as well as the chain rule of Theorem 3.11(ii) were then reproduced in [678] by using another device.

The infinite-dimensional setting is more diverse and comprehensively presented in the author's book [522] mostly dealing with multifunctions between Asplund spaces and mainly based on the previous publications [514, 532, 581, 584, 588]. Although the proofs in [522] went in the same direction as in [511], the techniques were more involved and the spectrum of the results obtained, and the assumptions imposed were essentially broader while reducing to [511] in finite dimensions. In [522] we developed parallel calculus rules for *both normal and mixed coderivatives* with the qualification conditions expressed in terms of the *mixed coderivative* having an essential advantage in comparison with the normal one from this viewpoint; see Exercises 3.59–3.61. In particular, we have the following *sum rule* for both coderivatives $D^* = D^*_N, D^*_M$ of the set-valued mappings F_1, F_2 between Asplund spaces:

$$D^*(F_1 + F_2)(\bar{x}, \bar{y})(y^*) \subset \bigcup_{(\bar{y}_1, \bar{y}_2) \in S(\bar{x}, \bar{y})} \Big[D^* F_1(\bar{x}, \bar{y}_1)(y^*) + D^* F_2(\bar{x}, \bar{y}_2)(y^*)\Big], \; y^* \in Y^*,$$

at $(\bar{x}, \bar{y}) \in \operatorname{gph}(F_1 + F_2)$ provided that the mapping S from (3.21) is *inner semicompact* at $(\bar{x}, \bar{y})$ as defined in Exercise 2.46, that

$$D^*_M F_1(\bar{x}, \bar{y}_1)(0)) \cap \big(- D^*_M F_2(\bar{x}, \bar{y}_2)\big) = \{0\}, \quad (\bar{y}_1, \bar{y}_2) \in S(\bar{x}, \bar{y}), \tag{3.82}$$

and that for each $(\bar{y}_1, \bar{y}_2) \in S(\bar{x}, \bar{y})$ either F_1 is PSNC at $(\bar{x}, \bar{y}_1)$ or F_2 is PSNC at $(\bar{x}, \bar{y}_2)$. It follows from (the necessity part of) the infinite-dimensional coderivative characterizations of well-posedness discussed above in Exercise 3.48 and the commentaries to Section 3.1 that both the mixed qualification condition (3.82) and the PSNC assumptions are satisfied if *either one* from F_1, F_2 is *Lipschitz-like* around the corresponding points of $S(\bar{x}, \bar{y})$. More diverse results of the inclusion and equality types were developed in [522] for coderivative *chain rules* and their consequences for set-valued and single-valued mappings between infinite-dimensional spaces.

Some of the discussed coderivative chain and sum rules for mappings between Asplund (or Fréchet smooth) spaces were then reproduced in the books [114, 637, 685], where the reader could also find certain “fuzzy” (neighborhood) versions. A number of calculus rules for (normal) coderivatives generated by normal cones/subdifferentials of other types were considered, e.g., in [369, 375, 400, 401, 533, 585] in appropriate Banach spaces. We specially emphasize the results by Ioffe [369, 371] and Jourani and Thibault [400, 401] who developed extended coderivative calculus rules for the normal coderivatives generated by various approximate subdifferential constructions (see Section 1.5) in arbitrary Banach spaces.

Section 3.3. This section is devoted to coderivative analysis and some applications of the parametric variational systems described as (3.34) in finite dimensions with infinite-dimensional extensions discussed in Section 3.4. The well-developed coderivative calculus plays a crucial role in this analysis and applications. The presented results are mostly taken from the book [522, Section 4.4] (coderivative calculus) and the subsequent papers [277, 524] (applications to metric regularity).

Coderivative calculus opens the gate to effectively evaluate the basic coderivative (1.15)—as well as its normal and mixed versions in infinite dimensions—of the parametric variation systems (3.34), which is the solution map to Robinson’s generalized equation (3.35). It allows us to implement these calculations to the obtained coderivative criteria for the well-posedness properties from Section 3.1 and their infinite-dimensional extensions, together with evaluating the corresponding exact bounds of moduli. In contrast to [522], where the main attention was drawn to Lipschitzian stability, now we mostly concentrate on metric regularity. The metric regularity property for (3.34) occurs to be more involved in comparison with its Lipschitz-like counterpart for PVS and often *fails* in the most natural PVS settings as shown in Subsection 3.3.3. This phenomenon was first revealed by the author in [524] and then was investigated in various publications dealing with different classes of PVS in both finite and infinite dimensions; see, e.g., [22, 41, 45, 155, 277, 404, 666, 719]. The proof of Proposition 3.24 follows [223]; cf. also [227, Theorem 3G.5].

The results of Section 3.3 and the exercises for it concerning the structural PVS with composite potentials (3.41) and composite fields (3.48) utilize the notion of the *second-order subdifferential* (or generalized Hessian) (3.42) introduced by the author in [508] and then developed and applied in numerous publications. The original motivation came from using the coderivative criterion and calculus rules to derive verifiable conditions for Lipschitzian stability of parametric variational systems; see [508, 509, 510, 511, 512, 513, 514]. Then important early contributions were made by Rockafellar and his collaborators [456, 642] to studying the remarkable notions of *tilt* and *full stability* of local minimizers in finite-dimensional optimization introduced therein; see also the book by Bonnans and Shapiro [96] concerning tilt perturbations and related *quadratic/second-order growth conditions*. Note that second-order growth conditions of this type were used in the pioneering paper by Zhang and Treiman [771] in connection with the upper Lipschitzian property of the inverse to the basic subdifferential of l.s.c. functions in finite dimensions while providing a complete characterization of the latter property for l.s.c. convex functions; see the exact formulations in Exercise 3.55. Similar growth conditions were developed by Aragón and Geoffroy [20, 21] for characterizations of metric regularity and subregularity properties as well as their strong counterparts for the convex subdifferential in Hilbert and Asplund spaces; see also Mordukhovich and Nghia [551] as well as Exercises 3.52 and 5.27. Some nonconvex versions for the basic subgradient mapping with quantitative interrelations between the corresponding constants were obtained in [551] and the joint author’s paper with Drusvyatskiy and Nghia [232] in the Asplund space setting; see Sections 5.3 and 5.4. Note that quite recent algorithmic applications of second-order growth conditions for metric subregularity of subgradient mappings arising in semidefinite programming have been developed by Cui, Sun, and Toh [186].

In recent years, the area of second-order variational analysis involving (3.42) and associated second-order constructions has drawn a strong and steadily growing attention with well-developed calculus and great many applications to various classes of finite- and infinite-dimensional problems in optimization, stability, equilibrium, control, mechanics, economics, electronic, etc., as well as to practical models described via first-order gradient/subgradient and normal cone data.

Among the most impressive developments and applications of this “dual-space” direction in second-order variational analysis, we mention the following (this list is by far incomplete): well-developed second-order calculus [517, 522, 539, 557, 558, 570, 625]; calculations of the second-order subdifferentials for maximum functions as well as for separable piecewise-$\mathcal{C}^2$, extended piecewise linear, and extended piecewise linear-quadratic functions [252, 557, 570, 573, 574]; calculations of the second-order subdifferentials of the indicator function (i.e., coderivatives of the normal cone) for polyhedral, generalized polyhedral, and polyhedric convex sets in finite and infinite dimensions [49, 225, 336, 342, 552, 741]; second-order characterizations of convexity and monotonicity properties of functions and mappings [150, 152, 153, 154, 555]; evaluations of the second-order constructions for broad classes of nonpolyhedral moving sets appeared in numerous applications [2, 5, 6, 8, 82, 143, 144, 151, 172, 174, 208, 286, 288, 339, 340, 361, 599, 621, 627, 654, 747]; characterizations of Robinson's strong regularity for variational systems [225, 336, 556, 570, 571, 742]; characterizations of tilt stability of local minimizers in various classes of optimization problems [230, 232, 243, 284, 461, 551, 554, 559, 743, 765, 780, 781]; characterizations of full Lipschitzian and Hölderian stability of local minimizers in nonlinear programming, conic programming, and optimal control [456, 552, 556, 563, 573, 571]; characterizations of full stability for solutions to general and particular classes of parametric variational systems [555]; characterizations of metric regularity, subregularity, and their strong counterparts for first-order subdifferentials and their relationships with second-order growth conditions [21, 43, 230, 232, 551, 554, 733, 734, 780]; characterizations of Kojima's strong stability of variational systems [552, 556, 576]; sensitivity and stability analysis with respect to robust and nonrobust properties of solution maps for constrained optimization, variational and quasivariational inequalities, and equilibrium problems [283, 285, 286, 287, 328, 336, 337, 338, 335, 339, 340, 451, 453, 454, 508, 513, 522, 558, 561, 562, 575, 576, 625, 654, 742]; characterizations of weak sharp minimizers and their stable higher-order extensions [780, 781]; no-gap second-order necessary and sufficient optimality conditions for some classes of constrained and vector optimization problems [358, 359]; necessary optimality and stationarity conditions for mathematical programming and control problems with equilibrium constraints [3, 78, 267, 314, 338, 341, 346, 290, 523, 620, 623, 745, 746, 780]; necessary optimality and stationarity conditions for equilibrium problems with equilibrium constraints [340, 342, 560, 622]; stability of discrete approximations and necessary optimality conditions for controlled sweeping processes [5, 127, 143, 144, 172, 173, 174]; stability and optimization of PDE systems [169, 346, 347, 623, 730]; qualitative characteristics of nonconvex gradient flows and nonlinear evolution equations [497, 681]; characterizations of critical multipliers in variational systems with eliminating slow convergence of primal-dual methods in optimization [577]; applications to second-order cone programming, semidefinite programming, circular cone programming, and second-order complementarity [208, 289, 390, 561, 562, 563, 747, 766, 784]; applications to bilevel programs, bilevel optimal control, and hierarchical optimization [51, 79, 80, 198, 199, 200, 202, 341, 767, 768, 769]; applications to viability issues for dynamical systems [266]; applications to numerical methods of optimization (proximal, trust-region, quasi-Newton ones, etc.) [21, 231, 451, 465, 539, 564, 652]; applications to various problems in mechanics [2, 174, 423, 557, 621]; applications to stochastic analysis and optimization [340, 626, 739]; applications to economic modeling [560, 622]; applications to electronics [6, 8, 127]; applications to micromagnetics and related topics [423]; applications to electricity spot markets [38, 340, 342]; applications to the crowd motion model of traffic flow [144]; etc. Second-order variational analysis and its applications are the subjects of the author's book in progress [527].

The class of prox-regular and subdifferentially continuous functions and their amenable subclasses were introduced by Poliquin and Rockafellar [641] in finite dimensions while playing a crucial role in second-order variational analysis. Besides the publications on second-order analysis listed above, prox-regular functions and associated sets as well as their favorable subclasses have been intensively studied and applied in a vast literature that covers both finite-dimensional and infinite-dimensional settings; see, e.g., [7, 45, 83, 175, 332, 678] and the references therein. Note that the notion of prox-regularity for closed sets in finite-dimensional spaces is equivalent to the notion well known in geometric measure theory as sets of *positive reach* introduced and largely studied by Federer [263].

Section 3.4. This section mostly presents some notions and results, which are extensions of the basic material (including approaches and proof techniques) developed in the main sections of this chapter in finite-dimensional spaces. The reader can find more discussions and references in the exercise hints. We particularly draw the reader's attention to the notions of *relative covering* and *semimetric regularity* of set-valued mappings from Exercises 3.50 and 3.56 introduced and partly investigated in the early author's work [505, 507] that require further developments and applications in both finite-dimensional and infinite-dimensional (Banach, metric) spaces. The properties of metric and strong metric *hemiregularity* defined in Exercise 3.58 and discussed in the commentaries above are also largely underinvestigated from both viewpoints of efficient certifications and applications. Yet another topic of profound interest concerns the notion of *outer derivative* of set-valued mappings between finite-dimensional spaces (3.70) introduced by Zhang and Treiman [771] who used it for the pointbased characterization of the upper Lipschitzian property of multifunctions and then applications to inverse subdifferential mappings presented in Exercise 3.55. As mentioned above, the upper Lipschitzian property—together with its calmness and metric subregularity counterparts—has recently drawn much attention in variational theory and numerous applications, and thus developing these lines of research would be very important in variational analysis. Some properties of the outer derivative are obtained in [771] (see, in particular, Exercise 3.54), but this is definitely not sufficient for desired applications, which require more developed calculus rules for (3.70) and its graphical version at $(\bar{x}, \bar{y}) \in \text{gph}\, F$ needed for the study of calmness and metric subregularity of set-valued mappings.

Chapter 4
First-Order Subdifferential Calculus

This chapter concerns generalized differential properties of *extended-real-valued functions* $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ that are assumed, unless otherwise stated, to be *lower semicontinuous* around references points. Our main purpose here is to develop comprehensive calculus rules for the *basic subdifferential* (1.24) and *singular subdifferential* (1.25) of such functions. Recall that general sum rules for them have been obtained in Section 2.4 as consequences of the intersection rule for basic normals; these results can be also deduced from the coderivative sum rules of Theorem 3.9. In this chapter we concentrate on deriving other major results of first-order subdifferential calculus including subdifferentiation of marginal/optimal value functions, general chain rules with their implementations to subdifferentiation of products, quotients, minimum, and maximum functions, and various versions of the subdifferential mean value theorem with some applications to variational analysis in nonsmooth settings.

4.1 Subdifferentiation of Marginal Functions

In this section we focus on evaluating basic and singular subgradients for a broad class of the *marginal functions* defined by

$$\mu(x) := \inf\big\{\varphi(x, y)\big|\ y \in G(x)\big\}, \tag{4.1}$$

where $\varphi\colon \mathbb{R}^n \times \mathbb{R}^m \to \overline{\mathbb{R}}$ is a (l.s.c.) extended-real-valued function and where $G\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is a (closed-graph) set-valued constraint mapping that described parameter-dependent/moving constraint sets. The given marginal function (4.1) can be interpreted, in particular, as the (optimal) *value function* in the problem of *parametric optimization* given by

$$\text{minimize}\ \ \varphi(x, y)\ \ \text{subject to}\ \ y \in G(x)$$

B. S. Mordukhovich, *Variational Analysis and Applications*, Springer Monographs in Mathematics, https://doi.org/10.1007/978-3-319-92775-6_4

with the *cost/objective* function φ and the *constraint* multifunction G, where y and x are the decision and parameter variables, respectively. A characteristic feature of marginal functions of type (4.1) is their *intrinsic nondifferentiability* regardless the smoothness of cost functions and the simplicity of moving constraint sets. As we'll see below, constructive evaluations of both the basic and singular subdifferentials under consideration are crucial for resolving major issues of subdifferential calculus as well as for deriving optimality conditions in various classes of optimization problems, sensitivity and stability analysis, as well as numerous applications in variational and nonvariational settings.

To evaluate basic and singular subgradients of the marginal function (4.1), define the *argminimum mapping* $M\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ by

$$M(x) := \big\{y \in G(x)\big|\ \varphi(x, y) = \mu(x)\big\} \tag{4.2}$$

and obtain subdifferential results of two kinds depending on either inner semicontinuity or local boundedness assumptions imposed on the mapping M.

Theorem 4.1 (Basic and Singular Subgradients of Marginal Functions). *For the marginal function* (4.1) *with* $\bar{x} \in \operatorname{dom}\mu$ *the following hold:*

(i) *Fix* $\bar{y} \in M(\bar{x})$ *from* (4.2)*, and suppose that* M *is inner semicontinuous at* $(\bar{x}, \bar{y})$ *and that the qualification condition*

$$\partial^\infty\varphi(\bar{x}, \bar{y}) \cap \big[-N\big((\bar{x}, \bar{y}); \operatorname{gph} G\big)\big] = \{0\} \tag{4.3}$$

is satisfied. Then we have the subdifferential upper estimates

$$\partial\mu(\bar{x}) \subset \bigcup_{(x^*, y^*)\in\partial\varphi(\bar{x},\bar{y})} \Big[x^* + D^*G(\bar{x}, \bar{y})(y^*)\Big], \tag{4.4}$$

$$\partial^\infty\mu(\bar{x}) \subset \bigcup_{(x^*, y^*)\in\partial^\infty\varphi(\bar{x},\bar{y})} \Big[x^* + D^*G(\bar{x}, \bar{y})(y^*)\Big]. \tag{4.5}$$

(ii) *Let the argminimum mapping* (4.2) *be locally bounded around* $\bar{x}$ *with* $M(\bar{x}) \neq \emptyset$*, and let condition* (4.3) *hold for any* $\bar{y} \in M(\bar{x})$*. Then*

$$\partial\mu(\bar{x}) \subset \bigcup_{\substack{(x^*, y^*)\in\partial\varphi(\bar{x},\bar{y})\\ \bar{y}\in M(\bar{x})}} \Big[x^* + D^*G(\bar{x}, \bar{y})(y^*)\Big],$$

$$\partial^\infty\mu(\bar{x}) \subset \bigcup_{\substack{(x^*, y^*)\in\partial^\infty\varphi(\bar{x},\bar{y})\\ \bar{y}\in M(\bar{x})}} \Big[x^* + D^*G(\bar{x}, \bar{y})(y^*)\Big].$$

Proof. To justify assertion (i), consider the function

$$\vartheta(x,y) := \varphi(x,y) + \delta\big((x,y);\operatorname{gph} G\big), \quad (x,y) \in \mathbb{R}^n \times \mathbb{R}^m,$$

and verify in its terms the subdifferential upper estimates

$$\partial\mu(\bar{x}) \subset \big\{x^* \big|\ (x^*,0) \in \partial\vartheta(\bar{x},\bar{y})\big\}, \quad \partial^\infty\mu(\bar{x}) \subset \big\{x^* \big|\ (x^*,0) \in \partial^\infty\vartheta(\bar{x},\bar{y})\big\}.$$

We first prove the upper estimate for $\partial\mu(\bar{x})$. Pick $x^* \in \partial\mu(\bar{x})$ and by (1.37) find sequences $x_k \stackrel{\mu}{\to} \bar{x}$ and $x_k^* \to x^*$ with $x_k^* \in \widehat{\partial}\mu(x_k)$, $k \in \mathbb{N}$. Hence for any $\varepsilon_k \downarrow 0$ there exists $\eta_k \downarrow 0$ such that for each fixed number $k \in \mathbb{N}$ we have

$$\langle x_k^*, x - x_k\rangle \le \mu(x) - \mu(x_k) + \varepsilon_k\|x - x_k\| \ \text{ whenever } \ x \in x_k + \eta_k\mathbb{B}.$$

This implies by the constructions of μ, ϑ, and M that

$$\langle (x_k^*,0),(x,y) - (x_k,y_k)\rangle \le \vartheta(x,y) - \vartheta(x_k,y_k) + \varepsilon_k(\|x - x_k\| + \|y - y_k\|)$$

for all $y_k \in M(x_k)$ and $(x,y) \in (x_k,y_k) + \eta_k\mathbb{B}$. Thus $(x_k^*,0) \in \widehat{\partial}_{\varepsilon_k}\vartheta(x_k,y_k)$. Since M is inner semicontinuous at $(\bar{x},\bar{y})$, we get a sequence of $y_k \in M(x_k)$ converging to $\bar{y}$. Observe that $\vartheta(x_k,y_k) \to \vartheta(\bar{x},\bar{y})$ due to $\mu(x_k) \to \mu(\bar{x})$, which therefore ensures that $(x^*,0) \in \partial\vartheta(\bar{x},\bar{y})$ by passing to the limit as $k \to \infty$ and so justify the claimed inclusion for $\partial\mu(\bar{x})$ via $\partial\vartheta(\bar{x},\bar{y})$. To deduce from here the subdifferential estimate (4.4), we apply to the sum in ϑ the basic subdifferential sum rule (2.35) under the qualification condition (2.34), which reduces in this case to the one assumed in (4.3).

To verify further the inclusion for $\partial^\infty\mu(\bar{x})$ via $\partial^\infty\vartheta(\bar{x},\bar{y})$, pick a singular subgradient $x^* \in \partial^\infty\mu(\bar{x})$, take any $\varepsilon_k \downarrow 0$, and by (1.38) find sequences $x_k \stackrel{\mu}{\to} \bar{x}$, $(x_k^*,\nu_k) \to (x^*,0)$, and $\eta_k \downarrow 0$ satisfying

$$\langle x_k^*, x - x_k\rangle + \nu_k(\alpha - \alpha_k) \le \varepsilon_k(\|x - x_k\| + |\alpha - \alpha_k|)$$

for all $(x,\alpha) \in \operatorname{epi}\mu$, $x \in x_k + \eta_k\mathbb{B}$, and $|\alpha - \alpha_k| \le \eta_k$. The inner semicontinuity of (4.2) ensures the existence of $y_k \xrightarrow{M(x_k)} \bar{y}$ and $\alpha_k \downarrow \vartheta(\bar{x})$ such that

$$(x_k^*,0,\nu_k) \in \widehat{N}_{\varepsilon_k}\big((x_k,y_k,\alpha_k);\operatorname{epi}\vartheta\big), \quad k \in \mathbb{N},$$

which gives us by passing to the limit as $k \to \infty$ that $(x^*,0) \in \partial^\infty\vartheta(\bar{x})$. We finish the proof of (i) by applying the singular subdifferential sum rule (2.36) to the sum in ϑ under the validity of the qualification condition (4.3). The verification of assertion (ii) is similar to the above. △

The main assumption of Theorem 4.1 is the qualification condition (4.3), which in fact holds in the following major settings due to the obtained coderivative/subdifferential characterizations of well-posedness. For brevity we discuss the qualification condition (4.3), which in fact holds in the following major settings due to the obtained coderivative/subdifferential characterizations this only in case (i) of Theorem 4.1.

Corollary 4.2 (Marginal Functions with Lipschitzian and Metrically Regular Data). *Given $\bar{y} \in M(\bar{x})$, suppose that the argminimum mapping (4.2) is inner semicontinuous at $(\bar{x}, \bar{y})$ and that either φ is locally Lipschitzian around $(\bar{x}, \bar{y})$ or $\varphi = \varphi(y)$ and G is metrically regular around $(\bar{x}, \bar{y})$. Then both inclusions (4.4) and (4.5) are satisfied.*

Proof. If φ is locally Lipschitzian around $(\bar{x}, \bar{y})$, we have $\partial^\infty \varphi(\bar{x}, \bar{y}) = \{0\}$ by Theorem 1.22, and thus (4.3) holds. For $\varphi = \varphi(y)$, the qualification condition (4.3) can be equivalently written as

$$\partial^\infty \varphi(\bar{y}) \cap \ker D^* G(\bar{x}, \bar{y}) = \{0\}$$

and hence holds by Theorem 3.3 if G is metrically regular around $(\bar{x}, \bar{y})$. △

Another useful consequence of Theorem 4.1 provides efficient conditions for locally *Lipschitz continuity* of a general class of marginal functions.

Corollary 4.3 (Local Lipschitz Continuity of Marginal Functions). *The following assertions hold for the class of marginal functions μ from (4.1):*

(i) *Assume that the argminimum mapping (4.2) is inner semicontinuous at some point $(\bar{x}, \bar{y}) \in \operatorname{gph} M$ and that the cost function φ is locally Lipschitzian around this point. Then μ is Lipschitz continuous around $\bar{x}$ provided that it is l.s.c. around $\bar{x}$ and that G is Lipschitz-like around $(\bar{x}, \bar{y})$.*

(ii) *Assume that M in (4.2) is locally bounded around $\bar{x} \in \operatorname{dom} M$ and that φ is locally Lipschitzian around $(\bar{x}, \bar{y})$ for any $\bar{y} \in M(\bar{x})$. Then μ is Lipschitz continuous around $\bar{x}$ provided that it is l.s.c. around this point and that G is Lipschitz-like around $(\bar{x}; \bar{y})$ whenever $\bar{y} \in M(\bar{x})$.*

Proof. It is sufficient to verify assertion (i), since the proof of (ii) is similar. The assumed local Lipschitz continuity of φ ensures the validity of the qualification condition (4.3) and reduces (4.5) to

$$\partial^\infty \mu(\bar{x}) \subset D^* G(\bar{x}, \bar{y})(0).$$

It follows from Theorem 3.3(iii) that $D^* G(\bar{x}, \bar{y})(0) = \{0\}$ by the Lipschitz-like property of G around $(\bar{x}, \bar{y})$. Thus $\partial^\infty \mu(\bar{x}) = \{0\}$, which yields the Lipschitz continuity of μ around $\bar{x}$ by Theorem 1.22. △

The next theorem, which can also be treated as a consequence of Theorem 4.1 with some elaborations, concerns subdifferentiation of the *infimal convolution* defined for two functions $\varphi_1, \varphi_2 \colon \mathbb{R}^n \to \overline{\mathbb{R}}$ by

$$\big(\varphi_1 \oplus \varphi_2\big)(x) := \inf\big\{\varphi_1(x_1) + \varphi_2(x_2)\big|\ x_1 + x_2 = x\big\}. \tag{4.6}$$

Let us associate with (4.6) the *convolution mapping* $C \colon \mathbb{R}^n \rightrightarrows \mathbb{R}^{2n}$ given by

$$C(x) := \big\{(x_1, x_2)\big|\ x_1 + x_2 = x,\ \varphi_1(x_1) + \varphi_2(x_2) = \big(\varphi_1 \oplus \varphi_2\big)(x)\big\}. \tag{4.7}$$

Theorem 4.4 (Subdifferentiation of Infimal Convolutions). *Given a point $\bar{x} \in \operatorname{dom} C$ for the mapping C from (4.7), the following assertions hold:*

(i) *Fix $(\bar{x}_1, \bar{x}_2) \in C(\bar{x})$, and assume that the convolution mapping (4.7) is inner semicontinuous at $(\bar{x}, \bar{x}_1, \bar{x}_2)$. Then we have the inclusions*

$$\partial\big(\varphi_1 \oplus \varphi_2\big)(\bar{x}) \subset \partial\varphi_1(\bar{x}_1) \cap \partial\varphi_2(\bar{x}_2),$$

$$\partial^\infty\big(\varphi_1 \oplus \varphi_2\big)(\bar{x}) \subset \partial^\infty\varphi_1(\bar{x}_1) \cap \partial^\infty\varphi_2(\bar{x}_2).$$

(ii) *If the convolution mapping (4.7) is locally bounded around $\bar{x}$, then*

$$\partial\big(\varphi_1 \oplus \varphi_2\big)(\bar{x}) \subset \bigcup_{(\bar{x}_1,\bar{x}_2)\in C(\bar{x})} \partial\varphi_1(\bar{x}_1) \cap \partial\varphi_2(\bar{x}_2),$$

$$\partial^\infty\big(\varphi_1 \oplus \varphi_2\big)(\bar{x}) \subset \bigcup_{(\bar{x}_1,\bar{x}_2)\in C(\bar{x})} \partial^\infty\varphi_1(\bar{x}_1) \cap \partial^\infty\varphi_2(\bar{x}_2).$$

Proof. It is sufficient to justify assertion (i) while noting that the proof of (ii) is similar. It follows from definition (4.6) that the infimal convolution admits the marginal function representation:

$$\big(\varphi_1 \oplus \varphi_2\big)(x) = \inf\big\{\varphi(x, x_1, x_2)\big|\ (x_1, x_2) \in G(x)\big\}, \quad x \in \mathbb{R}^n, \tag{4.8}$$

where $\varphi\colon \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \to \overline{\mathbb{R}}$ and $G\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n \times \mathbb{R}^n$ are given, respectively, by

$$\varphi(x, x_1, x_2) := \varphi_1(x_1) + \varphi_2(x_2), \quad G(x) := \big\{(x_1, x_2) \in \mathbb{R}^{2n}\big|\ x_1 + x_2 = x\big\},$$

and where the argminimum mapping (4.2) reduces to (4.7) in this case. To check now the qualification condition (4.3), observe that

$$\partial^\infty\varphi(\bar{x}, \bar{x}_1, \bar{x}_2) = \big(0, \partial^\infty\varphi_1(\bar{x}_1), \partial^\infty\varphi_2(\bar{x}_2)\big) \text{ and}$$

$$N\big((\bar{x}, \bar{x}_1, \bar{x}_2)); \operatorname{gph} G\big) = \big\{(v, -v, -v) \in \mathbb{R}^{3n}\big|\ v \in \mathbb{R}^n\big\},$$

and so (4.3) holds in the framework of (4.8). The latter formula yields

$$D^*G(\bar{x}, \bar{x}_1, \bar{x}_2)(v_1, v_2) = \begin{cases} \{v_1\} & \text{if } v_1 = v_2, \\ \emptyset & \text{otherwise.} \end{cases}$$

Substituting this into (4.4) and (4.5) with taking into account that

$$\partial\varphi(\bar{x}, \bar{x}_1, \bar{x}_2) = (0, \partial\varphi_1(\bar{x}_1), \partial\varphi_2(\bar{x}_2)\big),$$

we arrive at the claimed representations in (i). △

4.2 Subdifferentiation of Compositions

When the mapping $G = g\colon \mathbb{R}^n \to \mathbb{R}^m$ is single-valued in (4.1), the marginal function reduces to the (generalized) *composition*

$$\big(\varphi \circ g\big)(x) := \varphi\big(x, g(x)\big), \quad x \in \mathbb{R}^n, \tag{4.9}$$

and thus we can deduce from Theorem 4.1 extended *subdifferential chain rules* and their various consequences. The next theorem gives us also some cases of *equalities* and *subdifferential regularity* of compositions, which seem to be specifically related to single-valuedness of the constraint mapping in (4.1). Note that the first part of this theorem holds in the Asplund space setting, while the second part is valid in any Banach spaces; see Exercise 4.28.

Theorem 4.5 (Basic and Singular Subdifferentials of General Compositions). *Consider composition* (4.9) *with an extended-real-valued function $\varphi\colon \mathbb{R}^n \times \mathbb{R}^m \to \overline{\mathbb{R}}$ and a mapping $g\colon \mathbb{R}^n \to \mathbb{R}^m$ that is locally Lipschitzian around $\bar{x}$ with $\bar{y} = g(\bar{x})$. The following assertions hold:*

(i) *The qualification condition* (4.3) *with $G = g$ ensures the validity of the subdifferential upper estimates*

$$\partial\big(\varphi \circ g\big)(\bar{x}) \subset \bigcup_{(x^*, y^*) \in \partial\varphi(\bar{x}, \bar{y})} \Big[x^* + \partial\langle y^*, g\rangle(\bar{x})\Big], \tag{4.10}$$

$$\partial^\infty\big(\varphi \circ g\big)(\bar{x}) \subset \bigcup_{(x^*, y^*) \in \partial^\infty\varphi(\bar{x}, \bar{y})} \Big[x^* + \partial\langle y^*, g\rangle(\bar{x})\Big] \tag{4.11}$$

with the equality in (4.10) *if either the outer function φ is of class $\mathcal{C}^1$ around $(\bar{x}, \bar{y})$ or it is lower regular at $(\bar{x}, \bar{y})$ and the inner mapping g is of class $\mathcal{C}^1$ around $\bar{x}$; in the latter case, the composition $\varphi \circ g$ is lower regular at $\bar{x}$.*

(ii) *If φ is strictly differentiable at $(\bar{x}, \bar{y})$, then we always have the equality*

$$\partial\big(\varphi \circ g\big)(\bar{x}) = \nabla_x\varphi(\bar{x}, \bar{y}) + \partial\big\langle\nabla_y\varphi(\bar{x}, \bar{y}), g\big\rangle(\bar{x}). \tag{4.12}$$

Proof. To justify (i), observe that inclusions (4.10) and (4.11) reduce to (4.4) and (4.5), respectively, for locally Lipschitzian mappings g due to the scalarization formula of Theorem 1.32. We get furthermore the equalities

$$\partial\mu(\bar{x}) = \big\{x^* \big| \; (x^*, 0) \in \partial\vartheta(\bar{x}, \bar{y})\big\}, \quad \partial^\infty\mu(\bar{x}) = \big\{x^* \big| \; (x^*, 0) \in \partial^\infty\vartheta(\bar{x}, \bar{y})\big\}$$

via the function ϑ defined in the proof of Theorem 4.1, provided that $G = g$ is locally Lipschitzian around $\bar{x}$ without any additional assumptions. This can be verified similarly to the proof of Theorem 3.11(iii). Then the equality and regularity statements in (i) follow by applying the corresponding results of Proposition 1.30 and Theorem 2.19 to the sum form of ϑ.

To prove now assertion (ii), take an arbitrary sequence $\gamma_j \downarrow 0$ and get, by the assumed strict differentiability (1.19) of φ at $(\bar{x}, \bar{y})$, such $\eta_j \downarrow 0$ that

$$\begin{aligned}&\big|\varphi\big(u, g(u)\big) - \varphi\big(x, g(x)\big) - \big\langle \nabla_x\varphi(\bar{x}, \bar{y}), u - x\big\rangle - \big\langle \nabla_y\varphi(\bar{x}, \bar{y}), g(u) - g(x)\big\rangle\big|\\ &\le \gamma_j\big(\|u - x\| + \|g(u) - g(x)\|\big) \ \text{ for all } \ x, u \in B_{\eta_j}(\bar{x}), \quad j \in \mathbb{N}.\end{aligned}$$

Pick further $x^* \in \partial(\varphi\circ g)(\bar{x})$, and find by the first representation in (1.37) sequences $x_k \to \bar{x}$ and $x_k^* \to x^*$ with $x_k^* \in \widehat{\partial}(\varphi \circ g)(x_k)$, $k \in \mathbb{N}$. This allows us to select a subsequence $k_j \to \infty$ as $j \to \infty$ so that $\|x_{k_j} - \bar{x}\| \le \eta_j/2$ and

$$\varphi\big(x, g(x)\big) - \varphi\big(x_{k_j}, g(x_{k_j})\big) - \langle x^*_{k_j}, x - x_{k_j}\rangle \ge -\varepsilon_{k_j}\|x - x_{k_j}\|$$

whenever $\varepsilon_{k_j} \downarrow 0$ as $j \to \infty$ and $x \in x_{k_j} + (\eta_j/2)\mathbb{B}$. Combining the relationships above gives us the estimate

$$\begin{aligned}&\big\langle \nabla_y\varphi(\bar{x}, \bar{y}), g(x) - g(x_{k_j})\big\rangle - \big\langle x^*_{k_j} - \nabla_x\varphi(\bar{x}, \bar{y}), x - x_{k_j}\big\rangle\\ &\ge -\big[\varepsilon_{k_j} + \gamma_j(\ell + 1)\big]\|x - x_{k_j}\| \ \text{ for } \ x \in x_{k_j} + (\eta_j/2)\mathbb{B},\end{aligned}$$

where ℓ is a Lipschitz constant of g around $\bar{x}$. This yields

$$x^*_{k_j} - \nabla_x\varphi(\bar{x}, \bar{y}) \in \widehat{\partial}_{\nu_j}\big\langle \nabla_y\varphi(\bar{x}, \bar{y}), g\big\rangle(x_{k_j}) \ \text{ with } \ \nu_j := \varepsilon_{k_j} + \gamma_j(\ell + 1),$$

which ensures the validity of the inclusion "$\subset$" in (4.12) by passing to the limit as $j \to \infty$ and using the second representation in (1.37). To verify the opposite inclusion in (4.12), it suffices to employ the similar arguments to the above starting with an arbitrary subgradient $x^* \in \partial\langle\nabla_y\varphi(\bar{x}, \bar{y}), g\rangle(\bar{x})$. △

Next we derive several remarkable consequences of Theorem 4.5; see also exercises in Section 2.5 for more results in this direction. Let us start with the *chain rules* of the inclusion type for basic and singular subgradients of the standard compositions $\varphi \circ g = \varphi(g(x))$ in (4.9).

Corollary 4.6 (Chain Rules for Basic and Singular Subgradients). *Let $\varphi\colon \mathbb{R}^m \to \overline{\mathbb{R}}$ do not depend on the first variable in* (4.9)*, and let $g\colon \mathbb{R}^n \to \mathbb{R}^m$ be locally Lipschitzian around $\bar{x}$. Impose the qualification condition*

$$\partial^\infty\varphi(\bar{y}) \cap \ker \partial\langle\cdot, g\rangle(\bar{x}) = \{0\}.$$

Then we have the subdifferential chain rules

$$\partial(\varphi \circ g)(\bar{x}) \subset \bigcup_{y^* \in \partial\varphi(\bar{y})} \partial\langle y^*, g\rangle(\bar{x}), \quad \partial^\infty(\varphi \circ g)(\bar{x}) \subset \bigcup_{y^* \in \partial^\infty\varphi(\bar{y})} \partial\langle y^*, g\rangle(\bar{x}).$$

Proof. It follows from the above that the qualification condition (4.3) reduces to the one imposed here if $\varphi = \varphi(y)$ and $G = g$ is locally Lipschitzian. Then the claimed chain rules are specifications of (4.10) and (4.11). △

The next two corollaries of Theorem 4.5 present subdifferential *product* and *quotient rules* in inclusion and equality forms.

Corollary 4.7 (Subdifferential Product Rules). *Let $\varphi_i\colon \mathbb{R}^n \to \overline{\mathbb{R}}$, $i = 1, 2$, be Lipschitz continuous around $\bar{x}$. The we have the product rules*

$$\partial(\varphi_1 \cdot \varphi_2)(\bar{x}) = \partial\big(\varphi_2(\bar{x})\varphi_1 + \varphi_1(\bar{x})\varphi_2\big)(\bar{x}),$$

$$\partial\big(\varphi_1 \cdot \varphi_2\big)(\bar{x}) \subset \partial\big(\varphi_2(\bar{x})\varphi_1\big)(\bar{x}) + \partial\big(\varphi_1(\bar{x})\varphi_2\big)(\bar{x}),$$

where the latter holds as equality and the product $\varphi_1 \cdot \varphi_2$ is lower regular at $\bar{x}$ if both functions $\varphi_2(\bar{x})\varphi_1$ and $\varphi_1(\bar{x})\varphi_2$ are lower regular at this point.

Proof. To verify the first product rule, represent $\varphi_1 \cdot \varphi_2$ as composition (4.9) with $\varphi\colon \mathbb{R}^2 \to \mathbb{R}$ and $g\colon \mathbb{R}^n \to \mathbb{R}^2$ defined by

$$\varphi(y_1, y_2) := y_1 \cdot y_2 \ \text{ and } \ g(x) := \big(\varphi_1(x), \varphi_2(x)\big).$$

Then Theorem 4.5(ii) gives the claimed equality. Employing therein the subdifferential sum rule of Corollary 2.20 gives us the second product rule as inclusion, where the equality and regularity statements follow from the corresponding results of Theorem 2.19. △

Corollary 4.8 (Subdifferential Quotient Rules). *Let $\varphi_i\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ for $i = 1, 2$ be Lipschitz continuous around $\bar{x}$ with $\varphi_2(\bar{x}) \neq 0$. Then we have*

$$\partial\Big(\frac{\varphi_1}{\varphi_2}\Big)(\bar{x}) = \frac{\partial\big(\varphi_2(\bar{x})\varphi_1 - \varphi_1(\bar{x})\varphi_2\big)(\bar{x})}{[\varphi_2(\bar{x})]^2},$$

$$\partial\Big(\frac{\varphi_1}{\varphi_2}\Big)(\bar{x}) \subset \frac{\partial\big(\varphi_2(\bar{x})\varphi_1\big)(\bar{x}) - \partial\big(\varphi_1(\bar{x})\varphi_2\big)(\bar{x})}{[\varphi_2(\bar{x})]^2},$$

where the latter holds as equality and the quotient φ_1/φ_2 is lower regular at $\bar{x}$ if both functions $\varphi_2(\bar{x})\varphi_1$ and $-\varphi_1(\bar{x})\varphi_2$ are lower regular at this point.

Proof. Similar to Corollary 4.7 with $\varphi(y_1, y_2) := y_1/y_2$ therein. △

4.3 Subdifferentiation of Minima and Maxima

Next we proceed with evaluating basic and singular subdifferentials of *minima* and *maxima* of finitely many functions defined, respectively, by

$$\big(\min \varphi_i\big)(x) := \min\big\{\varphi_i(x)\big|\ i = 1, \ldots, s\big\}, \tag{4.13}$$

$$\big(\max \varphi_i\big)(x) := \max\big\{\varphi_i(x)\big|\ i = 1, \ldots, s\big\}, \tag{4.14}$$

where $\varphi_i\colon X \to \overline{\mathbb{R}}$ with $s \geq 2$. Functions of these two classes are *intrinsically nonsmooth* (even when all φ_i are linear), while their generalized differentiability properties are *very different* and cannot be reduced to each other by taking the negative sign; compare, e.g., the simplest functions $|x| = \max\{x, -x\}$ and $-|x| = \min\{x, -x\}$, and see Fig. 4.1. This issue has been well realized in convex analysis, while the difference can't be recognized by Clarke's generalized gradient (1.78) with its plus-minus symmetry, which implies the equality

$$\overline{\partial}\big(\min\varphi_i\big)(\bar{x}) = \overline{\partial}\big(\max\varphi_i\big)(\bar{x})$$

for arbitrary locally Lipschitzian functions φ_i. The usage of our nonconvex unilateral constructions fully recognizes this difference via the following calculus rules for evaluating subgradients of the minimum and maximum functions.

Let us start with the *minimum function* and define the set of *active indices*

$$I_{\min}(x) := \big\{i \in \{1,\ldots,s\}\big|\ \varphi_i(x) = \big(\min\varphi_i\big)(x)\big\}, \quad x \in \mathbb{R}^n.$$

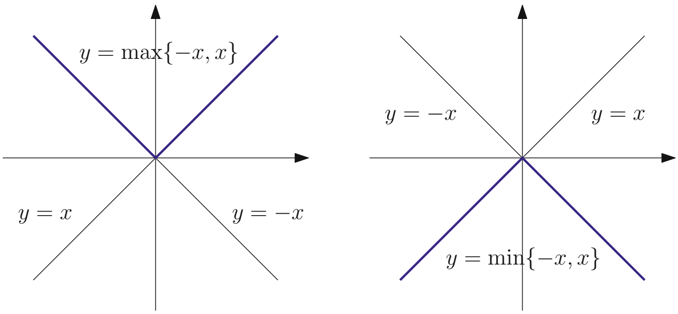


Fig. 4.1 Maximum and minimum functions.

Proposition 4.9 (Basic and Singular Subdifferentials of Minimum Functions). *Considering* (4.13)*, fix* $\bar{x} \in \cap_{i=1}^{s} \operatorname{dom}\varphi_i$. *Then we have*

$$\partial\big(\min\varphi_i\big)(\bar{x}) \subset \bigcup\Big\{\partial\varphi_i(\bar{x})\Big|\ i \in I_{\min}(\bar{x})\Big\}, \tag{4.15}$$

$$\partial^{\infty}\big(\min\varphi_i\big)(\bar{x}) \subset \bigcup\Big\{\partial^{\infty}\varphi_i(\bar{x})\Big|\ i \in I_{\min}(\bar{x})\Big\}. \tag{4.16}$$

Proof. We verify only (4.15), since representation (1.38) of the singular subdifferential allows us to proceed similarly in the proof of (4.16). Take a sequence of $x_k \in \mathbb{R}^n$ such that $x_k \to \bar{x}$ and $\varphi_i(x_k) \to (\min\varphi_i)(\bar{x})$ for $i \notin I_{\min}(\bar{x})$. Using the lower semicontinuity of φ_i at $\bar{x}$ (our standing assumption), we get $I_{\min}(x_k) \subset I_{\min}(\bar{x})$. It easily follows from definition (1.33) that

$$\widehat{\partial}\big(\min \varphi_i\big)(x_k) \subset \bigcup \big\{\widehat{\partial}\varphi_i(x_k)\big|\ i \in I_{\min}(\bar{x})\big\}, \quad k \in \mathbb{N}. \tag{4.17}$$

This yields by passing to the limit in (4.17) due to representation (1.37) of basic subgradients that (4.15) holds. The proof of (4.16) is similar by using the singular subdifferential representation (1.38). △

Although our standing assumption is the *lower* semicontinuity of the functions in question (unless otherwise stated), in the following theorem on subdifferentiation of the *maximum function* (4.14), we impose the *upper* semicontinuity (*u.s.c.*) of some functions under consideration. Denote

$$I_{\max}(\bar{x}) := \big\{i \in \{1,\ldots,s\}\big|\ \varphi_i(\bar{x}) = \big(\max \varphi_i\big)(\bar{x})\big\},$$

$$\Lambda(\bar{x}) := \Big\{(\lambda_1,\ldots,\lambda_s)\Big|\ \lambda_i \ge 0,\ \sum_{i=1}^{s}\lambda_i = 1,\ \lambda_i\Big(\varphi_i(\bar{x}) - \big(\max \varphi_i\big)(\bar{x})\Big) = 0\Big\}.$$

Theorem 4.10 (Subdifferentiation of Maximum Functions). *Let φ_i be l.s.c. around $\bar{x}$ for $i \in I_{\max}(\bar{x})$ and be u.s.c. at $\bar{x}$ for $i \notin I_{\max}(\bar{x})$. Then:*

(i) *Under the validity of the qualification condition (2.37) considered only for $i \in I_{\max}(\bar{x})$, we have the inclusions*

$$\partial\big(\max \varphi_i\big)(\bar{x}) \subset \bigcup \Big\{\sum_{i\in I_{\max}(\bar{x})} \lambda_i \circ \partial\varphi_i(\bar{x})\Big|\ (\lambda_1,\ldots,\lambda_s) \in \Lambda(\bar{x})\Big\},$$

$$\partial^{\infty}\big(\max \varphi_i\big)(\bar{x}) \subset \sum_{i\in I_{\max}(\bar{x})} \partial^{\infty}\varphi_i(\bar{x}),$$

where $\lambda \circ \partial\varphi(\bar{x})$ is defined as $\lambda\partial\varphi(\bar{x})$ for $\lambda > 0$ and as $\partial^{\infty}\varphi(\bar{x})$ for $\lambda = 0$. If furthermore each φ_i for $i \in I_{\max}(\bar{x})$ is epigraphically regular at $\bar{x}$, then the maximum function is also epigraphically regular at this point, and both inclusions above hold as equalities.

(ii) *Suppose that each φ_i, $i = 1,\ldots,s$, is Lipschitz continuous around $\bar{x}$. Then we have the inclusion*

$$\partial\big(\max \varphi_i\big)(\bar{x}) \subset \bigcup \Big\{\partial\Big(\sum_{i\in I_{\max}(\bar{x})} \lambda_i\varphi_i\Big)(\bar{x})\Big|\ (\lambda_1,\ldots,\lambda_s) \in \Lambda(\bar{x})\Big\},$$

where the equality holds and the maximum functions are lower regular at $\bar{x}$ if each φ_i, $i \in I_{\max}(\bar{x})$, is lower regular at this point.

Proof. Denoting $\bar{\alpha} := \big(\max \varphi_i\big)(\bar{x})$, observe that $(\bar{x},\bar{\alpha})$ is an interior point of the set $\operatorname{epi}\varphi_i$ for any $i \notin I_{\max}(\bar{x})$ due to the upper semicontinuity assumption. Then assertion (i) follows from the intersection rule of Corollary 2.17 applied to the epigraphs $\operatorname{epi}\varphi_i$, $i = 1, \ldots, s$, at $(\bar{x}, \bar{\alpha})$.

To verify assertion (ii), which provides a better upper estimate of the basic subdifferential for the case of Lipschitzian functions, we represent the maximum function as the composition $\varphi \circ g$ with

$$\varphi(y_1, \ldots, y_s) := \max\big\{y_1, \ldots, y_s\big\} \text{ and } g(x) := \big(\varphi_1(x), \ldots, \varphi_s(x)\big).$$

Then we apply to this composition the chain rule from Corollary 4.6 (with the equality and normal regularity statement therein) by taking into account the well-known formula for subdifferentiation of the convex function φ in the composition, which follows in turn from the equality in (i). △

4.4 Mean Value Theorems and Some Applications

It has been well recognized in mathematics that the classical *Lagrange mean value theorem* is one of the central results of real analysis that plays a crucial role in a variety of applications. This section contains several extended versions of the mean value theorem in the absence of differentiability. We also present some of their striking applications to important topics of variational analysis.

4.4.1 Mean Value Theorem via Symmetric Subgradients

Let us begin with deriving a generalized mean value theorem for *continuous* functions, which we obtain in the Lagrangian form with replacement of the classical gradient by a proper (actually *minimal* for such a form) subdifferential construction. This construction is the *symmetric subdifferential*

$$\partial^0\varphi(\bar{x}) := \partial\varphi(\bar{x}) \cup \big[-\partial(-\varphi)(\bar{x})\big] \tag{4.18}$$

some properties of which are discussed in Exercise 1.75. Its singular counterpart $\partial^{\infty,0}\varphi(\bar{x})$ from (1.75) is used to formulate the appropriate qualification condition needed for the validity of the following extended mean value theorem. For given $a, b \in \mathbb{R}^n$ we use the notation

$$(b-a)^{\perp} := \big\{x^* \in \mathbb{R}^n \big|\, \langle x^*, b-a\rangle = 0\big\}, \quad [a,b] := \big\{a + t(b-a)\big|\, 0 \le t \le 1\big\}$$

with (a,b), $(a,b]$, and $[a,b)$ defined accordingly.

Theorem 4.11 (Symmetric Subdifferential Mean Value Theorem for Continuous Functions). *Let* $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ *be continuous on an open set containing* $[a,b]$, *and let the qualification condition*

$$\partial^{\infty,0}\varphi(x) \cap (b-a)^{\perp} = \{0\} \text{ for every } x \in (a,b)$$

be satisfied. Then we have the mean value inclusion

$$\varphi(b)-\varphi(a)\in\big\langle\partial^0\varphi(c),b-a\big\rangle \text{ for some } c\in(a,b). \tag{4.19}$$

Proof. Let us first justify the existence of a real number $\theta\in(0,1)$ such that

$$\varphi(b)-\varphi(a)\in\partial_t^0\varphi\big(a+\theta(b-a)\big), \tag{4.20}$$

where on the right hand we have the symmetric subdifferential (1.75) of the function $t\to\varphi(a+t(b-a))$ at $t=\theta$. To proceed, define $\phi\colon[0,1]\to\mathbb{R}$ by

$$\phi(t):=\varphi\big(a+t(b-a)\big)+t\big(\varphi(a)-\varphi(b)\big),\quad 0\le t\le 1,$$

and observe that ϕ is continuous on $[0,1]$ with $\phi(0)=\phi(1)=\varphi(a)$. The classical Weierstrass theorem tells us that ϕ attains its minimum and maximum on $[0,1]$. Excluding the trivial case where ψ is constant on $[0,1]$, we get an *interior point* $\theta\in(0,1)$ at which ϕ attains either the minimal or maximal value over $[0,1]$. Then it follows from the generalized Fermat rule of Proposition 1.30(i) and its upper counterpart in the case of maxima that $0\in\partial^0\phi(\theta)$. Observing that ϕ is the sum of two functions one of which is smooth, we apply the elementary sum rule from Proposition 1.30(ii) and arrive at (4.20).

Represent now the function in (4.20) as the composition

$$\varphi\big(a+t(b-a)\big)=\big(\varphi\circ g\big)(t)\ \text{ with }\ g(t):=a+t(b-a),\quad 0\le t\le 1.$$

Applying finally the subdifferential chain rule of Corollary 4.6 and its upper counterpart to this composition gives us the mean value inclusion (4.19) with $c:=a+\theta(b-a)$ under the imposed qualification condition. △

Corollary 4.12 (Symmetric Subdifferential Mean Value Theorem for Lipschitzian Functions). *If φ be Lipschitz continuous on an open set containing $[a,b]$, then (4.19) holds. If in addition φ is lower regular on the interval (a,b), then we have the inclusion*

$$\varphi(b)-\varphi(a)\in\big\langle\partial\varphi(c),b-a\big\rangle \text{ for some } c\in(a,b). \tag{4.21}$$

Proof. It follows from Theorem 1.22 that the qualification condition of Theorem 4.11 is automatic for Lipschitzian functions. It remains to verify (4.21) under the assumed lower regularity. To this end we get from Theorem 4.5(i) that the lower regularity of φ at $c=a+\theta(b-a)$ yields the lower regularity of the function $t\to\varphi(a+t(b-a))=(\varphi\circ g)(t)$ at θ. Thus we get $\widehat\partial(\varphi\circ g)(\theta)=\partial(\varphi\circ g)(\theta)\ne\emptyset$ by Theorem 1.22 due to the Lipschitz continuity of $\varphi\circ g$. It easily implies that $\widehat\partial^+(\varphi\circ g)(\theta)\subset\widehat\partial(\varphi\circ g)(\theta)$; see Exercise 1.76(i). In this case it follows from the proof of (4.20) in Theorem 4.11 that

$$\varphi(b)-\varphi(a)\in\widehat\partial\big(\varphi\circ g\big)(\theta)\subset\partial\big(\varphi\circ g\big)(\theta),$$

and thus we arrive at (4.21) by using Corollary 4.6. △

Note that the lower regularity assumption is *essential* for the validity of the extended mean value theorem in form (4.21). A simple *counterexample* is provided by $\varphi(x) := -|x|$ on $[a,b] = [-1,1]$ with $\partial\varphi(0) = \{-1,1\}$ and $\partial^0\varphi(0) = [-1,1]$. This shows that (4.19) holds while (4.21) doesn't.

4.4.2 Approximate Mean Value Theorems

Next we present a mean value theorem of a *new type*, which has never appeared in the classical or convex analysis. Results of this type apply to *lower semicontinuous* extended-real-valued functions and are known as *approximate mean value theorems* (AMVT); see more commentaries in Section 4.6. Such results occur to be very instrumental in variational analysis, which is partly demonstrated in the rest of this section. The formulation of the following version of AMVT involves the regular subdifferential (1.33).

Theorem 4.13 (Approximate Mean Value Theorem for l.s.c. Functions). *Let $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ be finite at two given points $a \neq b$, and let $c \in [a,b)$ belong to the nonempty set of minimizers for the function*

$$\psi(x) := \varphi(x) - \frac{\varphi(b)-\varphi(a)}{\|b-a\|}\|x-a\|, \quad x \in [a,b].$$

Then there are sequences $x_k \xrightarrow{\varphi} c$ and $x_k^* \in \widehat{\partial}\varphi(x_k)$ satisfying

$$\liminf_{k\to\infty}\langle x_k^*, b-x_k\rangle \geq \frac{\varphi(b)-\varphi(a)}{\|b-a\|}\|b-c\|, \tag{4.22}$$

$$\liminf_{k\to\infty}\langle x_k^*, b-a\rangle \geq \varphi(b)-\varphi(a). \tag{4.23}$$

If furthermore $c \neq a$, then we have the equality

$$\lim_{k\to\infty}\langle x_k^*, b-a\rangle = \varphi(b)-\varphi(a).$$

Proof. Observe first that the function ψ defined in the theorem is l.s.c., and hence it attains its minimum over $[a,b]$ at some point c. Since $\psi(a) = \psi(b)$, we can always take $c \in [a,b)$. Suppose without loss of generality that $\varphi(a) = \varphi(b)$, which gives us $\psi(x) = \varphi(x)$ for all $x \in [a,b]$. The lower semicontinuity of φ ensures the existence of $r > 0$ such that φ is bounded from below on the set $\Theta := [a,b] + r\mathbb{B}$ by some $\gamma \in \mathbb{R}$. Thus the function $\vartheta(x) := \varphi(x) + \delta(x;\Theta)$ is l.s.c. and bounded from below on the whole space $\mathbb{R}^n$. For any fixed $k \in \mathbb{N}$ take $r_k \in (0,r)$ such that $\varphi(x) \geq \varphi(c) - k^{-2}$ whenever $x \in [a,b] + r_k\mathbb{B}$ and choose $t_k \geq k$ satisfying $\gamma + t_k r_k \geq \varphi(c) - k^{-2}$. Thus we have

$$\varphi(c) = \vartheta_k(c) \leq \inf_{x\in\mathbb{R}^n}\vartheta_k(x) + k^{-2} \ \text{ with } \ \vartheta_k(x) := \vartheta(x) + t_k\text{dist}\big(x;[a,b]\big).$$

Apply to $\vartheta_k(x)$ the Ekeland's variational principle from Corollary 2.13 with the parameters $\varepsilon = k^{-2}$ and $\lambda = k^{-1}$. In this way we find $x_k \in \mathbb{R}^n$ with

$$\|x_k - c\| \le k^{-1}, \quad \vartheta_k(x_k) \le \vartheta_k(c) = \varphi(c), \quad \vartheta_k(x_k) \le \vartheta_k(x) + k^{-1}\|x - x_k\|$$

for all $x \in \mathbb{R}^n$. The latter means that the function $\vartheta_k(x) + k^{-1}\|x - x_k\|$ attains its minimum at $x = x_k$. Hence we deduce from the subdifferential Fermat and sum rules of Proposition 1.30(i) and Corollary 2.20, respectively, that

$$0 \in \partial\vartheta_k(x_k) + k^{-1}\mathbb{B} \ \text{ for all } \ k \in \mathbb{N}$$

via the (dual) unit ball $\mathbb{B} \subset \mathbb{R}^n$ due to $\partial\|\cdot\|(0) = \mathbb{B}$. Now using the first representation in (1.37) of Theorem 1.28 for the basic subdifferential, applying again the sum rule from Corollary 2.20 to the sum in ϑ_k with taking into account that $x_k \in \operatorname{int}\Theta$ for large k, we find sequences $u_k \xrightarrow{\varphi} c$, $v_k \to c$, $u_k^* \in \widehat{\partial}\varphi(u_k)$, $v_k^* \in \partial\operatorname{dist}(v_k; [a,b])$, and $e_k^* \in \mathbb{B}$ such that

$$\|u_k^* + t_k v_k^* + k^{-1}e_k^*\| \to 0 \ \text{ as } \ k \to \infty, \tag{4.24}$$

where $\|v_k^*\| \le 1$ by Proposition 1.33 and, obviously,

$$\langle v_k^*, b - v_k\rangle \le \operatorname{dist}\big(b; [a,b]\big) - \operatorname{dist}\big(v_k; [a,b]\big) \le 0, \quad k \in \mathbb{N}.$$

Our next goal is to construct a point $w_k \in [a,b]$ for each $k \in \mathbb{N}$ so that it enjoys properties similar to v_k. We do it by picking an arbitrary projection $w_k \in \Pi(v_k; [a,b])$ and observing that

$$\begin{aligned}\langle v_k^*, b - w_k\rangle &= \langle v_k^*, b - v_k\rangle + \langle v_k^*, v_k - w_k\rangle \le \operatorname{dist}\big(b; [a,b]\big) - \operatorname{dist}\big(v_k; [a,b]\big)\\ &\quad + \|v_k^*\|\cdot\|v_k - w_k\| \le -\operatorname{dist}\big(v_k; [a,b]\big) + \operatorname{dist}\big(v_k; [a,b]\big) = 0.\end{aligned}$$

This yields $\langle v_k^*, b - a\rangle \le 0$ for large $k \in N$ since $w_k \to c \ne b$ and $(x-b)\|y-b\| = (y-b)\|x-b\|$ for $x, y \in [a,b]$. It follows now from (4.24) that

$$\liminf_{k\to\infty}\langle u_k^*, b - u_k\rangle \ge 0 \ \text{ and } \ \liminf_{k\to\infty}\langle u_k^*, b - a\rangle \ge 0,$$

which verify (4.22) and (4.23). If finally $c \ne a$, then $v_k \ne a$ for large $k \in \mathbb{N}$, and hence $\langle v_k^*, b - c\rangle = 0$. This readily implies that $\langle u_k^*, b - a\rangle \to 0$ by the above arguments and thus completes the proof of the theorem. △

Next we show that the crucial *mean value inequality* (4.23) holds even if $\varphi(b) = \infty$ and implies a useful estimate of the increment for a given l.s.c. function via its regular subdifferential. Furthermore, we establish the limiting counterparts of these relationships for Lipschitzian functions.

Corollary 4.14 (Mean Value Inequalities). *The following assertions hold:*

(i) *Let* $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ *be finite at* $a \in \mathbb{R}^n$. *Then for any* $b \in \mathbb{R}^n$ *there exist a point* $c \in [a, b]$ *and sequences* $x_k \overset{\varphi}{\to} c$, $x_k^* \in \widehat{\partial}\varphi(x_k)$ *satisfying the mean value inequality* (4.23). *Furthermore, for each* $\varepsilon > 0$ *we have the estimate*

$$|\varphi(b) - \varphi(a)| \le \|b - a\| \sup\big\{\|x^*\| \;\big|\; x^* \in \widehat{\partial}\varphi(c),\ c \in [a, b] + \varepsilon\mathbb{B}\big\}. \quad (4.25)$$

(ii) *If* φ *is Lipschitz continuous on an open set containing* $[a, b]$, *then*

$$\langle x^*, b - a\rangle \ge \varphi(b) - \varphi(a) \;\; \textit{for some} \;\; x^* \in \partial\varphi(c) \;\; \textit{with} \;\; c \in [a, b),$$

$$|\varphi(b) - \varphi(a)| \le \|b - a\| \sup\big\{\|x^*\| \;\big|\; x^* \in \partial\varphi(c),\ c \in [a, b)\big\}.$$

Proof. To verify the mean value inequality in (i), it remains to consider the case where $\varphi(b) = \infty$. In this case it suffices to apply (4.23) for each $s \in \mathbb{N}$ to the sequence of the modified functions

$$\phi_s(x) := \begin{cases} \varphi(x) & \text{if } x \ne b, \\ \varphi(a) + s & \text{if } x = b. \end{cases}$$

The increment estimate in (i) immediately follows from (4.23).

To justify (ii), employ Theorem 4.13 to find $c \in [a, b)$, $x_k \to c$, and $x_k^* \in \widehat{\partial}\varphi(x_k)$ satisfying (4.23), and observe by definition (1.33) of regular subgradients that the Lipschitz continuity of φ ensures the uniform boundedness of the sequence $\{x_k^*\}$. Thus it contains a convergence subsequence which limit x^* belongs to the basic subdifferential $\partial\varphi(c)$ due to (1.37). Then the mean value inequality in (ii) follows by passing to the limit in (4.23). It readily implies the increment estimate in (ii). $\triangle$

Note that the mean value inequality in Corollary 4.14(ii) provides a *unilateral version* (inequality vs. equality) of the extended mean value theorem for Lipschitzian functions from Corollary 4.12 by using only the basic subdifferential instead of its symmetric counterpart (4.18) without lower regularity.

4.4.3 Subdifferential Characterizations from AMVT

Finally in this section, we present several remarkable consequences of the approximate mean value theorem. The first one concerns subdifferential characterizations of *local Lipschitz continuity* of lower semicontinuous functions.

Theorem 4.15 (Subdifferential Characterizations of Local Lipschitz Continuity). *Given* $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ *with* $\bar{x} \in \operatorname{dom}\varphi$, *and given a constant* $\ell \ge 0$, *the following properties are equivalent:*

(a) *There exists a positive number* γ *such that*

$$\widehat{\partial}\varphi(x) \subset \ell\mathbb{B} \;\; \textit{whenever} \;\; \|x - \bar{x}\| < \gamma, \quad |\varphi(x) - \varphi(\bar{x})| < \gamma.$$

(b) *There is a neighborhood U of $\bar{x}$ on which $\widehat{\partial}\varphi(x) \subset \ell\mathbb{B}$.*
(c) *φ is Lipschitz continuous around $\bar{x}$ with modulus ℓ.*

Furthermore, the local Lipschitz continuity of φ around $\bar{x}$ with SOME MODULUS *$\ell \geq 0$ is equivalent to the singular subdifferential condition $\partial^\infty\varphi(\bar{x}) = \{0\}$.*

Proof. Suppose without loss of generality that $\bar{x} = 0$ and $\varphi(\bar{x}) = 0$, and verify first the validity of implication (a)$\Rightarrow$(b) with $U := \eta(\operatorname{int}\mathbb{B})$ in (b) for some constant $\eta > 0$. This means that in the setting of (a), there is $\eta > 0$ such that $|\varphi(x)| < \gamma$ for all $\|x\| < \eta$. Observe that the lower semicontinuity of φ around $\bar{x} = 0$ allows us to find $\nu > 0$ so that $\varphi(x) > -\gamma$ if $\|x\| < \nu$. Denote $\eta := \min\{\nu, \gamma, \gamma/\ell\}$, where the case of $\ell = 0$ is included and thus reduces η to $\min\{\nu, \gamma\}$, and then show that $\varphi(x) < \gamma$ whenever $\|x\| < \min\{\gamma, \gamma/\ell\}$. This would surely justify the claimed implication.

Suppose on the contrary that there exists $b \in \mathbb{R}^n$ with $\|b\| < \min\{\gamma, \gamma/\ell\}$ and $\varphi(b) \geq \gamma$. Consider the function

$$\phi(x) := \min\{\varphi(x), \gamma\} \quad \text{on } \mathbb{R}^n \text{ with } \ \phi(0) = 0, \ \phi(b) = \gamma$$

satisfying all the assumptions of Theorem 4.13, and apply to it the mean value inequality (4.23). This gives us $c \in [0, b)$ and $x_k \xrightarrow{\phi} c$, $x_k^* \in \widehat{\partial}\phi(x_k)$ with

$$\liminf_{k\to\infty}\langle x_k^*, b\rangle \geq \phi(b) - \phi(0) = \gamma, \qquad \liminf_{k\to\infty}\|x_k^*\| \geq \gamma/\|b\| > \ell.$$

Recall that the point c is a minimizer of the function

$$\psi(x) := \phi(x) - \|b\|^{-1}\|x\|\big(\phi(b) - \phi(0)\big) \ \text{ on } \ [0, b],$$

which yields $\phi(c) \leq \gamma\|b\|^{-1}\|c\| < \gamma$. Hence $\phi(x_k) < \gamma$ along $x_k \xrightarrow{\phi} c$ telling us that $\phi(x_k) = \varphi(x_k)$ for large k. We get furthermore that

$$\widehat{\partial}\phi(x_k) \subset \widehat{\partial}\varphi(x_k) \ \text{ by } \ \phi(x) \leq \varphi(x), \ x \in \mathbb{R}^n,$$

and so $x_k^* \in \widehat{\partial}\varphi(x_k)$. Since $\|x_k^*\| > \ell$, it contradicts (a) and verifies (a)$\Rightarrow$(b).

Implication (b)$\Rightarrow$(c) follows from the increment estimate in Corollary 4.14(i), implication (c)$\Rightarrow$(b) is an easy consequence of the definition, while (b)$\Rightarrow$(a) is trivial. It has been proved in Theorem 1.22 that $\partial^\infty\varphi(\bar{x}) = \{0\}$ for locally Lipschitzian functions; so it remains to verify the opposite implication. Due to the equivalence (a)$\Leftrightarrow$(c), it suffices to show that (a) holds with some $\ell, \gamma > 0$.

If it doesn't, find $x_k \xrightarrow{\varphi} \bar{x}$ and $x_k^* \in \widehat{\partial}\varphi(x_k)$ with $\|x_k^*\| \to \infty$. This yields

$$\Big(\frac{x_k^*}{\|x_k^*\|}, -\frac{1}{\|x_k^*\|}\Big) \in \widehat{N}\big((x_k, \varphi(x_k)); \operatorname{epi}\varphi\big), \qquad k \in \mathbb{N}.$$

Normalizing $\widetilde{x}_k^* := x_k^*/\|x_k^*\|$, select a subsequence of $\{\widetilde{x}_k^*\}$ that converges to some x^* with $\|x^*\| = 1$ and $(x^*, 0) \in N((\bar{x}, \varphi(\bar{x})); \operatorname{epi}\varphi)$. This contradicts the imposed condition $\partial^\infty\varphi(\bar{x}) = \{0\}$ and thus completes the proof. △

Theorem 4.15 readily implies a subdifferential extension of the fundamental result of classical analysis that bridges *differentiation* and *integration*, namely, the only function having zero derivative on an open set is constant.

Corollary 4.16 (Subdifferential Characterization of Constancy for l.s.c. Functions). *Consider $\varphi\colon U \to \overline{\mathbb{R}}$ on an open set $U \subset \mathbb{R}^n$. Then φ is locally constant on U if and only if we have*

$$x^* \in \widehat{\partial}\varphi(x) \Longrightarrow x^* = 0 \;\text{ for all }\; x \in U.$$

This is equivalent to φ being constant on U if U is connected.

Proof. Immediately follows from Theorem 4.15 with $\ell = 0$ therein. △

The next remarkable consequence of AMVT is the following subdifferential characterizations of *strictly differentiable* functions (1.19). The functions from Example 1.21 illustrate that imposing *Lipschitz* continuity as well as *strict* differentiability are essential for the validity of the obtained equivalences.

Theorem 4.17 (Subdifferential Characterizations of Strict Differentiability). *Given a (l.s.c.) function $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ finite at $\bar{x}$ and given a vector $\bar{x}^* \in \mathbb{R}^n$, the following properties are equivalent:*

(a) *φ is Lipschitz continuous around $\bar{x}$, and for every sequences $x_k \to \bar{x}$ and $x_k^* \in \widehat{\partial}\varphi(x_k)$ we have $x_k^* \to \bar{x}^*$ as $k \to \infty$.*

(b) *φ is Lipschitz continuous around $\bar{x}$ with $\partial\varphi(\bar{x}) = \{\bar{x}^*\}$.*

(c) *φ is strictly differentiable at $\bar{x}$ with $\nabla\varphi(\bar{x}) = \bar{x}^*$.*

Proof. Suppose without loss of generality that $\bar{x} = 0$, $\varphi(0) = 0$, and $\bar{x}^* = 0$. To verify (a)⇒(b), pick $x^* \in \partial\varphi(0)$ and find $x_k \to 0$ and $x_k^* \in \widehat{\partial}\varphi(x_k)$ with $x_k^* \to x^*$. It follows from (a) that $x^* = 0$, i.e., $\partial\varphi(0) = \{0\}$ and (b) holds.

To show next that (b)⇒(c), observe that the strict differentiability of φ at $\bar{x} \in \mathbb{R}^n$ with $x^* = \nabla\varphi(\bar{x})$ can be equivalently described as

$$\lim_{\substack{x \to \bar{x} \\ t \downarrow 0}} \left[\sup_{u \in C} \left| \frac{\varphi(x + tu) - \varphi(x)}{t} - \langle x^*, u \rangle \right| \right] = 0 \tag{4.26}$$

for any bounded and closed set $C \subset \mathbb{R}^n$. Arguing by contradiction, suppose that there is such C for which the limit in (4.26) either doesn't exist or is not zero. In both cases select subsequences $x_k \to 0$, $t_k \downarrow 0$, and $u_k \in C$ so that

$$\lim_{k \to \infty} \frac{\varphi(x_k + t_k u_k) - \varphi(x_k)}{t_k} := \alpha > 0.$$

Then the mean value inequality (4.23) gives us $c_k \in \mathbb{R}^n$ and $x_k^* \in \widehat{\partial}\varphi(c_k)$ with

$$\text{dist}\big(c_k; [x_k, x_k + t_k u_k]\big) \le k^{-1}, \quad \langle x_k^*, t_k u_k\rangle \ge \varphi(x_k + t_k u_k) - \varphi(x_k) - t_k k^{-1},$$

and thus $c_k \to 0$. By the compactness of C, find a subsequence of $\{u_k\}$ converging to some $u \in C$. Also, the boundedness of the sequence $\{x_k^*\}$ due to the local Lipschitz continuity of φ allows us to select its subsequence which converges to some $x^* \in \partial\varphi(0)$. Passing now to the limit above shows that

$$\|x^*\| \cdot \|u\| \ge \langle x^*, u\rangle \ge \lim_{k\to\infty} \frac{\varphi(x_k + t_k u_k) - \varphi(x_k)}{t_k} = \alpha > 0,$$

which tells us that $x^* \neq 0$ and hence contradicts (b).

To verify finally (c)⇒(a), recall first that the local Lipschitz continuity of φ around $\bar{x} = 0$ always follows from the strict differentiability of φ at this point; see Exercise 1.52. It remains to justify the limiting relationship in (a) with $\bar{x}^* = 0$. For any sequences $x_k \to 0$ and $x_k^* \in \widehat{\partial}\varphi(x_k)$, we have

$$\liminf_{x\to x_k} \frac{\varphi(x) - \varphi(x_k) - \langle x_k^*, x - x_k\rangle}{\|x - x_k\|} \ge 0$$

and hence for every $\gamma_k \downarrow 0$, find neighborhoods U_k of x_k with

$$\langle x_k^*, x - x_k\rangle \le \varphi(x) - \varphi(x_k) + \gamma_k\|x - x_k\| \text{ on } U_k.$$

Fix $u \in \mathbb{B}$ and take $t > 0$ so small that $x := x + k + tu \in U_k$. Then

$$\langle x_k^*, u\rangle \le \frac{\varphi(x_k + tu) - \varphi(x_k)}{t} + \gamma_k\|u\|.$$

Replacing u by $-u$ in the above inequality, we arrive at the estimates

$$\big|\langle x_k^*, u\rangle\big| \le \left|\frac{\varphi(x_k + tu) - \varphi(x_k)}{t}\right| + \gamma_k \text{ and}$$

$$\sup_{u\in\mathbb{B}} \Big\{\big|\langle x_k^*, u\rangle\big|\Big\} \le \sup_{u\in\mathbb{B}} \left[\left|\frac{\varphi(x_k + tu) - \varphi(x_k)}{t}\right|\right] + \gamma_k,$$

which imply in turn the limiting relationship

$$\lim_{k\to\infty} \|x_k^*\| \le \lim_{k\to\infty,\, t\downarrow 0} \left[\sup_{u\in\mathbb{B}} \left|\frac{\varphi(x_k + tu) - \varphi(x_k)}{t}\right|\right] + \lim_{k\to\infty} \gamma_k.$$

Thus we get $x_k^* \to 0$ as $k \to \infty$ and complete the proof of theorem. △

Starting from the next theorem and then continuing it in this section and also in Chapter 5, we proceed with the study of various kinds of *monotonicity* of functions and operators, which plays a fundamental role in many aspects of variational analysis and optimization.

The result below concerns monotonicity of extended-real-valued *functions*. It provides, in particular, a subdifferential extension of the classical fact, based on the Lagrange mean value theorem, that a smooth function whose derivative is nonpositive must itself be nonincreasing.

Theorem 4.18 (Subdifferential Characterization of Monotonicity for l.s.c. Functions). *Let $\varphi\colon U \to \overline{\mathbb{R}}$ be defined on an open subset $U \subset \mathbb{R}^n$, and let $K \subset \mathbb{R}^n$ be a cone with its polar $K^* = \{x^* \in \mathbb{R}^n \mid \langle x^*, x\rangle \le 0,\ x \in K\}$. Then the following properties are equivalent:*

(a) *The function φ is K-nonincreasing, i.e.,*

$$x, u \in U,\ u - x \in K \Longrightarrow \varphi(u) \le \varphi(x).$$

(b) *For every $x \in U$ we have $\widehat{\partial}\varphi(x) \subset K^*$.*

Proof. To verify (a)⇒(b), take any vectors $x \in U$ and $x^* \in \widehat{\partial}\varphi(x)$. Given $\gamma > 0$, find by the subgradient definition such $\eta > 0$ that $x + \eta\mathbb{B} \subset U$ and

$$\langle x^*, u - x\rangle \le \varphi(u) - \varphi(x) + \gamma\|u - x\| \text{ for all } u \in x + \eta\mathbb{B}.$$

Fix $w \in K$ and plug into this inequality $u := x + tw$ with $t > 0$ and $t\|w\| \le \eta$. Then the K-monotonicity property in (a) tells us that

$$\langle x^*, w\rangle \le \frac{\varphi(x + tw) - \varphi(x)}{t} + \gamma\|w\| \le \gamma\|w\|.$$

Since this holds for any $\gamma > 0$, we arrive at $\langle x^*, w\rangle \le 0$ and justify therefore the subdifferential inclusion in (b).

To verify the opposite implication (b)⇒(a), suppose the contrary, and find points $x, u \in U$ satisfying $u - x \in K$ and $\varphi(u) > \varphi(x)$. Applying the mean value inequality from Corollary 4.14(i), we get $c \in [x, u]$ and sequences $x_k \to c$, $x_k^* \in \widehat{\partial}\varphi(x_k)$ satisfying the conditions

$$\liminf_{k\to\infty}\langle x_k^*, u - x\rangle \ge \varphi(u) - \varphi(x) > 0, \quad k \in \mathbb{N}.$$

This yields $\langle x_k^*, u - x\rangle > 0$ for large k and contradicts (b). △

The last application of AMVT in the section is to the monotonicity of *subgradient mappings* generated by l.s.c. functions. Recall that a set-valued mapping $T\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is *globally monotone* on $\mathbb{R}^n$ if

$$\langle v_1 - v_2, u_1 - u_2\rangle \ge 0 \text{ for all } (u_1, v_1), (u_2, v_2) \in \operatorname{gph} T. \tag{4.27}$$

The mapping T is *globally maximal monotone* on $\mathbb{R}^n$ if $\operatorname{gph} T = \operatorname{gph} S$ for any monotone operator $S\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ with $\operatorname{gph} T \subset \operatorname{gph} S$.

It is well known in convex analysis that the subgradient mapping for a l.s.c. convex function $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ is globally maximal monotone. The next theorem shows that the global monotonicity (not even maximal) of either the regular subdifferential mapping or the basic subdifferential one for a l.s.c. function φ yields the convexity of φ.

Theorem 4.19 (Subdifferential Monotonicity and Convexity for l.s.c. Functions). *Given $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$, suppose that either $\widehat{\partial}\varphi$ or $\partial\varphi$ is a globally monotone operator on $\mathbb{R}^n$. Then the function φ must be convex on $\mathbb{R}^n$.*

Proof. As follows from the limiting representation of $\partial\varphi$ via $\widehat{\partial}\varphi$ in Theorem 1.28, it suffices to prove the claimed result just for $\widehat{\partial}\varphi$, since the global monotonicity of this mapping implies that for $\partial\varphi$.

Let us first show that the global monotonicity of the regular subdifferential mapping $\widehat{\partial}\varphi$ implies its representation

$$\widehat{\partial}\varphi(x) = \big\{v \in \mathbb{R}^n \big|\ \langle v, u - x\rangle \le \varphi(u) - \varphi(x)\ \text{ for all }\ u \in \mathbb{R}^n\big\} \tag{4.28}$$

in the *form* of the subdifferential of convex analysis whenever $x \in \operatorname{dom}\varphi$. Since the inclusion "$\supset$" in (4.28) is obvious, we proceed with the proof of the opposite inclusion by using AMVT. Pick $x, u \in \operatorname{dom}\varphi$ and $v \in \widehat{\partial}\varphi(x)$. Applying (4.22) gives us sequences $x_k \to c \in [u, x)$ and $v_k \in \widehat{\partial}\varphi(x_k)$ such that

$$\varphi(x) - \varphi(u) \le \frac{\|x-u\|}{\|x-c\|}\liminf_{k\to\infty}\langle v_k, x - x_k\rangle.$$

The global monotonicity of $\widehat{\partial}\varphi$ in (4.27) and the equality $\|x-u\|(x-c) = (x-u)\|x-c\|$ ensure the validity of the conditions

$$\varphi(x) - \varphi(u) \le \frac{\|x-u\|}{\|x-c\|}\liminf_{k\to\infty}\langle v, x - x_k\rangle = \langle v, x-u\rangle,$$

which justify the inclusion "$\subset$" in (4.28) and so the claimed representation.

Using (4.28) and employing AMVT again, we show next that φ is *convex*. For any $u, x \in \operatorname{dom}\varphi$ consider its convex combination $w := \lambda u + (1-\lambda)x$ with $0 < \lambda < 1$. It follows from the variational arguments of Theorem 2.14 (see also Exercise 2.25(i)) that the domain of $\widehat{\partial}\varphi$ is *dense in the graph* of φ. This gives us a sequence $u_k \overset{\varphi}{\to} u$ with $\widehat{\partial}\varphi(u_k) \ne \emptyset$. Fixing k, we can always suppose that $0 \in \widehat{\partial}\varphi(u_k)$. Let us verify that $w_k \in \operatorname{dom}\varphi$ for $w_k := \lambda u_k + (1-\lambda)x$. Assuming the contrary, take $\alpha > \varphi(x)$ and define the function

$$\psi(z) := \begin{cases} \varphi(z) & \text{if } z \ne w_k,\\ \alpha & \text{if } z = w_k.\end{cases}$$

Applying to it the mean value inequalities of Theorem 4.13 gives us $c \in [x, w_k)$ and sequences $z_m \to c$, $v_m \in \widehat{\partial}\psi(z_m)$ as $m \in \mathbb{N}$ such that

$$\liminf_{m\to\infty}\langle v_m, w_k - z_m\rangle \ge \frac{\|w_k - c\|}{\|w_k - x\|}\big(\alpha - \varphi(x)\big) > 0,$$

$$\liminf_{m\to\infty}\langle v_m, w_k - x\rangle \ge \alpha - \varphi(x).$$

We deduce from the monotonicity of $\widehat{\partial}\varphi$ and the choice of $0 \in \widehat{\partial}\varphi(u_k)$ that

$$\begin{aligned}0 &\geq \liminf_{m\to\infty}\langle v_m, u_k - z_m\rangle \geq \liminf_{m\to\infty}\langle v_m, w_k - z_m\rangle + \liminf_{m\to\infty}\langle v_m, u_k - w_k\rangle\\ &= \liminf_{m\to\infty}\langle v_m, w_k - z_m\rangle + \lambda^{-1}(1-\lambda)\liminf_{m\to\infty}\langle v_m, w_k - x\rangle\\ &\geq \lambda^{-1}(1-\lambda)\big(\alpha - \varphi(x)\big).\end{aligned}$$

It contradicts the assumption on $\alpha > \varphi(x)$ and hence shows that $w_k \in \operatorname{dom}\varphi$.

To continue verifying the convexity of φ, we split the subsequent proof into the consideration of two cases regarding the role of $w_k = \lambda u_k + (1-\lambda)x$ as local minimizers of φ. Suppose without loss of generality that the assumptions of either Case 1 or Case 2 are satisfied for all $k \in \mathbb{N}$.

Case 1: *Let w_k be a local minimizer of φ.* In this case we have $0 \in \widehat{\partial}\varphi(w_k)$. Then representation (4.28) gives us $\varphi(x) \geq \varphi(w_k)$ and $\varphi(u_k) \geq \varphi(w_k)$, which yields $\lambda\varphi(u_k) + (1-\lambda)\varphi(x) \geq \varphi(w_k)$. Letting $k \to \infty$, we arrive at

$$\lambda\varphi(u) + (1-\lambda)\varphi(x) \geq \varphi(w) = \varphi\big(\lambda u + (1-\lambda)x\big), \tag{4.29}$$

which justifies the convexity of φ in this case.

Case 2: *Let w_k be not a local minimizer of φ.* Select s_k so that $\|s_k - w_k\| < k^{-1}$ and $\varphi(s_k) < \varphi(w_k)$. For any fixed k, we apply again Theorem 4.13 to the function φ on the interval $[s_k, w_k]$. It gives us $c_k \in [s_k, w_k)$ and sequences $z_m \to c_k$ as $m \to \infty$ and $v_m \in \widehat{\partial}\varphi(z_m)$ satisfying the conditions

$$\liminf_{m\to\infty}\langle v_m, w_k - z_m\rangle \geq \frac{\|w_k - c_k\|}{\|w_k - s_k\|}\big(\varphi(w_k) - \varphi(s_k)\big) > 0,$$

which imply by representation (4.28) that

$$\varphi(x) - \varphi(z_m) \geq \langle v_m, x - z_m\rangle, \quad \varphi(u_k) - \varphi(z_m) \geq \langle v_m, u_k - z_m\rangle.$$

This readily yields by passing to the limit as $m \to \infty$ and using the imposed lower semicontinuity of φ that

$$\lambda\varphi(u_k) + (1-\lambda)\varphi(x) \geq \liminf_{m\to\infty}\big[\varphi(z_m) + \langle v_m, w_k - z_m\rangle\big] \geq \varphi(c_k), \quad k \in \mathbb{N}.$$

Letting finally $k \to \infty$ gives us (4.29), which verifies the convexity of φ in this case and thus completes the proof of the theorem. △

4.5 Exercises for Chapter 4

Exercise 4.20 (Subdifferentials of Marginal Functions in Infinite Dimensions). Consider the class of marginal functions of type (4.1), where the (locally l.s.c.) cost function $\varphi\colon X \to \overline{\mathbb{R}}$ and the (locally closed-graph) constraint mapping $G\colon X \rightrightarrows Y$ act in the Asplund space setting.

(i) Show that the results of Theorem 4.1(i) hold in this setting with $D^*G = D^*_N G$ if either φ is SNEC at $(\bar{x}, \bar{y})$ or G is SNC at this point.

(ii) Clarify whether the normal coderivative of G can be replaced by the mixed one in (i). Are any changes needed in the Asplund space version of Corollary 4.3?

(iii) Assuming that $\varphi = \varphi(y)$ and G^{-1} is PSNC at $(\bar{y}, \bar{x})$ instead of the SNC property property of G imposed in (i), verify that the inclusions

$$\partial\mu(\bar{x}) \subset \bigcup_{y^* \in \partial\varphi(\bar{y})} D^*_N G(\bar{x}, \bar{y})(y^*), \quad \partial^\infty\mu(\bar{x}) \subset \bigcup_{y^* \in \partial^\infty\varphi(\bar{y})} D^*_N G(\bar{x}, \bar{y})(y^*)$$

hold under validity of the mixed qualification condition

$$\partial^\infty\varphi(\bar{y}) \cap D^*_M G^{-1}(\bar{y}, \bar{x})(0) = \{0\}$$

replacing the normal one (4.3) in assertion (i).

(iv) Derive Asplund space versions of the results in Theorem 4.1(ii).

(v) Show that if φ is locally Lipschitzian around $(\bar{x}, \bar{y})$ and M is inner semicontinuous at this point, then we have

$$\partial^\infty\mu(\bar{x}) \subset D^*_M G(\bar{x}, \bar{y})(0). \tag{4.30}$$

Obtain a counterpart of this statement when M is merely inner semicompact at $\bar{x}$.

Hint: To justify (i,iii,iv), proceed similarly to the proof of Theorem 4.1 with employing the subdifferential sum rules in Asplund spaces. To verify (v), use the singular subdifferential description (1.71) from Exercise 1.68, and then apply the fuzzy sum rule from Exercise 2.26. Compare it with the proofs of [522, Theorem 3.38].

Exercise 4.21 (Extended Inner Semicontinuity and Inner Semicompactness of Set-Valued Mappings). Given $\mu\colon X \to \overline{\mathbb{R}}$ finite at $\bar{x}$, we say that a mapping $F\colon X \rightrightarrows Y$ between Banach spaces is *μ-inner semicontinuous* at $(\bar{x}, \bar{y}) \in$ gph if for every sequence $x_k \xrightarrow{\text{dom}\, F} \bar{x}$ with $\mu(x_k) \to \mu(\bar{x})$ there is a sequence of $y_k \in F(x_k)$ converging to $\bar{y}$. This mapping is *μ-inner semicompact* at $\bar{x}$ if for every sequence $x_k \xrightarrow{\mu} \bar{x}$ there is a sequence $y_k \in M(x_k)$ that contains a convergent subsequence.

(i) Obtain extensions of the results presented in Theorem 4.1 and Exercise 4.20 to the cases where the argminimum mapping M is assumed to be μ-inner semicontinuous and μ-inner semicompact, respectively.

(ii) Construct examples showing that the results from (i) under the extended inner semicontinuity and semicompactness assumptions strictly improve the corresponding ones from Theorem 4.1 and Exercise 4.20.

Exercise 4.22 (Equality Representations for Subgradients of Marginal Functions). Let X and Y be arbitrary Banach spaces, and let the cost function φ in (4.1) be Fréchet differentiable at $(\bar{x}, \bar{y}) \in \text{gph}\, M$. Assume that the argminimum mapping (4.2) admits an *upper Lipschitzian selector* near $(\bar{x}, \bar{y})$, i.e., there is $h\colon \text{dom}\, G \to Y$ such that $h(\bar{x}) = \bar{y}$ and $h(x) \in M(x)$ for all x in a neighborhood of $\bar{x}$.

(i) Show that in this case we have the equality

$$\widehat{\partial}\mu(\bar{x}) = \nabla_x\varphi(\bar{x}, \bar{y}) + \widehat{D}^*G\big(\bar{x}, \bar{y}\big)\big(\nabla_y\varphi(\bar{x}, \bar{y})\big).$$

(ii) Assume in addition that both X and Y are Asplund, that φ is strictly differentiable at $(\bar{x}, \bar{y})$, that M is μ-inner semicontinuous at $(\bar{x}, \bar{y})$, and that G is N-regular at this point. Show that in this case μ is lower regular at $\bar{x}$ and we have

$$\partial\mu(\bar{x}) = \nabla_x\varphi(\bar{x}, \bar{y}) + D^*_N G\big(\bar{x}, \bar{y}\big)\big(\nabla_y\varphi(\bar{x}, \bar{y})\big).$$

Hint: To verify (i), proceed by using the definitions. The inclusion "$\subset$" in the formula of (ii) is taken from Exercise 4.20, while the opposite inclusion therein follows from (i) under the imposed N-regularity assumption on G.

Exercise 4.23 (Regular Subgradients of Optimal Value Functions for Parametric Nonlinear Programs). Consider the marginal function (4.1) with the constraint mapping $G\colon X \rightrightarrows Y$ given by

$$G(x) := \Big\{y \in Y \Big|\ \varphi_i(x,y) \le 0 \ \text{ for } \ i = 1,\ldots,m, \quad \varphi_i(x,y) = 0 \ \text{ for } \ i = m+1,\ldots,m+r\Big\}, \tag{4.31}$$

where μ from (4.1) is known in this case as the (*optimal*) *value function* for mathematical programs with finitely many inequality and equality constraints.

(i) Let X and Y be Banach. Given $(\bar{x},\bar{y}) \in \operatorname{gph} M$, suppose that all the functions φ_i are Fréchet differentiable at $(\bar{x},\bar{y})$ and continuous around this point, and then define the following sets of Lagrange multipliers $\lambda = (\lambda_1,\ldots,\lambda_{m+r}) \in \mathbb{R}^{m+r}$ by

$$\Lambda(\bar{x},\bar{y}) := \Big\{\lambda \in \mathbb{R}^{m+r} \Big|\ \nabla_y\varphi(\bar{x},\bar{y}) + \sum_{i=1}^{m+r}\lambda_i\nabla_y\varphi_i(\bar{x},\bar{y}) = 0, \quad \lambda_i \ge 0,\ \lambda_i\varphi_i(\bar{x},\bar{y}) = 0 \ \text{ for } \ i = 1,\ldots,m\Big\}, \tag{4.32}$$

$$\Lambda(\bar{x},\bar{y},y^*) := \Big\{\lambda \in \mathbb{R}^{m+r} \Big|\ y^* + \sum_{i=1}^{m+r}\lambda_i\nabla_y\varphi_i(\bar{x},\bar{y}) = 0,\ \lambda_i \ge 0, \quad \lambda_i\varphi_i(\bar{x},\bar{y}) = 0 \ \text{ for } \ i = 1,\ldots,m\Big\}, \quad y^* \in Y^*.$$

Assuming that $\widehat{\partial}^+\varphi(\bar{x},\bar{y}) \ne \emptyset$ for the cost function in (4.1) and the LICQ condition from Exercise 3.73(i) holds for φ_i, $i = 1\ldots,m+r$, at $(\bar{x},\bar{y})$, prove the inclusion

$$\widehat{\partial}\mu(\bar{x}) \subset \bigcap_{(x^*,y^*)\in\widehat{\partial}^+\varphi(\bar{x},\bar{y})}\ \bigcup_{\lambda\in\Lambda(\bar{x},\bar{y},y^*)}\Big[x^* + \sum_{i=1}^{m+r}\lambda_i\nabla_x\varphi_i(\bar{x},\bar{y})\Big]. \tag{4.33}$$

(ii) Assuming in addition to (i) that φ is Fréchet differentiable at $(\bar{x},\bar{y})$ and the solution map (4.2) admits an upper Lipschitzian selector around this point, show that (4.33) holds as the equality:

$$\widehat{\partial}\mu(\bar{x}) = \bigcup_{\lambda\in\Lambda(\bar{x},\bar{y})}\Big[\nabla_x\varphi(\bar{x},\bar{y}) + \sum_{i=1}^{m+r}\lambda_i\nabla_x\varphi_i(\bar{x},\bar{y})\Big].$$

(iii) Let in the setting of (i) the spaces X and Y be Asplund, and let the functions φ_i, $i = 1,\ldots,m+r$, be strictly differentiable at $(\bar{x},\bar{y})$. Show that in this case we have inclusion (4.33) under the MFCQ condition from Exercise 2.53 imposed on the constraint functions φ_i of two variables. Does it hold if φ_i are merely Fréchet differentiable at $(\bar{x},\bar{y})$ and the spaces X and Y are Asplund?

Hint: To verify (i), first check that

$$\widehat{\partial}\mu(\bar{x}) \subset \bigcap_{(x^*,y^*)\in\widehat{\partial}^+\varphi(\bar{x},\bar{y})}\Big[x^* + \widehat{D}^*G(\bar{x},\bar{y})(y^*)\Big] \tag{4.34}$$

in the general framework of (4.1) provided that $\widehat{\partial}^+\varphi(\bar{x},\bar{y})\neq\emptyset$. This can be done by using the smooth variational description of regular subgradients from Theorem 1.27 and Exercise 1.64(i). Then employ in the setting of (4.31) the representation of regular normals to inverse images of graphs under Fréchet differentiable mappings with surjective derivatives; see Exercise 1.54(i). The equality representation in (ii) follows from Exercise 4.22(i) under the imposed assumptions. Assertion (iii) follows from (4.34) and Exercise 3.73(ii). Compare this with [547, Theorem 4 and Corollary 2].

Exercise 4.24 (Subgradients of Optimal Value Functions for Parametric Nondifferentiable Programs). Let $G\colon X\rightrightarrows Y$ in (4.1) be given by (4.31), let X and Y be Asplund, and let φ_i, $i=1,\ldots,m+r$, be locally Lipschitzian around $(\bar{x},\bar{y})\in\operatorname{gph}M$. Assume that only $(\lambda_1,\ldots,\lambda_{m+r})=0$ satisfies the relationships

$$\begin{aligned}&0\in\sum_{i=1}^{m}\lambda_i\partial\varphi_i(\bar{x},\bar{y})+\sum_{i=m+1}^{m+r}\lambda_i\big(\partial\varphi_i(\bar{x},\bar{y})\cup\partial(-\varphi_i)(\bar{x},\bar{y})\big),\\&(\lambda_1,\ldots,\lambda_{m+r})\in\mathbb{R}^{m+r}_+,\;\;\lambda_i\varphi_i(\bar{x},\bar{y})=0\;\text{ as }\;i=1,\ldots,m.\end{aligned}\tag{4.35}$$

(i) Show that if $\widehat{\partial}^+\varphi(\bar{x},\bar{y})\neq\emptyset$, then we have the inclusion

$$\begin{aligned}\widehat{\partial}\mu(\bar{x})\subset\bigcap_{(x^*,y^*)\in\widehat{\partial}^+\varphi(\bar{x},\bar{y})}\Big\{u^*\in X^*\Big|\;(u^*,0)\in(x^*,y^*)+\sum_{i=1}^{m}\lambda_i\partial\varphi_i(\bar{x},\bar{y})\\+\sum_{i=m+1}^{m+r}\lambda_i\big(\partial\varphi_i(\bar{x},\bar{y})\cup\partial(-\varphi_i)(\bar{x},\bar{y})\big)\Big\},\end{aligned}$$

where the multipliers $(\lambda_1,\ldots,\lambda_{m+r})$ are taken from (4.35).

(ii) Show that if φ is locally Lipschitzian around $(\bar{x},\bar{y})$ and M is μ-inner semicontinuous at this point, then with $(\lambda_1,\ldots,\lambda_{m+r})$ from (4.35), we have the inclusions

$$\begin{aligned}\partial\mu(\bar{x})\subset\Big\{u^*\in X^*\Big|\;(u^*,0)\in\partial\varphi(\bar{x},\bar{y})+\sum_{i=1}^{m}\lambda_i\partial\varphi_i(\bar{x},\bar{y})\\+\sum_{i=m+1}^{m+r}\lambda_i\big(\partial\varphi_i(\bar{x},\bar{y})\cup\partial(-\varphi_i)(\bar{x},\bar{y})\big)\Big\},\end{aligned}$$

$$\begin{aligned}\partial^\infty\mu(\bar{x})\subset\Big\{u^*\in X^*\Big|\;(u^*,0)\in\sum_{i=1}^{m}\lambda_i\partial\varphi_i(\bar{x},\bar{y})\\+\sum_{i=m+1}^{m+r}\lambda_i\big(\partial\varphi_i(\bar{x},\bar{y})\cup\partial(-\varphi_i)(\bar{x},\bar{y})\big)\Big\}.\end{aligned}$$

(iii) If φ and φ_i, $i=1,\ldots,m+r$, are strictly differentiable at $(\bar{x},\bar{y})$ and the MFCQ condition is satisfied at this point, we have the inclusions

$$\partial\mu(\bar{x})\subset\bigcup_{\lambda\in\Lambda(\bar{x},\bar{y})}\Big[\nabla_x\varphi(\bar{x},\bar{y})+\sum_{i=1}^{m+r}\lambda_i\nabla_x\varphi_i(\bar{x},\bar{y})\Big],$$

$$\partial^\infty\mu(\bar{x})\subset\bigcup_{\lambda\in\Lambda^\infty(\bar{x},\bar{y})}\Big[\sum_{i=1}^{m+r}\lambda_i\nabla_x\varphi_i(\bar{x},\bar{y})\Big],$$

where $\Lambda(\bar{x}, \bar{y})$ is taken from (4.32) and where by $\Lambda^\infty(\bar{x}, \bar{y})$ is defined by

$$\Big\{\lambda \in \mathbb{R}^{m+r} \Big| \sum_{i=1}^{m+r} \lambda_i \nabla_y \varphi_i(\bar{x}, \bar{y}) = 0,\ \lambda_i \ge 0,\ \lambda_i \varphi_i(\bar{x}, \bar{y}) = 0 \ \text{ for } \ i = 1, \ldots, m\Big\}.$$

Hint: Deduce it from the subdifferential inclusions for marginal functions given in (4.34) and Exercise 4.20(i), respectively, due to the coderivative for G in (4.31) taken from Exercise 3.74. Compare this with [547, Theorems 5 and 7].

Exercise 4.25 (Lipschitz Continuity of Marginal Functions in Finite and Infinite Dimensions). Consider the framework of Corollary 4.3, where $\varphi\colon X \to \overline{\mathbb{R}}$ and $G\colon X \rightrightarrows Y$ act between Asplund spaces.

(i) Present verifiable conditions in terms of the given data φ and G ensuring that the marginal function (4.1) is l.s.c. around $\bar{x} \in \operatorname{dom} \mu$.

(ii) Show that assertion (i) of Corollary 4.3 holds without any change in Asplund spaces, while assertion (ii) therein also holds with replacing the local boundedness of M around $\bar{x}$ by the the local semicompactness property of M at this point.

(iii) Verify that assertions (i) and (ii) of Corollary 4.3 are satisfied under the less restrictive μ-inner semicontinuity and μ-inner semicompactness assumptions on M at $(\bar{x}, \bar{y})$ and $\bar{x}$, respectively; see Exercise 4.21 for the definitions.

(iv) Derive sufficient conditions for local Lipschitz continuity of optimal value functions in problems of mathematical programming with equality and inequality constraints described by Lipschitzian and smooth functions. Show, in particular, that $\mu(x)$ is locally Lipschitzian around $\bar{x}$ under the Mangasarian-Fromovitz constraint qualification in the classical nonlinear programming in finite dimensions.

Hint: To verify (i)–(iii), proceed as in the proof of Corollary 4.3 by using the Asplund space results from Exercises 3.44, 3.45 and the subdifferential description of the SNEC property in Exercise 2.50; compare this with the proof of [532, Theorem 5.2]. To get (iv), use the inclusion for $\partial^\infty \mu(\bar{x})$ obtained in Exercise 4.24 together with the subdifferential characterization of local Lipschitz continuity from the last statement of Theorem 4.15 in $\mathbb{R}^n$ and Exercise 4.34(ii) in Asplund spaces.

Exercise 4.26 (Subdifferentials of Infimal Convolutions in Asplund Spaces). Establish extensions of Theorem 4.4 to infimal convolutions (4.6) of functions $\varphi_1, \varphi_2\colon X \to \overline{\mathbb{R}}$ defined on Asplund spaces. *Hint:* Proceed as in the proof of Theorem 4.4 with applying the corresponding results from Exercise 4.20.

Exercise 4.27 (Subdifferentiation of Marginal Functions and Infimal Convolutions in Finite-Dimensional and Infinite-Dimensional Convex Settings).

(i) Consider the class of marginal functions (4.1), where $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ is convex, and where $G\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ has the convex graph. Show that for any $\bar{x} \in \operatorname{dom} M$ we have the *equality* in (4.4) whenever $\bar{y} \in M(\bar{x})$ is such that the qualification condition (4.3) is satisfied (in particular, when the cost function φ is continuous at $(\bar{x}, \bar{y})$) without imposing any additional assumptions.

(ii) Verify that the first inclusion in Theorem 4.4(i) holds as *equality* for any $(\bar{x}_1, \bar{x}_2) \in C(\bar{x})$ without any additional assumptions.

(iii) Establish extensions of assertions (i) and (ii) to arbitrary Banach spaces.

Hint: To justify (i), proceed by the definitions of the convex constructions involved and the subdifferential sum rule of convex analysis. Derive (ii) as a consequence of (i), and compare this with the proofs of [537, Theorem 2.61 and Corollary 2.65]. Verify that this approach works in arbitrary Banach spaces.

Exercise 4.28 (Subgradients of Compositions in Infinite Dimensions).

(i) Show that the results of Theorem 4.5(i) hold in the case of Asplund spaces X and Y under the additional assumptions that either φ is SNEC at $(\bar{x}, \bar{y})$ or g is SNC at $\bar{x}$. Verify furthermore that the $\mathcal{C}^1$ property of φ and g in the regularity statements can be replaced by the strict differentiability requirement on φ and g at the corresponding points. Justify finally yet another case of the equality in (4.10): g is N-regular at $\bar{x}$ and $\dim Y < \infty$. *Hint:* Proceed as in the proof of Theorem 4.5(i) with employing infinite-dimensional extensions of the facts used therein, which are discussed in the exercises above. Compare this with the proof in [522, Theorem 3.41].

(ii) Verify that the equality in (4.12) of Theorem 4.5(ii) holds for any Banach spaces X and Y provided that φ is strictly differentiable at $(\bar{x}, \bar{y})$. Show in addition that $\varphi \circ g$ is lower regular at $\bar{x}$ if g is M-regular at this point. Is the latter assumption essential for the lower regularity of φ at $\bar{x}$ in finite dimensions? *Hint:* To justify (4.12), proceed as in the proof of Theorem 4.5(ii) and compare it with [522, Theorem 1.110].

Exercise 4.29 (Subdifferential Product and Quotient Rules in Infinite Dimensions). Show that the first equalities in the product and quotient rules of Corollaries 4.7 and 4.8 hold in arbitrary Banach spaces, while the inclusion and regularity statements therein are valid in the Asplund space setting. *Hint:* Proceed as in the proofs of Corollaries 4.7 and 4.8 with the usage of the infinite-dimensional chain and sum rules from Exercises 4.28(ii) and 2.54(i), respectively.

Exercise 4.30 (Partial Subgradients). Let both spaces X and Y be Asplund, and let the function $\varphi\colon X \times Y \to \overline{\mathbb{R}}$ enjoy the SNEC property at $(\bar{x}, \bar{y}) \in \operatorname{dom} \varphi$ and satisfy the qualification condition

$$\big[(0, y^*) \in \partial^\infty \varphi(\bar{x}, \bar{y})\big] \Longrightarrow y^* = 0.$$

(i) Prove that the following hold for partial basic and singular subdifferentials:

$$\partial_x \varphi(\bar{x}, \bar{y}) \subset \big\{x^* \in X^* \big|\ \exists y^* \in Y^* \text{ with } (x^*, y^*) \in \partial \varphi(\bar{x}, \bar{y})\big\}, \tag{4.36}$$

$$\partial_x^\infty \varphi(\bar{x}, \bar{y}) \subset \big\{x^* \in X^* \big|\ \exists y^* \in Y^* \text{ with } (x^*, y^*) \in \partial^\infty \varphi(\bar{x}, \bar{y})\big\}. \tag{4.37}$$

(ii) Check that both assumptions imposed above on φ are satisfied whenever φ is locally Lipschitzian around $(\bar{x}, \bar{y})$, and then give examples of non-Lipschitzian functions for which (4.36) and (4.37) hold.

(iii) Show that the inclusions in both (4.36) and (4.37) may be strict, while the equality holds in (4.36) if φ is lower regular at $(\bar{x}, \bar{y})$. Verify that the equality holds in (4.37) if φ is epigraphically regular at $(\bar{x}, \bar{y})$, and show furthermore that $\varphi(\cdot, \bar{y})$ is lower regular (resp. epigraphically regular) at $\bar{x}$ provided that φ possess the corresponding property at $(\bar{x}, \bar{y})$.

Hint: Represent $\varphi(x, \bar{y})$ in the composition form $(\varphi \circ g)(x)$ with $g(x) := (x, \bar{y})$, and then apply the results of Exercise 4.28(i).

Exercise 4.31 (Regular and Limiting Subgradients of Minimum Functions).

(i) Show that the inclusions in (4.15) and (4.16) hold in arbitrary Banach spaces. *Hint:* Proceed as in the proof of Proposition 4.9 by using representation (1.68) of basic subgradients and definition (1.70) of singular subgradients in Banach spaces.

(ii) Does inclusion (4.16) hold if the geometric representation(1.72) is taken as the definition of the singular subdifferential in Banach spaces?

(iii) Verify the following equality in terms of regular subgradients

$$\widehat{\partial}(\min \varphi_i)(\bar{x}) = \bigcap_{i \in I_{\min}(\bar{x})} \widehat{\partial} \varphi_i(\bar{x})$$

for any l.s.c. functions φ_i on Banach spaces. *Hint:* Compare [329, Proposition 2.5].

(iv) Give examples showing that neither the equalities in (4.15) and (4.16) nor counterparts of the equality in (iii) hold for basic and singular subgradients of l.s.c. functions defined on finite-dimensional spaces.

(v) Obtain precise equalities for calculating regular and basic subdifferentials for minima of two convex polyhedral functions in finite dimensions. *Hint:* Compare this with the results and proofs in [329, Theorems 3.1–3.3].

Exercise 4.32 (Subgradients of Maximum Functions on Asplund Spaces). Derive all the statements of Theorem 4.10 for functions defined on Asplund spaces and compare them with [522, Theorem 3.46].

Exercise 4.33 (Symmetric Subdifferential Mean Value Theorem in Asplund Spaces). Show that Theorem 4.11 holds in Asplund spaces provided that both φ and $-\varphi$ are SNEC at every $x \in (a, b)$. Are any changes needed in the formulation of Corollary 4.12 in the case of functions on Asplund spaces? *Hint:* Proceed as in the proofs of the finite-dimensional versions with the usage of the chain rule from Exercise 4.28(i). Compare this with [522, Theorem 3.47 and Corollary 3.48].

Exercise 4.34 (Approximate Mean Value Theorem and Some of Its Applications in the Framework of Asplund Spaces).

(i) Show that AMVT and its consequences from Section 4.4, with the exception of the last assertion of Theorem 4.15, hold without any changes in Asplund spaces.

(ii) Verify that all the assertions of Theorem 4.15 are equivalent to the validity of $\partial^\infty\varphi(\bar{x}) = \{0\}$ together with the SNEC property of φ at $\bar{x}$.

Hint: Proceed similarly to the proofs given in finite dimensions with employing the corresponding calculus results in Asplund spaces from the exercises above; compare with [522, Subsection 3.2.2].

Exercise 4.35 (Approximate Mean Value Theorem via Basic Subgradients).

(i) Show that AMVT and its consequences in Section 4.4 hold with replacing regular subgradients by basic subgradients.

(ii) Are the latter results in terms of basic subgradients are equivalent to those given via regular subgradients in finite dimensions and in Asplund spaces?

Exercise 4.36 (Relationships Between Basic Normals and Subgradients and Their Clarke Counterparts in Asplund Spaces). Show that the following assertions hold in any Asplund space X:

(i) Let $\varphi\colon X \to \overline{\mathbb{R}}$ be locally Lipschitzian around $\bar{x}$. Then

$$\overline{\partial}\varphi(\bar{x}) = \mathrm{cl}^*\mathrm{co}\,\partial\varphi(\bar{x}),$$

where generalized gradient $\overline{\partial}\varphi(\bar{x})$ of a locally Lipschitz function is defined by (1.78) and where the basic subdifferential $\partial\varphi(\bar{x})$ is represented by (1.69) in Asplund spaces.

(ii) Let $\bar{x} \in \Omega \subset X$, where Ω is locally closed around $\bar{x} \in \Omega$ as in our standing standing assumption. Then we have

$$\overline{N}(\bar{x}; \Omega) = \mathrm{cl}^*\mathrm{co}\,N(\bar{x}; \Omega),$$

where the Clarke normal cone $\overline{N}(\bar{x}; \Omega)$ is taken from (1.80).

(iii) Let $\varphi\colon X \to \overline{\mathbb{R}}$ be l.s.c. around $\bar{x}$ as in our standing assumption. Then

$$\overline{\partial}\varphi(\bar{x}) = \mathrm{cl}^*\mathrm{co}\big[\partial\varphi(\bar{x}) + \partial^\infty\varphi(\bar{x})\big],$$

where $\overline{\partial}\varphi(\bar{x})$ is defined by (1.81) and where $\partial^\infty\varphi(\bar{x})$ is taken from (1.71).

Hint: First justify (i) by applying AMVT to the limiting description (1.77) of the generalized directional derivative $\varphi^\circ(\bar{x}; h)$ in the generalized gradient constriction (1.78), and then proceed with (ii) and (iii) by using the definitions therein; compare this with the proof of [522, Theorem 3.57].

4.6 Commentaries to Chapter 4

Sections 4.1–4.3. *Marginal/optimal value functions* constitute one of the most fundamental objects of variational analysis. They have never been seriously investigated in frameworks of classical analysis due to their intrinsic nonsmoothness, which is always the case unless quite restrictive and unnatural assumptions are imposed. This was actually what L. C. Young meant by observing in the 1930s that, roughly speaking, the limitations of many results of the calculus of variations came from the absence of an adequate nonsmooth analysis; see [753]. It would not be an exaggeration to say that marginal functions manifest the essence of modern techniques in variational analysis involving *perturbation and approximation procedures* with the subsequent passing to the limit. Subdifferentiation of marginal functions evaluates *rates of change* under parameter perturbations, which is crucial for sensitivity analysis while in fact leads us to a much larger scope of applications as, in particular, shown above. Besides sensitivity issues, subdifferential analysis of marginal functions has been recognized as an important machinery for the study of viscosity and minimax solutions of Hamilton-Jacobi equations, deterministic and stochastic dynamic programming, feedback control design, differential game theory, deterministic and stochastic optimal control, bilevel programming, economic growth modeling, etc.; see, e.g., [67, 93, 100, 117, 165, 167, 195, 198, 199, 215, 268, 271, 416, 425, 522, 540, 629, 698, 699, 712, 713, 729, 748] with more discussions and references therein.

The principal result of Sections 4.1–4.3 is Theorem 4.1 on the subdifferential estimates for marginal functions (4.1). It was obtained in full generality of finite-dimensional spaces in the author's paper [508], while the basic subdifferential estimate (4.4) with $\varphi(x, y) = \varphi(y)$ was established by the author earlier [505, 507]. In the unconstrained case of $G(x) = \mathbb{R}^m$ in (4.1), both basic and singular subdifferential estimates were given by Rockafellar [675]; cf. also [672]. The full Asplund space extension of Theorem 4.1 can be found in the paper by the author and Shao [580], while some previous results were derived by Thibault [706] in Fréchet smooth spaces; see also [14, 117, 532, 546, 547] for more recent developments and applications.

When the mapping G in (4.1) is *single-valued*, the subdifferential formulas of Theorems 4.1 and 4.5 evaluate basic and singular subgradients of *generalized compositions*. Moreover, the singular subdifferential estimate (4.5) in the *set-valued* case of G allows us to obtain verifiable conditions ensuring the local *Lipschitz continuity* of marginal functions due to its singular subdifferential characterization of Theorem 1.22; see more discussions in Section 1.5. The latter direction has been largely explored, e.g., in [472, 508, 512, 513, 522, 532, 600, 603, 672, 675, 678, 729].

The subdifferential *chain rules*, where $\varphi(x, y) = \varphi(y)$ in the composition, and related results presented in Sections 4.1–4.3 under the general assumptions imposed therein are mainly based on the author's developments from [505, 507]. Their Lipschitzian counterparts were derived by Kruger [428, 430] in Fréchet smooth spaces; see also Ioffe [365] for parallel Lipschitzian results concerning certain versions of the "approximate" subdifferentials in Banach spaces. An upper estimate of $\partial(\varphi \circ g)(\bar{x})$ for non-Lipschitzian functions was obtained in [368] under a tangential qualification condition essentially more restrictive in comparison with that in [505]. Asplund space versions of the subdifferential calculus results given in these sections were established by the author and Shao in [580] and then further elaborated in [522, 588]. Let us mention more recent results on calculating the basic subdifferential of the minimum and maximum functions (including equality therein) obtained in [329, 680] with rather surprising applications in [329] to deriving necessary and sufficient conditions for DC (difference of convex) optimization problems. We also refer the reader to [114, 135, 167, 369, 375, 376, 398, 399, 610, 637, 678, 685, 729] for other calculus results involving limiting and "approximate" subgradients.

Note that putting $\varphi(y) = \delta(y; \Theta)$ in the obtained chain rule formulas for either $\partial(\varphi \circ g)$ or $\partial^\infty(\varphi \circ g)$ allows us to evaluate the normal cone to the *inverse image* $N\big(\bar{x}; g^{-1}(\Theta)\big)$ of the set Θ under the mapping g, which in fact was derived in Corollary 3.13 even for set-valued mappings G as a consequence of the chain rule for coderivatives. The first results of the inclusion type for representing the normal cone to *direct images* $G(\Theta)$ of sets under smooth single-valued mappings between finite-dimensional spaces were obtained by Rockafellar [675]; see also [678, Theorem 6.43]. They were significantly extended, for both single-valued and set-valued mappings G, in the joint paper by the author with Nam and Wang [545] in Asplund and general Banach spaces, being also new in finite dimensions. An important role in this derivation (different from [675, 678]) was played by the notion of *restrictive metric regularity* introduced and investigated by the author and Wang [591]. Some of these results in Asplund spaces have been recently reproduced by Penot [637].

Section 4.4. The first mean value theorem for nonsmooth Lipschitzian functions was obtained by Lebourg [449] in terms of Clarke's generalized gradient. The *nonconvex* subdifferential versions as in Theorem 4.11 and Corollary 4.12 go back to Kruger and Mordukhovich (see [428, 431, 505, 507]) and are based on the corresponding subdifferential chain rules for Lipschitzian and non-Lipschitzian functions. The Asplund space version of Theorem 4.11 from [505] was given in [580]; see also [522, Theorem 3.47]. Note that this result requires *two-sided* generalized differential constructions $\partial^0\varphi$ and $\partial^{\infty,0}\varphi$ in both mean value inclusion (4.19) and the supporting qualification condition of the theorem. Nevertheless, it provides an essential improvement of Lebourg's mean value theorem, since the symmetric subdifferential $\partial^0\varphi$ may be much smaller than the generalized gradient even for simple nonsmooth Lipschitzian functions as those considered in Example 1.31.

Approximate mean value theorems (AMVT) of the type presented in Theorem 4.13 are new in analysis being significantly different from the conventional Lagrangian framework. The major difference is that the results of the new type apply to the general class of l.s.c. *extended-real-valued* functions providing *mean value inequalities* instead of equalities or inclusions as in (4.19). The first result of this type was obtained in variational analysis by Zagrodny [756] in terms of Clarke's subgradients of l.s.c. functions defined on Banach spaces. Then Thibault observed [707] that Zagrodny's approach led us in fact to appropriate versions of AMVT for a broad class of subdifferentials (called "presubdifferentials" in [710]) satisfying natural requirements in suitable Banach spaces. The AMVT version in terms of regular and limiting subgradients can also be found in Loewen [471, 472] for l.s.c. functions on Fréchet smooth spaces, while the mean value inequality (4.25) for Lipschitzian functions was obtained earlier by Borwein and Preiss [108] in the same framework. The full Asplund space version of Theorem 4.13 and Corollary 4.14 was given by the author and Shao [580] (see also [522]) with the variational proof presented above and being different in some essential points from those given in [108, 472, 756]. More recently [714] Trang has shown that the Asplund property of the space in question is also *necessary* for the validity of AMVT in the form of [580]. Mean value inequalities of the so-called *multidirectional* type were initiated by Clarke and Ledyaev [166] and further developed in [39, 114, 167, 637] and other publications.

The regular subdifferential characterizations (a) and (b) of local Lipschitz continuity in Theorem 4.15 were given by Loewen [472] in Fréchet smooth spaces and then by the author and Shao [580] in Asplund spaces. The limiting subdifferential characterization of Theorem 4.15(c) in finite dimensions was also obtained by another way in Theorem 1.22 of Chapter 1 and was discussed in Section 1.5. Its Asplund space version (with the additional SNEC property of φ in the last assertion of Theorem 4.15) was given in [522, Theorem 3.52].

The results of Theorems 4.17 and 4.18 are also taken from Loewen [472] (with simplified proofs), where the conditions of Theorem 4.17 were proved to characterize *strict Hadamard differentiability* of functions defined on Fréchet smooth spaces; the latter notion reduces to the usual (Fréchet) strict differentiability in finite dimensions. Asplund space versions of both Theorems 4.17 and 4.18 were given in [580]; see [522] for more details. A proximal subdifferential counterpart of Theorem 4.18 was derived in [168] for l.s.c. functions on Hilbert spaces.

Monotonicity of set-valued mappings has been widely recognized as one of the most important concepts in variational analysis and its applications. We refer the reader to the monograph by Rockafellar and Wets [678, Chapter 12] for a variety of results on monotonicity and detailed comments on the history and genesis of major ideas; see also [37, 70, 112, 116, 126, 130, 185, 323, 486, 638, 656, 689, 690] for some additional material and further applications. A fundamental result of convex analysis and monotone operator theory, which goes back to Minty, Moreau, and finally

Rockafellar (see [678] for more details), is the *maximal monotonicity* of subdifferential mappings $\partial\varphi$ generated by l.s.c. *convex* functions $\varphi\colon X \to \overline{\mathbb{R}}$. Theorem 4.19, which is mainly based on the paper by Correa, Jofré, and Thibault [181] (see also the references therein for the previous results in this direction), shows that the convexity of φ is in fact *necessary* for the monotonicity (even not the maximal one) of $\partial\varphi$. Its Asplund space version was presented in [522, Theorem 3.56]. Furthermore, Daniilidis and Georgiev [189] established the equivalence between the *approximate convexity* of a locally Lipschitzian function on an arbitrary Banach space and the *submonotonicity* of its (Clarke) generalized gradient at the point in question.

Section 4.5. As in the case of the previous chapters, the material included in this exercise section presents some additional results and infinite-dimensional extensions of the basic facts and proofs given in Sections 4.1–4.3. The reader can find more information in the references included in the hints to the corresponding exercises and also in the above commentaries on the main theorems.

Chapter 5
Coderivatives of Maximal Monotone Operators

In this chapter we employ the tools of variational analysis and generalized differentiation developed above to study global and local monotonicity of set-valued operators. Our main attention is paid to the properties of *global maximal monotonicity* and *strong local maximal monotonicity*, which both have been well recognized as fundamental notions in many areas of nonlinear analysis, optimization, variational inequalities, and numerous applications. The main results below provide complete *coderivative characterizations* of the monotonicity concepts under consideration for the general class of set-valued operators. Although we present these characterizations in finite dimensions, they hold with minimal adjustments (if any) in the framework of Hilbert spaces. Among other things, the mean value inequality from Corollary 4.14(i) plays a crucial role in the proofs of the obtained coderivative characterizations.

5.1 Coderivative Criteria for Global Monotonicity

We begin with the study of *global monotonicity* while recalling that the definitions of (globally) *monotone* and *maximal monotone* operators $T\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ have been already presented in (4.27) of Section 4.4, where we characterized monotonicity of subdifferential operators.

5.1.1 Maximal Monotonicity via Regular Coderivative

The following *hypomonotonicity* properties of set-valued operators play a significant role in the subsequent results of this chapter.

B. S. Mordukhovich, *Variational Analysis and Applications*, Springer Monographs in Mathematics, https://doi.org/10.1007/978-3-319-92775-6_5

Definition 5.1 (Hypomonotonicity). *Let $T\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be a set-valued mapping, and let $I\colon \mathbb{R}^n \to \mathbb{R}^n$ be the identity operator on $\mathbb{R}^n$. We say that:*

(i) *T is* GLOBALLY HYPOMONOTONE *on $\mathbb{R}^n$ if there is $r > 0$ such that the mapping $T + rI$ is monotone on $\mathbb{R}^n$. This means that the inequality*

$$\langle v_1 - v_2, u_1 - u_2 \rangle \ge -r\|u_1 - u_2\|^2 \tag{5.1}$$

holds for all pairs $(u_1, v_1), (u_2, v_2) \in \operatorname{gph} T$.

(ii) *T is* SEMILOCALLY HYPOMONOTONE *around $\bar{x} \in \operatorname{dom} T$ if there exist a neighborhood U of $\bar{x}$ and $r > 0$ such that (5.1) holds for all $(u_1, v_1), (u_2, v_2) \in \operatorname{gph} T \cap (U \times \mathbb{R}^n)$. We say that T is semilocally hypomonotone* ON *$\Omega \subset \mathbb{R}^n$ if it is semilocally hypomonotone around each $\bar{x} \in \Omega \cap \operatorname{dom} T$.*

(iii) *T is* LOCALLY HYPOMONOTONE *around $(\bar{x}, \bar{v}) \in \operatorname{gph} T$ if there exist a neighborhood $U \times V$ of $(\bar{x}, \bar{v})$ and a number $r > 0$ such that (5.1) holds for all pairs $(u_1, v_1), (u_2, v_2) \in \operatorname{gph} T \cap (U \times V)$.*

Note that the classes of hypomonotone operators of all the three types defined above are fairly broad while containing, in particular, locally monotone operators, Lipschitz continuous single-valued mappings, and also subgradient mappings generated by continuously prox-regular functions that are especially important in the framework of second-order variational analysis; see Section 3.5 for more discussions and references.

The following theorem characterizes the global maximal monotonicity of set-valued operators via their global hypomonotonicity and the *positive-semidefiniteness* condition for their regular coderivatives (1.16).

Theorem 5.2 (Regular Coderivative and Global Hypomonotonicity Criterion for Maximal Monotonicity). *Let $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be a set-valued mapping with closed graph. The following assertions are equivalent:*

(i) *T is globally maximal monotone on $\mathbb{R}^n$.*

(ii) *T is globally hypomonotone on $\mathbb{R}^n$ and we have*

$$\langle z, w \rangle \ge 0 \ \textit{ whenever } \ z \in \widehat{D}^*T(u, v)(w) \ \textit{ and } \ (u, v) \in \operatorname{gph} T. \tag{5.2}$$

Proof. To verify (i)⇒(ii), it suffices to show that the maximal monotonicity of T implies the positive-semidefiniteness condition (5.2) by taking into account that the hypomonotonicity of T in (ii) obviously follows from its monotonicity. We proceed by recalling the classical Minty theorem (see, e.g., [70, Theorem 21.1]), which tells us that the maximal monotonicity of T ensures that for any $\lambda > 0$ the *resolvent* $R_\lambda = (I + \lambda T)^{-1}$ is *single-valued* and *nonexpansive* (i.e., globally Lipschitz continuous on its domain with constant $\ell = 1$) and that $\operatorname{dom} R_\lambda = \mathbb{R}^n$. Picking an arbitrary pair $(w, z) \in \operatorname{gph} \widehat{D}^*T(u, v)$, we deduce from the sum rule for the regular coderivative in Exercise 3.59(i) that

$$-\lambda^{-1}w \in \widehat{D}^*R_\lambda(u + \lambda v, u)(-z - \lambda^{-1}w).$$

Since R_λ is nonexpansive, it follows from the neighborhood version of the coderivative criterion for the Lipschitz-like property in Theorem 3.3(iii) (see Exercise 3.41 and [522, Theorem 4.7] in Asplund spaces) that $\|-\lambda^{-1}w\| \le \|-z-\lambda^{-1}w\|$, which clearly implies that

$$\lambda^{-2}\|w\|^2 \le \|-z-\lambda^{-1}w\|^2 = \|z\|^2 + 2\lambda^{-1}\langle z, w\rangle + \lambda^{-2}\|w\|^2$$

and yields in turn that $0 \le \lambda\|z\|^2 + 2\langle z, w\rangle$ for all $\lambda > 0$. Letting $\lambda \downarrow 0$ tells us that $\langle z, w\rangle \ge 0$ and thus justifies (5.2).

To verify the converse implication (ii)$\Rightarrow$(i), suppose that T is hypomonotone and that condition (5.2) is satisfied. Then there is some number $r > 0$ such that $T + rI$ is monotone. Take any $s > r$ and define $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ by $\operatorname{gph} F := \operatorname{gph}(T + sI)^{-1}$. For any $(v_i, u_i) \in \operatorname{gph} F$, $i = 1, 2$, we have $(u_i, v_i - su_i) \in \operatorname{gph} T$ and thus deduce from (5.1) that

$$\langle v_1 - su_1 - v_2 + su_2, u_1 - u_2\rangle \ge -r\|u_1 - u_2\|^2.$$

The latter implies in turn that the inequalities

$$\|v_1 - v_2\| \cdot \|u_1 - u_2\| \ge \langle v_1 - v_2, u_1 - u_2\rangle \ge (s - r)\|u_1 - u_2\|^2$$

hold, which allow us to arrive at the estimate

$$\|u_1 - u_2\| \le \frac{1}{s - r}\|v_1 - v_2\| \tag{5.3}$$

verifying that F is single-valued and Lipschitz continuous on its domain with modulus $(s - r)^{-1}$. Fix now any $z \in \mathbb{R}^n$ and define $\varphi_z : \mathbb{R}^n \to \overline{\mathbb{R}}$ by

$$\varphi_z(v) := \begin{cases} \langle z, F(v)\rangle & \text{if } v \in \operatorname{dom} F, \\ \infty & \text{otherwise.} \end{cases} \tag{5.4}$$

Since $\operatorname{gph} T$ is closed, it is easy to check that $\operatorname{gph} F$ is also closed in $\mathbb{R}^n \times \mathbb{R}^n$. Next we show that φ_z is l.s.c. on $\mathbb{R}^n$. Arguing by contradiction, suppose that there exist $\varepsilon > 0$ and a sequence v_k converging to some $v \in \mathbb{R}^n$ such that $\varphi_z(v_k) < \varphi_z(v) - \varepsilon$. If $\varphi_z(v) = \infty$, then $v \notin \operatorname{dom} F$ while $v_k \in \operatorname{dom} F$. It follows from (5.3) that $\|F(v_k) - F(v_j)\| \le (s - r)^{-1}\|v_k - v_j\|$, and so $\{F(v_k)\}$ is a Cauchy sequence converging to some $u \in \mathbb{R}^n$. Hence the sequence $(v_k, F(v_k)) \in \operatorname{gph} F$ converges to $(v, u) \in \operatorname{gph} F$ due to the closedness of $\operatorname{gph} F$. This gives us $F(v) = u$ and contradicts the condition $v \notin \operatorname{dom} F$. In the remaining case of $\varphi_z(v) < \infty$, we get from (5.3) and (5.4) the estimates

$$|\varphi_z(v_k) - \varphi_z(v)| \le \|z\| \cdot \|F(v_k) - F(v)\| \le \|z\| \cdot \frac{1}{s - r}\|v_k - v\| \to 0,$$

which also contradict the assumption $\varphi_z(v_k) < \varphi_z(v) - \varepsilon$. This justifies the lower semicontinuity of φ_z on the space $\mathbb{R}^n$ for any fixed $z \in \mathbb{R}^n$.

To prove now that T is *monotone*, pick two pairs $(u_i, v_i) \in \operatorname{gph} T$, and get

$$(y_i, u_i) \in \operatorname{gph} F \ \text{ with } \ y_i := v_i + su_i, \ i = 1, 2.$$

Applying the mean value inequality (4.25) to φ_z tells us that

$$\begin{aligned} &|\langle z, u_1 - u_2\rangle| = |\varphi_z(y_1) - \varphi_z(y_2)| \\ &\le \|y_1 - y_2\| \sup\big\{\|w\| \;\big|\; w \in \widehat{\partial}\varphi_z(y), \ y \in [y_1, y_2] + \varepsilon\mathbb{B}\big\} \end{aligned} \tag{5.5}$$

for any fixed $\varepsilon > 0$. Since $\widehat{\partial}\varphi_z(y) = \emptyset$ if $y \notin \operatorname{dom}\varphi_z$, it suffices to consider the case where $y \in \operatorname{dom}\varphi_z \cap ([y_1, y_2] + \varepsilon\mathbb{B}) = \operatorname{dom} F \cap ([y_1, y_2] + \varepsilon\mathbb{B})$ in (5.5). Take any y from the latter set, and observe that

$$w \in \widehat{D}^* F(y)(z) \ \text{ whenever } \ w \in \widehat{\partial}\varphi_z(y). \tag{5.6}$$

Indeed, it follows from the definition of $w \in \widehat{\partial}\varphi_z(y)$ that

$$\liminf_{v \to y} \frac{\varphi_z(v) - \varphi_z(y) - \langle w, v - y\rangle}{\|v - y\|} \ge 0,$$

which can be equivalently written by the construction of φ_z in (5.4) as

$$\liminf_{v \xrightarrow{\operatorname{dom} F} y} \frac{\langle z, F(v)\rangle - \langle z, F(y)\rangle - \langle w, v - y\rangle}{\|v - y\|} \ge 0.$$

The latter readily implies that

$$\liminf_{(v,u) \xrightarrow{\operatorname{gph} F} (y, F(y))} \frac{\langle z, u - F(y)\rangle - \langle w, v - y\rangle}{\|v - y\| + \|u - F(y)\|} \ge 0.$$

Hence we get from the definitions in (1.33), (1.5), and (1.16) that

$$(w, -z) \in \widehat{N}\big((y, F(y)); \operatorname{gph} F\big) \Longleftrightarrow w \in \widehat{D}^* F(y)(z) = \widehat{D}^*(T + sI)^{-1}(y)(z),$$

and therefore $-z \in \widehat{D}^*(T + sI)(F(y), y)(-w)$. It easily follows from the elementary sum rule for the regular coderivative in Exercise 3.59(i) that

$$-z + sw \in \widehat{D}^* T\big(F(y), y - sF(y)\big)(-w). \tag{5.7}$$

Combining this with (5.2) tells us that $\langle -z + sw, -w\rangle \ge 0$, which yields

$$\|z\| \cdot \|w\| \ge \langle z, w\rangle \ge s\|w\|^2 \tag{5.8}$$

and implies furthermore together with the estimate (5.5) that

$$|\langle z, u_1 - u_2\rangle| \le s^{-1}\|z\| \cdot \|y_1 - y_2\|.$$

Since this inequality holds for all $z \in \mathbb{R}^n$, we get

$$\|u_1 - u_2\| \le s^{-1}\|y_1 - y_2\| = s^{-1}\|v_1 + su_1 - v_2 - su_2\|$$

and then deduce by the Euclidean norm property that

$$\begin{aligned} s^2\|u_1 - u_2\|^2 &\le \|(v_1 - v_2) + s(u_1 - u_2)\|^2 \\ &= \|v_1 - v_2\|^2 + 2s\langle v_1 - v_2, u_1 - u_2\rangle + s^2\|u_1 - u_2\|^2. \end{aligned}$$

Therefore we arrive at the inequality

$$0 \le \frac{1}{2s}\|v_1 - v_2\|^2 + \langle v_1 - v_2, u_1 - u_2\rangle \ \text{ for any } \ s > r.$$

Passing there to the limit as $s \to \infty$ shows that

$$0 \le \langle v_1 - v_2, u_1 - u_2\rangle \ \text{ for all } \ (u_1, v_1), (u_2, v_2) \in \operatorname{gph} T$$

and thus justifies the monotonicity of the operator T.

It remains to prove that T is *maximal monotone*. Since T is proper, there exists a pair $(u_0, v_0) \in \operatorname{gph} T$ such that

$$u_0 = (T + sI)^{-1}(y_0) \ \text{ with } \ y_0 := v_0 + su_0.$$

Applying again the mean value inequality (4.25) to the function φ_z defined in (5.4), we verify that the estimate

$$|\varphi_z(y) - \varphi_z(y_0)| \le \|y - y_0\| \sup\Big\{\|w\| \ \Big| \ w \in \widehat{\partial}\varphi_z(x), \ x \in [y, y_0] + \varepsilon\mathbb{B}\Big\}$$

is valid for any $y \in \mathbb{R}^n$. It follows similarly to the proof of (5.8) that $\|w\| \le s^{-1}\|z\|$ for all $w \in \widehat{\partial}\varphi_z(x)$ with $x \in \operatorname{dom} F \cap ([y, y_0] + \varepsilon\mathbb{B})$. This gives us due to the above mean value inequality that

$$|\varphi_z(y) - \varphi_z(y_0)| \le s^{-1}\|z\| \cdot \|y - y_0\|.$$

Hence $\varphi_z(y) < \infty$ and so $F(y) \ne \emptyset$ for all $y \in \mathbb{R}^n$, which means that $\operatorname{dom}(T + sI)^{-1} = \mathbb{R}^n$. Employing again the aforementioned Minty theorem and taking into account the monotonicity of T justified above, we conclude that T is maximal monotone and thus complete the proof of the theorem. △

5.1.2 Maximal Monotone Operators with Convex Domains

Our next goal is to obtain another version of the coderivative characterization in Theorem 5.2 with replacing the global hypomonotonicity of T in assertion (ii) by a *semilocal* hypomonotonicity. Establishing such a result requires an additional *convexity* assumption on the *domain* of T, which is shown below to be essential by providing a counterexample. To proceed in this direction, we first present the following lemma, where the *semilocal monotonicity* of $T:\mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is defined as in Definition 5.1(ii) with $r = 0$ in (5.1).

Lemma 5.3 (Semilocal Monotonicity of Set-Valued Mappings with Convex Domains). *Let $T:\mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be semilocally monotone on $\mathbb{R}^n$, and let its domain* $\operatorname{dom} T$ *be convex. Then T is globally monotone on $\mathbb{R}^n$.*

Proof. Pick any $(u_1, v_1), (u_2, v_2) \in \operatorname{gph} T$, and get $[u_1, u_2] \subset \operatorname{dom} T$ by the assumed convexity of $\operatorname{dom} T$. Since T is semilocally monotone, for each vector $x \in [u_1, u_2]$ there is a number $\gamma_x > 0$ such that

$$\langle y_1 - y_2, x_1 - x_2\rangle \ge 0 \ \text{ if } \ (x_1, y_1), (x_2, y_2) \in \operatorname{gph} T \cap \big(\operatorname{int} B_{\gamma_x}(x) \times \mathbb{R}^n\big). \tag{5.9}$$

By compactness of $[u_1, u_2]$, find $x_i \in [u_1, u_2]$ with $i = 1, \ldots, m$ satisfying

$$[u_1, u_2] \subset \bigcup_{i=1}^{m} \operatorname{int}\big(x_i + \gamma_{x_i}\mathbb{B}\big).$$

Thus there exist numbers $0 = t_0 < t_1 < \ldots < t_k = 1$ such that

$$[\widehat{u}_j, \widehat{u}_{j+1}] \subset \operatorname{int}\big(x_i + \gamma_{x_i}\mathbb{B}\big) \ \text{ with some } \ i := i_j \in \big\{1, \ldots, m\big\}$$

for each $j \in \{0, \ldots, k-1\}$, where $\widehat{u}_j := u_1 + t_j(u_2 - u_1)$. Since we have $\widehat{u}_j \in [u_1, u_2] \subset \operatorname{dom} T$ for each $j \in \{0, \ldots, k\}$, there are vectors $\widehat{v}_j \in T(\widehat{u}_j)$ satisfying $\widehat{v}_0 = v_1$ and $\widehat{v}_k = v_2$. It follows from (5.9) that

$$(t_{j+1} - t_j)\langle \widehat{v}_{j+1} - \widehat{v}_j, u_2 - u_1\rangle = \langle \widehat{v}_{j+1} - \widehat{v}_j, \widehat{u}_{j+1} - \widehat{u}_j\rangle \ge 0,$$

which implies that $\langle \widehat{v}_{j+1} - \widehat{v}_j, u_2 - u_1\rangle \ge 0$ whenever $j \in \{0, \ldots, k-1\}$. Hence

$$\langle v_2 - v_1, u_2 - u_1\rangle = \sum_{j=0}^{k-1} \big\langle \widehat{v}_{j+1} - \widehat{v}_j, u_2 - u_1\big\rangle \ge 0,$$

which justifies the global monotonicity of the operator T. △

Now we are ready to obtain a semilocal counterpart of the coderivative characterization in Theorem 5.2 under the convexity assumption on $\operatorname{dom} T$. Example 5.5 below demonstrates that the latter assumption cannot be dropped. Since the proof of the following theorem is similar in some places to that of Theorem 5.2, we omit the corresponding details.

Theorem 5.4 (Regular Coderivative and Semilocal Hypomonotonicity Criterion for Maximal Monotonicity). *Let $T:\mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be a set-valued mapping of closed graph and convex domain. Then the following are equivalent:*

(i) *T is globally maximal monotone on $\mathbb{R}^n$.*

(ii) *T is semilocally hypomonotone on $\mathbb{R}^n$, and the positive-semidefiniteness regular coderivative condition (5.2) is satisfied.*

Proof. Implication (i)⇒(ii) follows from Theorem 5.2. To verify the converse implication, suppose that condition (5.2) holds and that T is semilocally hypomonotone. This allows us to find, for each $\bar{x} \in \operatorname{dom} T$, numbers $\gamma, r > 0$ such that (5.1) is fulfilled whenever $(u_1, v_1), (u_2, v_2) \in \operatorname{gph} T \cap (B_\gamma(\bar{x}) \times \mathbb{R}^n)$. Take now any $s > r$, and define the mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ by

$$\operatorname{gph} F := \operatorname{gph} (T + sI)^{-1} \cap \big(\mathbb{R}^n \times (\bar{x} + \gamma\mathbb{B})\big).$$

Picking arbitrary pairs $(v_i, u_i) \in \operatorname{gph} F$, $i = 1, 2$, we have $(u_i, v_i - su_i) \in \operatorname{gph} T \cap (\mathbb{B}_\gamma(\bar{x}) \times \mathbb{R}^n)$. It follows from the semilocal hypomonotonicity that

$$\langle v_1 - su_1 - v_2 + su_2, u_1 - u_2\rangle \geq -r\|u_1 - u_2\|^2.$$

Similarly to (5.5) we deduce from the latter that

$$\|u_1 - u_2\| \leq \frac{1}{s-r}\|v_1 - v_2\| \ \text{ for all } \ (v_1, u_1), (v_2, u_2) \in \operatorname{gph} F. \tag{5.10}$$

This implies that F is single-valued and Lipschitz continuous on $\operatorname{dom} F$. For any fixed vector $z \in \mathbb{R}^n$ define the function $\varphi_z : \mathbb{R}^n \to \overline{\mathbb{R}}$ as in (5.4) and prove similarly to Theorem 5.2 that φ_z is l.s.c. on $\mathbb{R}^n$.

Pick further arbitrary pairs $(u_1, v_1), (u_2, v_2) \in \operatorname{gph} T \cap (\operatorname{int} B_\gamma(\bar{x}) \times \mathbb{R}^n)$ and fix $\bar{v} \in T(\bar{x})$. Then $F(y_i) = u_i \in \mathbb{B}_\gamma(\bar{x})$ with $y_i := v_i + su_i$. Applying the mean value inequality (4.25) for any $\varepsilon \in (0, \sqrt{s})$ gives us estimate (5.5). Similarly to (5.6) we get $\widehat{\partial}\varphi_z(y) \subset \widehat{D}^*F(y)(z)$ if $y \in \operatorname{dom} F \cap ([y_1, y_2] + \varepsilon\mathbb{B})$ and then for any $y \in \operatorname{dom} F \cap ([y_1, y_2] + \varepsilon\mathbb{B})$ find $y_0 \in \varepsilon\mathbb{B}$ and $t \in [0, 1]$ with $y = ty_1 + (1-t)y_2 + y_0$. Since $F(\bar{v} + s\bar{x}) = \bar{x}$, it follows from (5.10) that

$$\begin{aligned}
&\|F(y) - \bar{x}\| = \|F(ty_1 + (1-t)y_2 + y_0) - F(\bar{v} + s\bar{x})\| \\
&\leq \frac{1}{s-r}\|ty_1 + (1-t)y_2 + y_0 - \bar{v} - s\bar{x}\| \\
&= \frac{1}{s-r}\|t(v_1 + su_1) + (1-t)(v_2 + su_2) + y_0 - \bar{v} - s\bar{x}\| \\
&= \frac{1}{s-r}\|t(v_1 - \bar{v}) + st(u_1 - \bar{x}) + (1-t)(v_2 - \bar{v}) + s(1-t)(u_2 - \bar{x}) + y_0\| \\
&\leq \frac{1}{s-r}\Big[t\|v_1 - \bar{v}\| + (1-t)\|v_2 - \bar{v}\| + st\|u_1 - \bar{x}\| + s(1-t)\|u_2 - \bar{x}\| + \|y_0\|\Big] \\
&\leq \frac{1}{s-r}\Big[\max\big\{\|v_1 - \bar{v}\|, \|v_2 - \bar{v}\|\big\} + \varepsilon\Big] + \frac{s}{s-r}\max\big\{\|u_1 - \bar{x}\|, \|u_2 - \bar{x}\|\big\} \\
&\leq \frac{1}{s-r}\Big[\max\big\{\|v_1 - \bar{v}\|, \|v_2 - \bar{v}\|\big\} + \sqrt{s}\Big] + \frac{s}{s-r}\max\big\{\|u_1 - \bar{x}\|, \|u_2 - \bar{x}\|\big\}.
\end{aligned}$$

Taking now into account that the choice of $(u_1, v_1), (u_2, v_2), (\bar{x}, \bar{v}) \in \operatorname{gph} T \cap (\operatorname{int} B_\gamma(\bar{x}) \times \mathbb{R}^n)$ was independent of the parameter $s > r$ and that $\max\{\|u_1 - \bar{x}\|, \|u_2 - \bar{x}\|\} < \gamma$, we can find a large number $M > 0$ for which

$$\frac{1}{s-r}\max\big\{\|v_1 - \bar{v}\|, \|v_2 - \bar{v}\| + \sqrt{s}\big\} + \frac{s}{s-r}\max\big\{\|u_1 - \bar{x}\|, \|u_2 - \bar{x}\|\big\} < \gamma$$

whenever $s > M$. This together with the above estimate of $\|F(y) - \bar{x}\|$ ensures the inclusion $F(y) \in \operatorname{int} B_\gamma(\bar{x})$ and thus the equalities

$$\begin{aligned}\widehat{N}\big((y, F(y)); \operatorname{gph} F\big) &= \widehat{N}\big((y, F(y)); \operatorname{gph} (T + sI)^{-1} \cap (\mathbb{R}^n \times B_\gamma(\bar{x})\big)\\ &= \widehat{N}\big((y, F(y)); \operatorname{gph} (T + sI)^{-1}\big),\end{aligned}$$

which clearly imply in turn the relationship

$$\widehat{D}^* F(y)(z) = \widehat{D}^*(T + sI)^{-1}\big(y, F(y)\big)(z).$$

Arguing similarly to (5.7), for any $w \in \widehat{\partial}\varphi_z(y) \subset \widehat{D}^* F(y)(z)$, we get from the latter equality that $-z + sw \in \widehat{D}^* T(F(y), y - sF(y))(-w)$. It follows from (5.2) that $\langle -z + sw, -w\rangle \geq 0$, which yields

$$\|z\| \cdot \|w\| \geq \langle z, w\rangle \geq s\|w\|^2, \quad \text{i.e.,} \quad \|z\| \geq s\|w\|.$$

This together with (5.5) tells us that

$$\langle z, u_1 - u_2\rangle \leq \frac{1}{s}\|y_1 - y_2\| \cdot \|z\|.$$

Since the obtained estimate holds for any $z \in \mathbb{R}^n$, we have

$$\|u_1 - u_2\|^2 \leq \frac{1}{s^2}\|y_1 - y_2\| = \frac{1}{s^2}\|v_1 + su_1 - v_2 - su_2\|^2 = \frac{1}{s^2}\|(v_1 - v_2) + s(u_1 - u_2)\|^2$$

and hence arrive at the inequality

$$0 \leq \frac{1}{s}\|v_1 - v_2\|^2 + 2\langle v_1 - v_2, u_1 - u_2\rangle \quad \text{when} \quad s > M.$$

Passing there to the limit as $s \to \infty$ shows that

$$0 \leq \langle v_1 - v_2, u_1 - u_2\rangle \quad \text{for all} \quad (u_1, v_1), (u_2, v_2) \in \operatorname{gph} T \cap \big(\operatorname{int} B_\gamma(\bar{x}) \times \mathbb{R}^n)\big),$$

which verifies the semilocal monotonicity of T at any $\bar{x} \in \operatorname{dom} T$. Since the domain of T is assumed to be convex, Lemma 5.3 tells us that T is globally monotone. Now we are in a position to apply Theorem 5.2 and conclude therefore that T is a globally maximal monotone operator on $\mathbb{R}^n$. △

It is well known in monotone operator theory that the maximal monotonicity of T always yields the *convexity* of the *closure* of the domain $\operatorname{cl}(\operatorname{dom} T)$; see, e.g., [70, Corollary 21.12]. This naturally gives a raise to the question whether Theorem 5.4 is true when the condition on the convexity of $\operatorname{dom} T$ is replaced by the convexity of $\operatorname{cl}(\operatorname{dom} T)$. The following simple example shows that it is not true, and consequently that the convexity assumption on $\operatorname{dom} T$ in Theorem 5.4 cannot be dropped.

Example 5.5 (Semilocal Monotonicity Doesn't Yield the Convexity of the Domain). Define the mapping $T : \mathbb{R} \rightrightarrows \mathbb{R}$ by

$$T(x) := \begin{cases} -x^{-1} & \text{if } \; x \in \mathbb{R}\backslash\{0\}, \\ \emptyset & \text{if } \; x = 0. \end{cases}$$

Observe that the operator T is semilocally monotone on $\mathbb{R}$, its graph $\operatorname{gph} T$ is closed, its domain $\operatorname{dom} T = \mathbb{R}\backslash\{0\}$ is nonconvex, while the closure of the domain $\operatorname{cl}(\operatorname{dom} T) = \mathbb{R}$ is convex (Fig. 5.1). Moreover, it is obvious that all the conditions in (ii) of Theorem 5.4 hold, but T is not globally monotone on $\mathbb{R}$.

5.1.3 Maximal Monotonicity via Limiting Coderivative

The next theorem provides other coderivative characterizations of global maximal monotonicity, where the regular coderivative condition (5.2) is replaced by the positive-semidefiniteness imposed on our *basic/limiting coderivative* (1.15). These characterizations are clearly equivalent to those presented in Theorems 5.2 and 5.4, but it is more convenient here to derive them by passing to the limit in (5.2). Note that the limiting coderivative characterizations have a *strong advantage* in comparison with (5.2) due to *comprehensive calculus rules* for (1.15) presented in Sections 3.2 and 3.4, which are not available for its regular (precoderivative) counterpart (1.16).

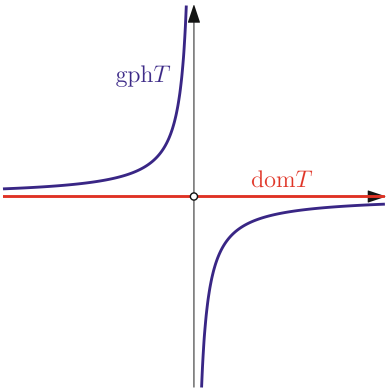


Fig. 5.1 Semilocal monotonicity but not global monotonicity.

Theorem 5.6 (Limiting Coderivative Characterizations of Global Maximal Monotonicity). *Let $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be a set-valued mapping with closed graph. The following assertions are equivalent:*

(i) *T is globally maximal monotone on $\mathbb{R}^n$.*

(ii) *T is globally hypomonotone on $\mathbb{R}^n$, and for any $(u, v) \in \operatorname{gph} T$ we have*

$$\langle z, w\rangle \geq 0 \ \ whenever \ \ z \in D^*T(u,v)(w), \ \ w \in \mathbb{R}^n. \tag{5.11}$$

If in addition the operator domain dom T *is convex, then the global hypomonotonicity in assertion* (ii) *can be equivalently replaced by the semilocal one.*

Proof. Implication (ii)⇒(i) is straightforward from Theorem 5.2 due to

$$\widehat{D}^*T(u,v)(w) \subset D^*T(u,v)(w) \ \text{ for all } \ (u,v) \in \text{gph}\, T \ \text{ and } \ w \in \mathbb{R}^n.$$

Thus (5.2) follows from (5.11), and T is maximal monotone by Theorem 5.2.

To justify the converse implication (i)⇒(ii), suppose that (i) holds, and so (5.2) is valid due to Theorem 5.2. Picking any vectors $(u,v) \in \text{gph}\, T$ and $z \in D^*T(u,v)(w)$ and using definition (1.15) of the basic coderivative, we find sequences $(u_k, v_k) \overset{\text{gph}\, T}{\to} (u,v)$ with $z_k \to z$ and $w_k \to w$ satisfying the inclusion $z_k \in \widehat{D}^*T(u_k,v_k)(w_k)$ for all $k \in \mathbb{N}$. It follows from (5.2) that $\langle z_k, w_k\rangle \geq 0$. Letting $k \to \infty$ implies that $\langle z, w\rangle \geq 0$, which verifies (5.11).

Assuming now the convexity of the domain dom T and employing Lemma 5.3 allow us to replace the global hypomonotonicity in (ii) by the semilocal hypomonotonicity while using Theorem 5.4 instead of Theorem 5.2. △

Remark 5.7 (Preservation of Maximal Monotonicity). Well-developed coderivative calculus presented in Sections 3.2 and 3.4 opens the gate to derive via (5.11) verifiable conditions ensuring the preservation of maximal monotonicity under various operations performed over maximal monotone operators. The results in this direction involve qualification conditions for the validity of the corresponding coderivative calculus rules.

The following one-dimensional example shows that the hypomonotonicity conditions in (ii) of Theorems 5.2, 5.4, and 5.6 are essential for the obtained coderivative characterizations of maximal monotonicity.

Example 5.8 (Hypomonotonicity Conditions Are Essential). Given positive η, define the set-valued mapping $T\colon \mathbb{R} \rightrightarrows \mathbb{R}$ with full domain by

$$T(x) := \eta x + [0,1] \ \text{ for all } \ x \in \mathbb{R}.$$

It is easy to calculate directly by the definitions that

$$D^*T(u,v)(w) = \widehat{D}^*T(u,v)(w) = \begin{cases} \{0\} & \text{if } \ w = 0, \ \ v - \eta u \in (0,1), \\ \{\eta w\} & \text{if } \ w \geq 0, \ \ v - \eta u = 0, \\ \{\eta w\} & \text{if } \ w \leq 0, \ \ v - \eta u = 1, \\ \emptyset & \text{otherwise.} \end{cases}$$

Thus both coderivative conditions (5.2) and (5.11) are satisfied. However, T is not globally monotone on $\mathbb{R}$ (Fig. 5.2). The reason is that this mapping is not semilocally (and hence not globally) hypomonotone.

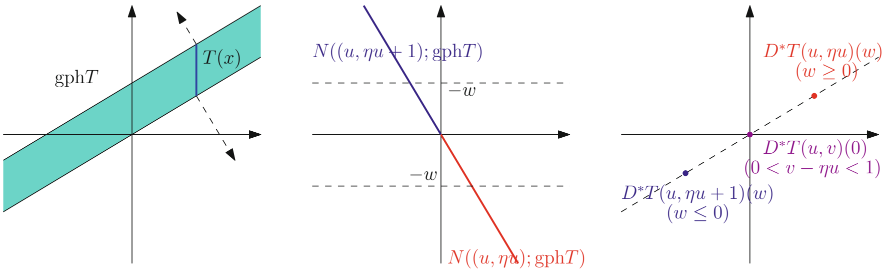


Fig. 5.2 Coderivative of $T(x) = \eta x + [0, 1]$ with $\eta > 0$.

Finally in this section, we derive from the obtained results complete coderivative characterizations of a stronger version of global monotonicity for set-valued mappings. We say that $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is *strongly globally maximal monotone* on $\mathbb{R}^n$ with modulus $\kappa > 0$ if it is globally maximal monotone and the shifted mapping $T - \kappa I$ is globally monotone on $\mathbb{R}^n$, i.e.,

$$\langle v_1 - v_2, u_1 - u_2\rangle \geq \kappa\|u_1 - u_2\|^2 \;\text{ for all }\; (u_1, v_1), (u_2, v_2) \in \text{gph}\, T.$$

Minty's theorem ensures that T is strongly globally maximal monotone on $\mathbb{R}^n$ with $\kappa > 0$ if and only if $T - \kappa I$ is globally maximal monotone on $\mathbb{R}^n$.

Corollary 5.9 (Coderivative Characterizations of Strong Global Maximal Monotonicity). *Let $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be a set-valued mapping with closed graph. The following are equivalent:*

(i) *T is strongly globally maximal monotone on $\mathbb{R}^n$ with modulus $\kappa > 0$.*

(ii) *T is globally hypomonotone on $\mathbb{R}^n$, and for any $(u, v) \in \text{gph}\, T$ we have*

$$\langle z, w\rangle \geq \kappa\|w\|^2 \;\textit{ whenever }\; z \in \widehat{D}^*T(u, v)(w),\; w \in \mathbb{R}^n.$$

(iii) *T is globally hypomonotone on $\mathbb{R}^n$, and we have*

$$\langle z, w\rangle \geq \kappa\|w\|^2 \;\textit{ whenever }\; z \in D^*T(u, v)(w),\; w \in \mathbb{R}^n$$

for any $(u, v) \in \text{gph}\, T$. If in addition the operator domain dom T *is convex, then the global hypomonotonicity in assertions* (ii) *and* (iii) *can be equivalently replaced by its semilocal counterpart.*

Proof. Define $S := T - \kappa I$, and immediately deduce from the coderivative sum rules in Exercise 3.59(i,ii) the following equalities

$$\begin{aligned}\widehat{D}^*T(u, v)(w) &= \widehat{D}^*S(u, v - \kappa u)(w) + \kappa w,\\ D^*T(u, v)(w) &= D^*S(u, v - \kappa u)(w) + \kappa w\end{aligned}$$

holding for all $(u, v) \in \text{gph}\, T$ and $w \in \mathbb{R}^n$. Thus the validity of (ii) (resp. (iii)) for T is equivalent to the fulfillment of all the conditions in Theorem 5.2(ii) (resp. in

Theorem 5.4(ii)) for the operator S; it is obvious for hypomonotonicity. Applying now Theorem 5.2 and Theorem 5.4, respectively, we get that either assertion (ii) or (iii) of this corollary is equivalent to the global maximal monotonicity of S. Since the latter is equivalent to the global strong maximal monotonicity of T with modulus κ, we complete the proof. △

5.2 Coderivative Criteria for Strong Local Monotonicity

In this section we study *strong local* monotonicity properties of set-valued operators and provide their complete *coderivative characterizations* of their *maximality*. Similarly to the global monotonicity investigated in Section 5.1, the techniques developed here also use a variational approach and generalized differentiation while being largely different from and more involved in comparison with global maximal monotonicity. We essentially exploit now the strong local nature of the maximal monotonicity under consideration.

5.2.1 Strong Local Monotonicity and Related Properties

The following local monotonicity properties of operators are studied below. Recall that we have already used local monotonicity in Subsection 3.3.3.

Definition 5.10 (Locally Monotone and Strongly Monotone Operators). *Let* $T\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *with* $(\bar{x}, \bar{v}) \in \operatorname{gph} T$. *We say that:*

(i) *T is* LOCALLY MONOTONE *around* $(\bar{x}, \bar{v})$ *if there exists a neighborhood* $U \times V$ *of this point such that*

$$\langle v_1 - v_2, u_1 - u_2 \rangle \geq 0 \ \textit{ for all } \ (u_1, v_1), (u_2, v_2) \in \operatorname{gph} T \cap (U \times V).$$

(ii) *T is* STRONGLY LOCALLY MONOTONE *around* $(\bar{x}, \bar{v})$ *with modulus* $\kappa > 0$ *if there exists a neighborhood* $U \times V$ *of* $(\bar{x}, \bar{v})$ *such that for any pair* $(u_1, v_1), (u_2, v_2) \in \operatorname{gph} T \cap (U \times V)$ *we have the estimate*

$$\langle v_1 - v_2, u_1 - u_2 \rangle \geq \kappa \|u_1 - u_2\|^2. \tag{5.12}$$

(iii) *T is* STRONGLY LOCALLY MAXIMAL MONOTONE *around* $(\bar{x}, \bar{v})$ *with modulus* $\kappa > 0$ *if there exists a neighborhood* $U \times V$ *of* $(\bar{x}, \bar{v})$ *such that* (ii) *holds and that* $\operatorname{gph} T \cap (U \times V) = \operatorname{gph} S \cap (U \times V)$ *for any globally monotone operator* $S\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ *satisfying the inclusion* $\operatorname{gph} T \cap (U \times V) \subset \operatorname{gph} S$.

In what follows we present complete coderivative characterizations of the *strong local maximal monotonicity* of set-valued operators while connecting this property, via coderivatives, with *local hypomonotonicity* from Definition 5.1(iii). To proceed, consider first the notions of *single-valued localizations* of set-valued mappings that are important for the study and applications of strong local monotonicity and related properties.

Definition 5.11 (Single-Valued Localizations). *Given $F\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ with $(\bar{x}, \bar{y}) \in \operatorname{gph} F$, we say that F admits a* SINGLE-VALUED LOCALIZATION *around $(\bar{x}, \bar{y})$ if there is a neighborhood $U \times V \subset \mathbb{R}^n \times \mathbb{R}^m$ of $(\bar{x}, \bar{y})$ such that the mapping $\widehat{F}\colon U \to V$ defined via $\operatorname{gph} \widehat{F} := \operatorname{gph} F \cap (U \times V)$ is single-valued on U with $\operatorname{dom} \widehat{F} = U$. Furthermore, F admits a* LIPSCHITZIAN SINGLE-VALUED LOCALIZATION *around $(\bar{x}, \bar{y})$ if the mapping $\widehat{F}$ is Lipschitz continuous on U.*

If the mapping $\widehat{F}$ in Definition 5.11 is generally set-valued, it is said to be just a *localization* of F relatively to $U \times V$.

Using the second part of Definition 5.11, we formulate now the following well-posedness property related to our study in Section 3.1. Theorem 3.2(ii) tells us that this property can be viewed as a Lipschitzian single-valued localization of metric regularity for the mapping in question.

Definition 5.12 (Strong Metric Regularity). *We say that $F\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is* STRONGLY METRICALLY REGULAR *around $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ with modulus $\ell > 0$ if the inverse mapping F^{-1} admits a single-valued localization around $(\bar{y}, \bar{x})$, which is Lipschitz continuous around $\bar{y}$ with modulus ℓ.*

The next result characterizes the strong local maximal monotonicity of T via Lipschitzian single-valued localizations of the inverse T^{-1} that indeed distinguishes strong local *maximal* monotonicity from merely strong local monotonicity. In addition to qualitative characterizations, the theorem below provides some quantitative relationships between the corresponding moduli.

Theorem 5.13 (Strong Local Maximal Monotonicity via Lipschitzian Localization). *Given a set-valued operator $T\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ with $(\bar{x}, \bar{v}) \in \operatorname{gph} T$ and given $\kappa > 0$, the following assertions are equivalent:*

(i) *T is strongly locally maximal monotone around $(\bar{x}, \bar{v})$ with modulus κ.*

(ii) *T is strongly locally monotone around $(\bar{x}, \bar{v})$ with modulus κ, and its inverse T^{-1} admits a Lipschitzian single-valued localization around $(\bar{v}, \bar{x})$.*

(iii) *T^{-1} admits a single-valued localization ϑ relative to some neighborhood $V \times U$ of $(\bar{v}, \bar{x})$ such that for all $v_1, v_2 \in V$ we have*

$$\big\|(v_1 - v_2) - 2\kappa\big[\vartheta(v_1) - \vartheta(v_2)\big]\big\| \le \|v_1 - v_2\|, \tag{5.13}$$

which implies that ϑ is locally Lipschitzian around $(\bar{v}, \bar{x})$ with modulus κ^{-1}, and so T is strongly metrically regular around $(\bar{x}, \bar{v})$ with the same modulus.

Proof. To verify (i)$\Rightarrow$(ii), take by (i) a neighborhood $U \times V$ of $(\bar{x}, \bar{v})$ such that (5.12) holds and we have $\operatorname{gph} T \cap (U \times V) = \operatorname{gph} S \cap (U \times V)$ for any globally monotone operator $S\colon \mathbb{R}^n \times \mathbb{R}^n$ with $\operatorname{gph} T \cap (U \times V) \subset \operatorname{gph} S$. Define

$$J_\kappa(u, v) := (u, v - \kappa u) \ \text{ on } \ \mathbb{R}^n \times \mathbb{R}^n, \quad W := J_\kappa(U \times V)$$

and deduce from (5.12) that the operator $F\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ constructed by $\operatorname{gph} F := \operatorname{gph}(T - \kappa I) \cap W$ via the identity mapping I is globally monotone on $\mathbb{R}^n$. Indeed, whenever $(u_i, v_i) \in \operatorname{gph} F$ we get

$$(u_i, v_i + \kappa u_i) \in \operatorname{gph} T \cap J_\kappa^{-1}(W) = \operatorname{gph} T \cap (U \times V) \ \text{ for } \ i = 1, 2.$$

It follows from the strong local monotonicity (5.12) of T that

$$\langle v_1 + \kappa u_1 - v_2 - \kappa u_2, u_1 - u_2\rangle \geq \kappa \|u_1 - u_2\|^2,$$

which yields $\langle v_1 - v_2, u_1 - u_2\rangle \geq 0$ and thus verifies the global monotonicity of F. Consider now the (global) *maximal monotone extension* R of F (see, e.g., [70, Theorem 20.21]) for which we have the inclusion

$$\operatorname{gph}(F + \kappa I) \cap (U \times V) = \operatorname{gph} T \cap (U \times V) \subset \operatorname{gph}(R + \kappa I).$$

The local maximal monotonicity of T relative to the neighborhood $U \times V$ implies that $\operatorname{gph} T \cap (U \times V) = \operatorname{gph}(R + \kappa I) \cap (U \times V)$, and therefore

$$\operatorname{gph} T^{-1} \cap (V \times U) = \operatorname{gph}(R + \kappa I)^{-1} \cap (V \times U). \tag{5.14}$$

The aforementioned Minty theorem tells us that $\operatorname{dom}(R + \kappa I)^{-1} = \mathbb{R}^n$ and that the operator $(R + \kappa I)^{-1}$ is single-valued and Lipschitz continuous on $\mathbb{R}^n$. Combining this with (5.14) ensures that the set

$$V_1 := (R + \kappa I)(U) \cap V = \big[(R + \kappa I)^{-1}\big]^{-1}(U) \cap V$$

is a neighborhood of $\bar{v}$ by noting from (5.14) that $(\bar{v}, \bar{x}) \in \operatorname{gph}(R + \kappa I)^{-1} \cap (V \times U)$ and using the fact that V_1 is the inverse image of the neighborhood U via the continuous map $(R + \kappa I)^{-1}$. Furthermore, it follows from (5.14) that $T^{-1}(v) = (R + \kappa I)^{-1}(v)$ for all $v \in V_1$. Thus the localization $S\colon V_1 \to U$ defined via $\operatorname{gph} S := \operatorname{gph} T^{-1} \cap (V_1 \times U)$ is single-valued and Lipschitz continuous on V_1. This verifies implication (i)$\Rightarrow$(ii).

To justify (ii)$\Rightarrow$(iii), find by (ii) a neighborhood $U \times V$ of $(\bar{x}, \bar{v})$ on which (5.12) holds and such that the localization ϑ of T^{-1} is single-valued and Lipschitz continuous on $V \times U$. Then it follows from (5.12) that

$$\begin{aligned}\|v_1 - v_2 - 2\kappa(u_1 - u_2)\|^2 &= \|v_1 - v_2\|^2 - 4\kappa\big[\langle v_1 - v_2, u_1 - u_2\rangle - \kappa\|u_1 - u_2\|^2\big]\\ &\leq \|v_1 - v_2\|^2 \ \text{ if } \ (v_1, u_1), (v_2, u_2) \in \operatorname{gph} \cap (V \times U),\end{aligned}$$

which yields (5.13) and thus verifies the main statement in (iii). To show further that (5.13) readily implies that ϑ is locally Lipschitz continuous around $(\bar{v}, \bar{x})$ with modulus κ^{-1} (and hence T is strongly metrically regular around $(\bar{x}, \bar{v})$ with the same modulus according to Definition 5.12), take $u_i := \vartheta(v_i)$, $i = 1, 2$, and deduce from (5.13) that

$$0 \le \|v_1 - v_2\|^2 - \|v_1 - v_2 - 2\kappa(u_1 - u_2)\|^2$$
$$= 4\kappa\big[\langle v_1 - v_2, u_1 - u_2\rangle - \kappa\|u_1 - u_2\|^2\big].$$

This implies in turn the estimates

$$\|v_1 - v_2\| \cdot \|u_1 - u_2\| \ge \langle v_1 - v_2, u_1 - u_2\rangle \ge \kappa\|u_1 - u_2\|^2, \tag{5.15}$$

which therefore justify the additional claims in (iii).

It remains to verify implication (iii)⇒(i). Having by (iii) the neighborhood $V \times U$ of $(\bar{v}, \bar{x})$ on which T^{-1} admits the single-valued localization ϑ and (5.12) holds, pick any $(u_1, v_1), (u_2, v_2) \in \text{gph}\, T \cap (U \times V)$ for which we get $u_i = \vartheta(v_i), i = 1, 2$, and the strong local monotonicity condition (5.12) is satisfied as proved in (5.15). Let us finally check the local *maximality* of T.

To proceed, take any globally monotone operator $S\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ satisfying the inclusion $\text{gph}\, T \cap (U \times V) \subset \text{gph}\, S$, and conclude by (5.13) that

$$\langle y - v, \vartheta(y) - u\rangle \ge 0 \ \text{ for any } \ y \in V,\ (u, v) \in \text{gph}\, S \cap (U \times V). \tag{5.16}$$

Fix an arbitrary vector $z \in \mathbb{R}^n$, and find $\varepsilon > 0$ such that $v + \varepsilon z \in V$. Since $\vartheta(V) \subset U$, we have $\vartheta(v + \varepsilon z) \in U$. This tells us together with (5.16) that

$$\big\langle v + \varepsilon z - v, \vartheta(v + \varepsilon z) - u\big\rangle = \varepsilon\big\langle z, \vartheta(v + \varepsilon z) - u\big\rangle \ge 0,$$

which clearly yields $\langle z, \vartheta(v+\varepsilon z)-u\rangle \ge 0$. Letting now $\varepsilon \downarrow 0$ implies that $\langle z, \vartheta(v)-u\rangle \ge 0$ due to the continuity of ϑ shown above. Since this holds for any $z \in \mathbb{R}^n$, we get $\vartheta(v) = u$, i.e., $(u, v) \in \text{gph}\, T \cap (U \times V)$. Therefore

$$\text{gph}\, S \cap (U \times V) \subset \text{gph}\, T \cap (U \times V),$$

which verifies the strong local maximal monotonicity of T relative to $U \times V$ and thus completes the proof of the theorem. △

5.2.2 Strong Local Maximal Monotonicity via Coderivatives

The next theorem presents the principal result of this section on characterizing strongly locally maximal monotone (closed-graph) operators via their local hypomonotonicity coupled with a *strengthened positive-definiteness condition* expressed in terms of the precoderivative/regular coderivative (1.16) at neighborhood points. The result obtained provides also a quantitative relation involving *moduli* of strong local maximal monotonicity.

Theorem 5.14 (Neighborhood Coderivative Characterization of Strong Local Maximal Monotonicity). *Given a set-valued mapping $T\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ with $(\bar{x}, \bar{v}) \in \text{gph}\, T$, fix a number $\kappa > 0$. The following are equivalent:*

(i) *T is strongly locally maximal monotone around $(\bar{x}, \bar{v})$ with modulus κ.*

(ii) *T is locally hypomonotone around $(\bar{x}, \bar{v})$, and there is $\eta > 0$ such that*

$$\langle z, w \rangle \geq \kappa \|w\|^2 \quad \text{if} \quad z \in \widehat{D}^* T(u, v)(w), \ (u, v) \in \operatorname{gph} T \cap B_\eta(\bar{x}, \bar{v}). \quad (5.17)$$

Proof. To justify (i)⇒(ii), note that the local hypomonotonicity is trivial under (i), and by Theorem 5.13 find a single-valued localization ϑ of T^{-1} relative to a neighborhood $V \times U$ of $(\bar{v}, \bar{x})$ such that (5.13) holds. As proved above, this yields the Lipschitz continuity of ϑ on V with modulus κ^{-1}. To proceed with verifying the coderivative condition (5.17), choose $\eta > 0$ satisfying $B_\eta(\bar{x}, \bar{v}) \subset U \times V$ and then pick $(u, v) \in \operatorname{gph} T \cap B_\eta(\bar{x}, \bar{v})$ and $z \in \widehat{D}^* T(u, v)(w)$. Given any $\varepsilon > 0$ and using (1.16), we select η so small that

$$\langle z, x - u \rangle - \langle w, y - v \rangle \leq \varepsilon \big(\|x - u\| + \|y - v\| \big) \quad (5.18)$$

for all $(x, y) \in \operatorname{gph} T \cap B_\eta(u, v)$. When $t > 0$ is also small, consider $u_t := \vartheta(v_t)$ with $v_t := v + t(z - 2\kappa w) \in V$, and get from the continuity of ϑ that $(u_t, v_t) \to (u, v)$ as $t \downarrow 0$. Suppose without loss of generality that $(u_t, v_t) \in B_\eta(u, v)$ for all $t > 0$. Replacing (x, y) in (5.18) by (u_t, v_t) and using (5.13) give us

$$\begin{aligned}
\varepsilon \big(\|u_t - u\| + \|v_t - v\| \big) &\geq \langle z, u_t - u \rangle - \langle w, v_t - v \rangle \\
&= \langle t^{-1}(v_t - v) + 2\kappa w, u_t - u \rangle - t \langle w, z - 2\kappa w \rangle \\
&\geq \kappa t^{-1} \|u_t - u\|^2 + 2\kappa \langle w, u_t - u \rangle - t \langle w, z - 2\kappa w \rangle \\
&\geq \kappa t^{-1} \|u_t - u\|^2 - 2\kappa \|w\| \cdot \|u_t - u\| \\
&\quad + t\kappa \|w\|^2 - t \langle w, z - \kappa w \rangle \\
&\geq -t \langle w, z - \kappa w \rangle = -t \langle z, w \rangle + t\kappa \|w\|^2.
\end{aligned}$$

Since ϑ is Lipschitz continuous on V with modulus κ^{-1}, we have

$$\begin{aligned}
\varepsilon \big(\|u_t - u\| + \|v_t - v\| \big) &= \varepsilon \big(\|\vartheta(v_t) - \vartheta(v)\| + \|v_t - v\| \big) \\
&\leq \varepsilon \big(\kappa^{-1} \|v_t - v\| + \|v_t - v\| \big) \\
&= \varepsilon (\kappa^{-1} + 1) \|v_t - v\| = \varepsilon t (\kappa^{-1} + 1) \|z - 2\kappa w\|,
\end{aligned}$$

which together with the estimates above yields

$$\langle z, w \rangle + \varepsilon (\kappa^{-1} + 1) \|z - 2\kappa w\| \geq \kappa \|w\|^2.$$

Passing to the limit as $\varepsilon \downarrow 0$ gives us $\langle z, w \rangle \geq \kappa \|w\|^2$ and thus justifies (5.17).

To verify next the converse implication (ii)⇒(i), observe that by Theorem 5.13 we only need to show that the inverse operator T^{-1} admits a Lipschitz continuous single-valued localization ϑ around $(\bar{v}, \bar{x})$ satisfying estimate (5.13). This is done in the following two claims.

Claim 1. *T^{-1} admits a Lipschitz continuous localization ϑ around $(\bar{v}, \bar{x})$.*

To justify this claim, choose $\eta > 0$ so small that the set $\operatorname{gph} T \cap B_\eta(\bar{x}, \bar{v})$ is closed and there is a positive number r for which

$$\langle v_1 - v_2, x_1 - x_2 \rangle \geq -r \|x_1 - x_2\|^2 \quad (5.19)$$

if $(x_1, v_1), (x_2, v_2) \in \operatorname{gph} T \cap B_\eta(\bar{x}, \bar{v})$. Pick any $s > r$ and define

$$J_s(u, v) := (v + su, u) \ \text{ for } \ (u, v) \in \mathbb{R}^n \times \mathbb{R}^n.$$

Denoting further $W_s := J_s(B_\eta(\bar{x}, \bar{v}))$, observe that $\operatorname{int} W_s = J_s(\operatorname{int} B_\eta(\bar{x}, \bar{v}))$ is a neighborhood of $(\bar{v} + s\bar{x}, \bar{x})$. It follows from (5.19) that for any pair $(v_1, x_1), (v_2, x_2) \in \operatorname{gph}(T + sI)^{-1} \cap W_s$, we have $(x_i, v_i - sx_i) \in \operatorname{gph} T \cap J_s^{-1}(W_s) = \operatorname{gph} T \cap \mathbb{B}_\eta(\bar{x}, \bar{v})$. Thus (5.19) tells us that

$$\langle v_1 - sx_1 - v_2 + sx_2, x_1 - x_2 \rangle \ge -r\|x_1 - x_2\|^2,$$

which clearly implies the estimates

$$\|v_1 - v_2\| \cdot \|x_1 - x_2\| \ge \langle v_1 - v_2, x_1 - x_2 \rangle \ge (s - r)\|x_1 - x_2\|^2 \tag{5.20}$$

showing that the mapping $(T + sI)^{-1}$ admits a single-valued localization denoted by f. Taking now any $(v, u) \in \operatorname{gph} f \cap (\operatorname{int} W_s)$ with $u = f(v)$ and any $(w, z) \in X \times X$ with $w \in \widehat{D}^* f(v)(z)$, we get that $w \in \widehat{D}^*(T + sI)^{-1}(v, u)(z)$ and hence $-z \in \widehat{D}^*(T + sI)(u, v)(-w)$. It follows from the equality sum rule for the regular coderivative in Exercise 3.59(i) that $-z + sw \in \widehat{D}^* T(u, v - su)(-w)$. Since $(u, v - su) = J_s^{-1}(v, u) \in J_s^{-1}(\operatorname{int} W_s) = \operatorname{int} B_\eta(\bar{x}, \bar{v})$, we deduce from (5.17) that $\langle -z + sw, -w \rangle \ge \kappa\|w\|^2$ and thus

$$\|z\| \cdot \|w\| \ge \langle z, w \rangle \ge (\kappa + s)\|w\|^2.$$

To proceed further, for any $z \in \mathbb{B}$ define the function $\varphi_z \colon \mathbb{R}^n \to \overline{\mathbb{R}}$ by

$$\varphi_z(v) := \begin{cases} \langle z, f(v) \rangle & \text{if} \quad v \in \operatorname{dom} f, \\ \infty & \text{otherwise} \end{cases}$$

and verify similarly to the proof of Theorem 5.2 that it is l.s.c. on $\mathbb{R}^n$. Applying the mean value inequality (4.25) to φ_z, fix $\gamma \in (0, \eta/3)$ and pick two pairs $(u_i, v_i) \in \operatorname{gph} T \cap B_\gamma(\bar{x}, \bar{v})$, $i = 1, 2$. By the construction of f, we get $(y_i, u_i) \in \operatorname{gph} f$ with $y_i := v_i + su_i$. Taking any $\varepsilon \in (0, \gamma)$ and applying the increment estimate (4.25) to φ_z on $[y_1, y_2]$ with the chosen ε give us

$$|\varphi_z(y_1) - \varphi_z(y_2)| \le \|y_1 - y_2\| \sup\big\{\|w\| \ \big| \ w \in \widehat{\partial}\langle z, f \rangle(y), \ \ y \in [y_1, y_2] + \varepsilon\mathbb{B}\big\}.$$

For any $y \in \operatorname{dom} f \cap \big([y_1, y_2] + \varepsilon\mathbb{B}\big)$, there are some $t \in [0, 1]$ and $y_0 \in \varepsilon\mathbb{B}$ such that $y = ty_1 + (1 - t)y_2 + y_0$. Then it follows that

$$\begin{aligned}
\|y - \bar{v} - s\bar{x}\| &= \|ty_1 + (1 - t)y_2 + y_0 - \bar{v} - s\bar{x}\| \\
&= \|t(y_1 - \bar{v} - s\bar{x}) + (1 - t)(y_2 - \bar{v} - s\bar{x}) + y_0\| \\
&= \|t(v_1 + su_1 - \bar{v} - s\bar{x}) + (1 - t)(v_2 + su_2 - \bar{v} - s\bar{x}) + y_0\| \\
&\le t\big(\|v_1 - \bar{v}\| + s\|u_1 - \bar{x}\|\big) + (1 - t)\big(\|v_2 - \bar{v}\| + s\|u_2 - \bar{x}\|\big) + \|y_0\| \\
&\le t(\gamma + s\gamma) + (1 - t)(\gamma + s\gamma) + \varepsilon = (1 + s)\gamma + \varepsilon < (2 + s)\gamma.
\end{aligned}$$

We easily get from the latter estimate and (5.20) that

$$\begin{aligned}\|f(y)-\bar{x}\| &= \|f(y)-f(\bar{v}+s\bar{x})\| \le (s-r)^{-1}\|y-\bar{v}-s\bar{x}\|\\ &\le (s-r)^{-1}(2+s)\gamma.\end{aligned} \tag{5.21}$$

Furthermore, it follows from the above that

$$\begin{aligned}\|y-sf(y)-\bar{v}\| &= \big\|y-\bar{v}-s\bar{x}-s\big(f(y)-\bar{x}\big)\big\|\\ &\le (2+s)\gamma+s(s-r)^{-1}(2+s)\gamma.\end{aligned} \tag{5.22}$$

By choosing γ sufficiently small, we deduce from (5.21) and (5.22) that

$$J_s^{-1}\big(y,f(y)\big)=\big(f(y),y-sf(y)\big)\in \operatorname{int} B_\eta(\bar{x},\bar{v}),$$

which tells us that $(y,f(y))\in J_s\big(\operatorname{int} B_\eta(\bar{x},\bar{v})\big)=\operatorname{int} W_s$. Moreover, it is easy to see from the definitions that $\widehat{\partial}\langle z,f\rangle(y)\subset \widehat{D}^*f(y)(z)$. Taking into account that $(y,f(y))\in \operatorname{gph}(T+sI)^{-1}\cap \operatorname{int} W_s$ and the constructions of f and φ_z, we conclude that $\widehat{D}^*f(y)(z)=\widehat{D}^*(T+sI)^{-1}(y,f(y))(z)$, which ensures by the increment estimate above that

$$|\langle z,f(y_1)-f(y_2)\rangle|=|\varphi_z(y_1)-\varphi_z(y_2)|\le \|y_1-y_2\|(\kappa+s)^{-1}\|z\|$$

for all $z\in \mathbb{B}$. Remembering the definitions of y_i above implies that

$$\begin{aligned}\|u_1-u_2\| &= \|f(y_1)-f(y_2)\|\\ &\le (\kappa+s)^{-1}\|y_1-y_2\|=(\kappa+s)^{-1}\|v_1+su_1-v_2-su_2\|,\end{aligned}$$

which yields in turn the inequality

$$(\kappa+s)\|u_1-u_2\|\le \|(v_1-v_2)+s(u_1-u_2)\|\le \|v_1-v_2\|+s\|u_1-u_2\|.$$

Thus we arrive at the estimate

$$\kappa\|u_1-u_2\|\le \|v_1-v_2\| \ \text{ if } \ (u_1,v_1),\ (u_2,v_2)\in \operatorname{gph} T\cap B_\gamma(\bar{x},\bar{v}). \tag{5.23}$$

It remains to show that T^{-1} admits a single-valued Lipschitzian localization around $(\bar{v},\bar{x})$. To verify it, observe from (5.17) that

$$\|z\|\ge \kappa\|w\| \ \text{ for all } \ z\in \widehat{D}^*T(u,v)(w),\ (u,v)\in \operatorname{gph} T\cap B_\eta(\bar{x},\bar{v}),$$

which is a neighborhood version of the coderivative characterization of *metric regularity* of T around $(\bar{x},\bar{v})$ in Theorem 3.3(ii) (see Exercise 3.47 and [522, Theorem 4.5] for more details). This allows us to find positive numbers μ and ν, where μ can be taken as κ^{-1}, such that

$$\operatorname{dist}\big(\bar{x};T^{-1}(v)\big)\le \mu\operatorname{dist}\big(v;T(\bar{x})\big)\le \mu\|v-\bar{v}\| \ \text{ for all } \ v\in B_\nu(\bar{v})$$

which ensures that $T^{-1}(v) \cap \text{int}\, B_{\mu\nu}(\bar{x}) \neq \emptyset$ for all $v \in \text{int}\, B_\nu(\bar{v})$. Defining finally the mapping ϑ from $\text{int}\, B_\nu(\bar{v})$ into $\text{int}\, B_{\mu\nu}(\bar{x})$ by

$$\text{gph}\,\vartheta := \text{gph}\, T^{-1} \cap \big(\text{int}\, B_\nu(\bar{v}) \times \text{int}\, B_{\mu\nu}(\bar{x})\big) \ \text{ with } \ \text{dom}\,\vartheta = \text{int}\, B_\nu(\bar{v}),$$

we get $\text{dom}\,\vartheta = \text{int}\, B_\nu(\bar{v})$. It follows directly from (5.23) that ϑ is single-valued and Lipschitz continuous on its domain with modulus κ^{-1}.

Claim 2. *The single-valued Lipschitzian localization ϑ of T^{-1} taken from Claim* 1 *satisfies the additional condition* (5.13).

To verify this claim, for any $z \in \mathbb{B}$ define $\xi_z(v) := \langle z, v - 2\kappa\vartheta(v)\rangle$ as $v \in \text{int}\, B_\nu(\bar{v})$. Pick any $\alpha, \varepsilon > 0$ with $\alpha + \varepsilon < \nu$ and any $v_1, v_2 \in B_\alpha(\bar{v})$. Similarly to the proof of the increment estimate for φ_z in Claim 1 with the chosen ε therein, we deduce from the mean value inequality that

$$|\xi_z(v_1) - \xi_z(v_2)| \leq \|v_1 - v_2\| \sup\big\{\|w\| \ \big|\ w \in \widehat{\partial}\xi_z(v),\ v \in [v_1, v_2] + \varepsilon\mathbb{B}\big\}.$$

Since $v \in \text{int}\, B_\nu(\bar{v})$ for each $v \in [v_1, v_2] + \varepsilon\mathbb{B}$, it is easy to get from the construction of ξ_z and the elementary calculus rules as above that

$$w \in \widehat{\partial}\xi_z(v) \subset z - 2\kappa\widehat{D}^*\vartheta(v)(z) = z - 2\kappa\widehat{D}^*T^{-1}(v)(z),$$

which tells us that $(2\kappa)^{-1}(z - w) \in \widehat{D}^*T^{-1}(v)(z)$ or equivalently

$$-z \in \widehat{D}^*T\big(\vartheta(v), v\big)\big((2\kappa)^{-1}(w - z)\big).$$

It thus follows from the coderivative condition (5.17) that

$$\langle -z, (2\kappa)^{-1}(w - z)\rangle \geq \kappa\|(2\kappa)^{-1}(w - z)\|^2,$$

which easily implies in turn that $\|w\| \leq \|z\|$. This together with the increment estimate for ξ_z established above ensures that

$$|\xi_z(v_1) - \xi_z(v_2)| = \langle z, v_1 - 2\kappa\vartheta(v_1) - v_2 + 2\kappa\vartheta(v_2)\rangle \leq \|v_1 - v_2\| \cdot \|z\| \ \text{ for all } \ z \in \mathbb{B}.$$

Remembering the definition of ξ_z, we derive from the latter that

$$\|v_1 - v_2 - 2\kappa[\vartheta(v_1) - \vartheta(v_2)]\| \leq \|v_1 - v_2\| \ \text{ whenever } \ v_1, v_2 \in B_\alpha(\bar{v}),$$

which verifies condition (5.13) and hence justifies Claim 2. The proof of the theorem is complete by combining Claim 1 and Claim 2. △

Let us present a remarkable consequence of Theorem 5.14 for detecting *strong metric regularity* of set-valued mappings.

Corollary 5.15 (Sufficient Conditions for Strong Metric Regularity). *The conditions of Theorem* 5.14(ii) *ensure that the mapping T is strongly metrically regular around* $(\bar{x}, \bar{v})$ *with modulus κ^{-1}.*

Proof. By Theorem 5.13 it follows from assertion (i) of Theorem 5.14, while it has also been deduced directly from hypomonotonicity and the coderivative condition (5.17) in the proof of the latter theorem. △

5.2.3 Pointbased Coderivative Characterizations

We conclude this section by deriving a *pointbased* characterization of strong maximal monotonicity for Lipschitz-like mappings via *positive-definiteness* of the basic/limiting *coderivative* (1.15). The next theorem and its corollary below are natural extensions to set-valued and nonsmooth mappings of the classical characterization of the strong local monotonicity for a smooth mapping via positive-definiteness of its Jacobian matrix.

Theorem 5.16 (Pointbased Coderivative Conditions for Strong Local Maximal Monotonicity). *Let $T\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$. The following assertions hold:*

(i) *The strong local maximal monotonicity of T around $(\bar{x}, \bar{v}) \in \operatorname{gph} T$ implies that the coderivative $D^*T(\bar{x}, \bar{v})$ is em positive-definite, i.e.,*

$$\langle z, w\rangle > 0 \ \textit{for any} \ z \in D^*T(\bar{x}, \bar{v})(w) \ \textit{with} \ w \neq 0. \tag{5.24}$$

(ii) *If T is single-valued and Lipschitz continuous around $\bar{x}$, then the positive-definiteness condition (5.24) is necessary and sufficient for the strong local maximal monotonicity of T around this point.*

Proof. It is easy to see by passing to the limit that (5.17) always implies (5.24), and thus we get (i) by employing Corollary 5.15. Let us now verify assertion (ii) assuming that T is single-valued and locally Lipschitzian around $\bar{x}$. First we check that T is automatically locally hypomonotone around $\bar{x}$ in this case. Indeed, take a Lipschitz constant $\ell > 0$ of T around $\bar{x}$, and define $T\colon \mathbb{R}^n \to \mathbb{R}^n$ by $g(u) := T(u) + \ell u$. Then we have

$$\begin{aligned}\langle g(u_1) - g(u_2), u_1 - u_2\rangle &= \langle T(u_1) - T(u_2), u_1 - u_2\rangle + \ell\|u_1 - u_2\|^2\\ &\geq -\|T(u_1) - T(u_2)\| \cdot \|u_1 - u_2\| + \ell\|u_1 - u_2\|^2 \geq 0,\end{aligned}$$

which therefore yields the local hypomonotonicity of T around $\bar{x}$. Invoking Theorem 5.14, it remains to show that (5.24) ensures the validity of (5.17). Arguing by contradiction, suppose that (5.24) holds while (5.17) doesn't. Hence we find a sequence (u_k, w_k, z_k) satisfying

$$u_k \to \bar{x}, \ \ z_k \in \widehat{D}^*T(u_k)(w_k), \ \ \langle z_k, w_k\rangle < k^{-1}\|w_k\|^2 \ \text{ for all } \ k \in \mathbb{N}.$$

Letting $\bar{w}_k := w_k/\|w_k\|$, $\bar{z}_k := z_k/\|w_k\|$ and using the Lipschitz property of T with modulus $\ell \geq 0$ yield as in the proof of Theorem 3.3 that

$$\|\bar{z}_k\| \leq \ell\|\bar{w}_k\| = \ell, \quad k \in \mathbb{N}.$$

Select convergent subsequences of $\{\bar{w}_k\}$ and $\{\bar{z}_k\}$ and then find $(\bar{w}, \bar{z})$ such that $(\bar{w}_k, \bar{z}_k) \to (\bar{w}, \bar{z})$. Passing now to the limit as $k \to \infty$ and using the limiting coderivative representation (1.17), we get $\bar{z} \in D^*T(\bar{x})(\bar{w})$ with $\|\bar{w}\| = 1$. Furthermore, it follows from $\langle \bar{z}_k, \bar{w}_k \rangle < k^{-1}$ that $\langle \bar{z}, \bar{w} \rangle \le 0$, which contradicts (5.24) and thus completes the proof of the theorem. △

Remark 5.17 (Pointbased Coderivative Criteria for Strong Local Maximal Monotonicity of Mappings). A natural question arises about the possibility to extend the pointbased coderivative characterization of Theorem 5.16(ii) to *set-valued* mappings. Observe first that the answer is *negative* if the mapping T in question is *Lipschitz-like* around $(\bar{x}, \bar{v})$, provided that we keep the local hypomonotonicity assumptions coming from Theorem 5.14. As shown by Levy and Poliquin [455], the simultaneous validity of the Lipschitz-like and local hypomonotonicity properties of a mapping around the reference point is *equivalent* to the *single-valuedness* and Lipschitz continuity of the mapping around this point.

On the other hand, it is shown by the author and Nghia [555] that the pointbased condition (5.24) completely characterizes, together with the local hypomonotonicity of T, the strong local maximal monotonicity of *subgradient* (highly non-Lipschitzian) mappings $T = \partial\varphi$ generated by extended-real-valued functions $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ that belong to a broad class of continuously *prox-regular* functions. As mentioned in Section 3.5, the latter class plays a crucial role in second-order variational analysis, optimization, and their numerous applications; see the books [527, 678] and the discussions above.

5.3 Exercises for Chapter 5

Exercise 5.18 (Hypomonotonicity of Single-Valued Mappings).

(i) Let $T\colon X \to X$ be a single-valued mapping on a Hilbert space X. Prove that T is locally hypomonotone around $\bar{x}$ provided that it is locally Lipschitzian around this point. *Hint:* Compare it with the proof of [455, Theorem 1.2] given in the case of finite-dimensional spaces.

(ii) Does (i) hold if $T\colon X \rightrightarrows X$ is a set-valued mapping admitting a Lipschitzian single-valued localization around $(\bar{x}, \bar{v}) \in \operatorname{gph} T$?

(ii) Is (i) valid for continuous single-valued mappings?

Exercise 5.19 (Hypomonotonicity of Subgradient Mappings). Let $f\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ be l.s.c. as in our standing assumption.

(i) Show that $\partial\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is semilocally hypomonotone around $\bar{x}$ if and only if there is a positive number ρ such that the function $f + \rho\|\cdot\|^2$ is convex on a neighborhood of $\bar{x}$. *Hint:* Deduce it from [678, Theorem 12.17].

(ii) Does the characterization in (i) hold in Hilbert spaces?

Exercise 5.20 (Calculus of Hypomonotonicity). Let X be a Hilbert space, and let $T_1\colon X \to X$ be continuous around $\bar{x}$ with $v_1 := T(\bar{x})$. Show that:

(i) If $T_2\colon X \rightrightarrows X$ is locally hypomonotone around $(\bar{x}, \bar{v}_2) \in \operatorname{gph} T_2$, then the sum $T_1 + T_2$ is locally hypomonotone around $(\bar{x}, \bar{v}_1 + \bar{v}_2)$.

(ii) If $T_2\colon X \rightrightarrows X$ is semilocally hypomonotone around $\bar{x}$, then the sum $T_1 + T_2$ is semilocally hypomonotone around this point.

(iii) Formulate and prove a version of this rule for global hypomonotonicity.

Hint: Proceed by the definitions of the corresponding hypomonotonicity notions.

Exercise 5.21 (Global Maximal Monotonicity for Mappings with Convex Domains). Reconstruct all the details in the proof of Theorem 5.4.

Exercise 5.22 (Coderivative Characterizations of Global Maximal Monotonicity for Set-Valued Mappings in Infinite Dimensions). Let $T: X \rightrightarrows X$ be a (closed-graph) set-valued operator defined on a Hilbert space X.

(i) Check that Theorem 5.2 and Theorem 5.4 hold true in Hilbert spaces.

(ii) Formulate and prove a Hilbert space version of Theorem 5.6 in terms of the mixed coderivative of T; cf. [153]. Give an example that the usage of the normal coderivative in (5.11) doesn't provide a necessary condition for global maximal monotonicity in infinite dimensions.

Exercise 5.23 (Coderivative Characterizations of Global Monotonicity for Single-Valued Continuous Mappings).

(i) Show that hypomonotonicity requirements are not needed in the coderivative characterizations of Theorems 5.2, 5.4, and 5.6 for single-valued continuous mappings $T: \mathbb{R}^n \to \mathbb{R}^n$; cf. [154]. Is it true for $T: X \to X$ in Hilbert spaces?

(ii) Find general conditions unifying hypomonotonicity of set-valued mappings and continuity of single-valued ones for the validity of the coderivative characterizations of global monotonicity in Section 5.1 in finite and infinite dimensions.

(iii) Using the symmetric subdifferential mean value theorem in Asplund spaces discussed in Exercise 4.33, derive appropriate versions of (i) for single-valued continuous mappings $T: X \to X^*$ in the case of Asplund spaces X.

Exercise 5.24 (Coderivative Characterizations of Global Strong Maximal Monotonicity in Hilbert Spaces). Derive Hilbert space extensions of the coderivative characterizations of this property presented in Corollary 5.9. *Hint:* Follow the proof of Corollary 5.9 by using the results taken from Exercise 5.22.

Exercise 5.25 (Preservation of Global Maximal Monotonicity and Strong Monotonicity Under Sums and Compositions).

(i) Based on the pointwise coderivative characterizations of global maximal monotonicity and strong monotonicity obtained in Subsection 5.2.3 and their infinite-dimensional versions from the exercises above, establish verifiable conditions for the preservation of these properties under sums and compositions by using the pointwise coderivative calculus developed in Chapter 4.

(ii) Compare the results obtained via (i) with known conditions for preserving maximal monotonicity; in particular, with Rockafellar's theorem [668] about the maximal monotonicity of sums under certain interiority assumptions.

Exercise 5.26 (Coderivative Characterizations of Local and Semilocal Maximal Monotonicity). Investigate the possibilities for deriving coderivative characterizations of the types given in Theorems 5.2, 5.4, and 5.6 for the notions of local and semilocal maximal monotonicity introduced similarly to the corresponding notions of hypomonotonicity from Definition 5.1(ii,iii).

Exercise 5.27 (Strong Metric Regularity of the Convex Subdifferential). Let X be a Banach space, and let $\varphi: X \to \overline{\mathbb{R}}$ be a l.s.c. convex function.

(i) Prove that for any $(\bar{x}, \bar{v}) \in \operatorname{gph} \partial\varphi$ the following assertions are equivalent:

• The subgradient mapping $\partial\varphi: X \rightrightarrows X^*$ is strongly metrically regular around $(\bar{x}, \bar{v})$ with modulus $\kappa > 0$.

• There are neighborhoods U of $\bar{x}$ and V of $\bar{v}$ such that the mapping $(\partial\varphi)^{-1}$ admits a single-valued localization $\vartheta: V \to U$ around $(\bar{v}, \bar{x})$ and that for any pair $(v, u) \in \operatorname{gph} \vartheta = \operatorname{gph} (\partial\varphi)^{-1} \cap (V \times U)$ we have the second-order growth condition

$$\varphi(x) \geq \varphi(u) + \langle v, x - u\rangle + \frac{1}{2\kappa}\|x - u\|^2 \quad \text{whenever} \quad x \in U. \tag{5.25}$$

Hint: Use the maximal monotonicity of the subdifferential mapping $\partial\varphi: X \rightrightarrows X^*$ and Fenchel duality; compare it with the proof of [551, Theorem 3.1].

(ii) Is strong metric regularity equivalent to metric regularity of $\partial\varphi$ around the same point in this setting? *Hint:* Use Kenderov's theorem [408], and compare this with the characterization of the metric regularity of $\partial\varphi$ given in Exercise 3.52.

Exercise 5.28 (Metric Regularity and Strong Metric Regularity of the Basic Subdifferential). Let $\varphi\colon X \to \overline{\mathbb{R}}$ be an arbitrary l.s.c. function on an Asplund space X, and let $(\bar{x}, \bar{v}) \in \operatorname{gph} \partial\varphi$ for the basic subgradient mapping (1.69).

(i) Prove that the following statements are equivalent:

• The subdifferential $\partial\varphi$ is metrically regular around $(\bar{x}, \bar{v})$ with modulus $\kappa > 0$, and there exist a real number $r \in [0, \kappa^{-1})$ and neighborhoods U of $\bar{x}$ and V of $\bar{v}$ such that for any pair $(u, v) \in \operatorname{gph} \partial\varphi \cap (U \times V)$ we have

$$\varphi(x) \geq \varphi(u) + \langle v, x - u\rangle - \frac{r}{2}\operatorname{dist}^2\big(x; (\partial\varphi)^{-1}(v)\big) \quad \text{whenever} \quad x \in U.$$

• There exist neighborhoods U of $\bar{x}$ and V of $\bar{v}$ such that for any $v \in V$ there is a point $u \in (\partial\varphi)^{-1}(v) \cap U$ satisfying (5.25)

• The subdifferential $\partial\varphi$ is metrically regular around $(\bar{x}, \bar{v})$ with modulus $\kappa > 0$, and there are neighborhoods U of $\bar{x}$ and V of $\bar{v}$ such that

$$\varphi(x) \geq \varphi(u) + \langle v, x - u\rangle \ \text{ for all } \ x \in U \ \text{ and } \ (u, v) \in \operatorname{gph} \partial\varphi \cap (U \times V).$$

• The point $\bar{x}$ is a local minimizer of the function $x \mapsto \varphi(x) - \langle v, x\rangle$, and the subdifferential $\partial\varphi$ is strongly metrically regular around $(\bar{x}, \bar{v})$ with modulus κ.

Hint: Proceed as in the proofs of [551, Theorem 3.2] with the usage of Ekeland's variational principle, the semi-Lipschitzian sum rule for the basic subdifferential, and the maximal monotonicity of the subdifferential of convex analysis.

(ii) Let X be Hilbert. Then all the statements above are equivalent to:

• ∂f is metrically regular around $(\bar{x}, \bar{v})$ with modulus $\kappa > 0$, and there are some $r \in [0, \kappa^{-1})$ and neighborhoods U of $\bar{x}$, V of $\bar{v}$ such that

$$\varphi(x) \geq \varphi(u) + \langle v, x - u\rangle - \frac{r}{2}\|x - u\|^2 \ \text{ for all } \ x \in U, \ (u, v) \in \operatorname{gph} \partial\varphi \cap (U \times V).$$

Hint: Arguing by contradiction, consider the function $\psi(x) := \psi(x) + \frac{r}{2}\|x - \bar{x}\|^2$, and show that $\partial\psi$ is metrically regular around $(\bar{x}, \bar{x}^*)$ with modulus $\frac{\kappa}{1-r\kappa}$. Then apply (i) and use the parallelogram law in Hilbert spaces; cf. [232, Corollary 3.8].

Exercise 5.29 (Equivalent Regularity Properties for $\mathcal{C}^2$-Smooth Functions). Let $\varphi\colon \mathbb{R}^n \to \mathbb{R}$ be twice continuously differentiable around its local minimizer $\bar{x}$. Check that the following properties are equivalent:

(a) The gradient mapping $\nabla\varphi\colon \mathbb{R}^n \to \mathbb{R}^n$ is metrically regular around $(\bar{x}, 0)$.

(b) The gradient mapping $\nabla\varphi$ is strongly metrically regular around $(\bar{x}, 0)$.

(c) The Hessian matrix $\nabla^2\varphi(\bar{x})$ is positive-definite.

(d) $\ker \nabla^2\varphi(\bar{x}) = \{0\}$ for the Hessian kernel $\ker \nabla^2\varphi(\bar{x}) := \{u|\ \nabla^2\varphi(\bar{x})u = 0\}$.

Hint: Deduce it from well-known facts of nonlinear analysis.

Exercise 5.30 (Equivalent Second-Order Conditions for Regularity Properties of Prox-Regular Functions). Let $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ be both prox-regular and subdifferentially continuous at $\bar{x}$ for $0 \in \partial\varphi(\bar{x})$.

(i) Prove that the following conditions are equivalent:

(a) The subgradient mapping $\partial\varphi\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is metrically regular around $(\bar{x}, 0)$ and the generalized Hessian $\partial^2\varphi(\bar{x}, 0)$ is positive-semidefinite in the sense that

$$\langle v, u\rangle \geq 0 \ \text{ whenever } \ v \in \partial^2\varphi(\bar{x}, 0)(u), \ u \in \mathbb{R}^n.$$

(b) The subgradient mapping $\partial\varphi$ is strongly metrically regular around $(\bar{x}, 0)$ and $\bar{x}$ is a local minimizer of φ.

(c) The generalized Hessian $\partial^2\varphi(\bar{x}, 0)$ is positive-definite in the sense that

$$\langle v, u\rangle > 0 \text{ whenever } v \in \partial^2\varphi(\bar{x}, 0)(u),\ u \neq 0.$$

(d) $\ker \partial^2\varphi(\bar{x}, 0) = \{0\}$ and $\partial^2\varphi(\bar{x}, 0)$ is positive-semidefinite.

Hint: Use the coderivative criterion for metric regularity from Theorem 3.3(ii), and proceed as in the proof of [232, Theorem 4.13].

(ii) Show that in the case where $\varphi \in \mathcal{C}^2$ around a local minimizer $\bar{x}$ of φ the equivalent conditions in (i) reduce to those in Exercise 5.29.

(iii) Given an example showing that the positive-semidefiniteness of the generalized Hessian $\partial^2\varphi(\bar{x}, 0)$ is not necessary for the local optimality of φ at $\bar{x}$ even in the case of fully amenable functions $\varphi\colon \mathbb{R}^2 \to \overline{\mathbb{R}}$.

Exercise 5.31 (Equivalence of Metric Regularity to Strong Metric Regularity of the Basic Subdifferential). Let $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ be prox-regular and subdifferentially continuous at $\bar{x}$ for $\bar{v} = 0$, and let $\bar{x}$ be a local minimizer of φ.

(i) Prove or disprove that the basic subgradient mapping $\partial\varphi$ is metrically regular around $(\bar{x}, \bar{v})$ if and only if it is strongly metrically regular around this point. *Hint:* Compare it with the results presented in Exercises 5.29 and 5.30, and see the corresponding discussions in Section 5.4.

(ii) Show that the equivalence (i) is certainly false outside the class of prox-regular and subdifferentially continuous functions on $\mathbb{R}^2$.

Exercise 5.32 (Strong Metric Subregularity and Isolated Calmness). A mapping $F\colon X \rightrightarrows Y$ between Banach spaces is STRONGLY METRICALLY SUBREGULAR at $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ with a positive modulus μ if there exist neighborhood U of $\bar{x}$ and V of $\bar{y}$ such that we have the estimate

$$\|x - \bar{x}\| \le \mu \operatorname{dist}\big(\bar{y}; F(x) \cap V\big) \text{ for all } x \in U.$$

The mapping $F\colon X \rightrightarrows Y$ enjoys the *isolated calmness* property at $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ with modulus $\ell \ge 0$ if there are neighborhoods U of $\bar{x}$ and V of $\bar{y}$ such that

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

If also $F(x) \cap V \neq \emptyset$ for all $x \in U$, then F has the *robust isolated property* at $(\bar{x}, \bar{y})$.

(i) Show that the isolated calmness of F at $(\bar{x}, \bar{y})$ is equivalent to the strong metric subregularity of the inverse mapping F^{-1} at $(\bar{y}, \bar{x})$. What about relationships between moduli and their exact bounds? *Hint:* Proceed similarly the proof of Theorem 3.2, and compare it with [227, Theorem 3I.3].

(ii) Find conditions on F ensuring that the isolated calmness of F at $(\bar{x}, \bar{y})$ agrees with its robust counterpart, and give an example that this fails in general.

(iii) Which property of the inverse mapping F^{-1} at $(\bar{y}, \bar{x})$ is equivalent to the robust isolated calmness of F at $(\bar{x}, \bar{y})$?

Exercise 5.33 (Graphical Derivative Characterizations of Isolated Calmness of Multifunctions). Let $F\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, and let $(\bar{x}, \bar{y}) \in \operatorname{gph} F$.

(i) Show that the graphical derivative condition

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

in terms of (1.87) is necessary and sufficient for the isolated calmness of F at $(\bar{x}, \bar{y})$. *Hint:* Proceed directly by the definitions of isolated calmness and graphical derivative, and compare it with the proofs of sufficiency in [417] and of necessity in [453].

(ii) Verify the possibility to deduce the calmness characterization in (i) from that for the upper Lipschitzian property via the outer derivative (3.70) from [771] presented in Exercise 3.55(i). *Hint:* Use the result of Exercise 3.54(i).

(ii) Derive a formula for the exact bound of isolated calmness in (i). *Hint:* Compare it with the result and proof in [227, Theorem 4E.1] for the equivalent property of strong metric regularity.

(iii) Do the results in (i) and (ii) hold in infinite dimensions?

Exercise 5.34 (Strong Metric Subregularity and Strong Local Monotonicity of the Convex Subdifferential). Let $\varphi\colon X \to \overline{\mathbb{R}}$ be a l.s.c. convex function on a Banach space X, and let $\bar{v} \in \partial\varphi(\bar{x})$. Prove that the following are equivalent:

(a) The subdifferential mapping $\partial\varphi$ is strongly metrically subregular at $(\bar{x}, \bar{v})$.

(b) There exist a neighborhood U of $\bar{x}$ and a constant $\gamma > 0$ such that

$$\varphi(x) \geq \varphi(\bar{x}) + \langle \bar{v}, x - \bar{x}\rangle + \gamma\|x - \bar{x}\|^2 \quad \text{for all} \quad x \in U.$$

(c) There exist a neighborhood U of $\bar{x}$ and a constant $\gamma > 0$ such that

$$\langle \bar{v} - v, x - \bar{x}\rangle \geq \gamma\|x - \bar{x}\|^2 \quad \text{for all} \quad x \in U, \ v \in \partial\varphi(\bar{x}).$$

Hint: Deduce it from Exercise 5.27(i) and the definitions; cf. [21, Theorem 3.6].

Exercise 5.35 (Strong Metric Subregularity of the Basic Subdifferential). Given a l.s.c. function $\varphi\colon X \to \overline{\mathbb{R}}$ on an Asplund space X and given a pair $(\bar{x}, \bar{v} \in \operatorname{gph}\partial\varphi$, consider the following two statements:

(a) The subdifferential $\partial\varphi$ is strongly metrically subregular at $(\bar{x}, \bar{v})$ with modulus $\kappa > 0$, and there are real numbers $r \in (0, \kappa^{-1})$ and $\nu > 0$ such that

$$\varphi(x) \geq \varphi(\bar{x}) + \langle \bar{v}, x - \bar{x}\rangle - \frac{r}{2}\|x - \bar{x}\|^2 \quad \text{for all} \quad x \in \bar{x} + \nu\mathbb{B}.$$

(b) There are real numbers $\alpha, \eta > 0$ such that

$$\varphi(x) \geq \varphi(\bar{x}) + \langle \bar{v}, x - \bar{x}\rangle + \frac{\alpha}{2}\|x - \bar{x}\|^2 \quad \text{for all} \quad x \in \bar{x} + \eta\mathbb{B}.$$

(i) Prove that (a)$\Longrightarrow$(b) holds, where α may be chosen arbitrarily in $(0, \kappa^{-1})$.

(ii) Verify that the converse implication (b)$\Longrightarrow$(a) also holds if in addition there is some number $\beta \in [0, \alpha)$ with

$$\varphi(\bar{x}) \geq \varphi(x) + \langle v, \bar{x} - x\rangle - \frac{\beta}{2}\|x - \bar{x}\|^2 \quad \text{for all} \quad (x, v) \in \operatorname{gph}\partial\varphi \cap \big[(\bar{x}, \bar{v}) + \eta\mathbb{B}].$$

Hint: Deduce both conclusions from the results presented in Exercise 5.28(i) by proceeding similarly to the proof of [232, Corollary 3.3].

Exercise 5.36 (Strong Local Maximal Monotonicity in Hilbert Spaces). Let $T\colon X \rightrightarrows X$ be a set-valued operator defined on a Hilbert space X.

(i) Analyzing the proofs of Theorem 5.13 and Theorem 5.14, check that these results hold in infinite dimensions.

(ii) Is it true for Theorem 5.16 in terms of the mixed coderivative D^*_M?

Exercise 5.37 (Coderivative Conditions for Strong Metric Regularity). Construct examples in finite-dimensional spaces showing that the conditions of Corollary 5.15 and of Theorem 5.16 are not necessary for strong metric regularity in both set-valued and single-valued cases.

Exercise 5.38 (Limiting Coderivative Characterization of Local Strong Maximal Monotonicity). Let $T\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be a (closed-graph) set-valued mapping with $(\bar{x}, \bar{v}) \in \operatorname{gph} T$. Prove or disprove the following *conjecture*: T is locally strongly maximal monotone around $(\bar{x}, \bar{v})$ if and only if T is locally hypomonotone around $(\bar{x}, \bar{v})$ and the positive-definiteness condition (5.24) is satisfied; cf. Remark 5.17.

Exercise 5.39 (Coderivative Characterizations of Strong Semilocal Monotonicity). Investigate the possibilities for deriving coderivative characterizations of the types given in Theorems 5.14 and 5.16 for the notion of strong semilocal monotonicity introduced similarly to semilocal hypomonotonicity in Definition 5.1(ii).

5.4 Commentaries to Chapter 5

Section 5.1. As already mentioned in Section 4.6, global monotonicity and especially its maximal manifestation have been highly recognized among the most fundamental developments of nonlinear and variational analysis with great many applications to theoretical and numerical aspects of optimization-related and equilibrium problems; see, e.g., the references given above. The approaches and results presented in this chapter are based on quite recent developments while communicating new ideas in the study and applications of monotonicity by using the appropriate tools of generalized differentiation in variational analysis.

The main results of Section 5.1 (Theorems 5.2, 5.4, and 5.6), as well as their infinite-dimensional versions formulated in the exercises, are taken from the author's joint paper with Chieu, Lee, and Nghia [153], which contains complete *coderivative characterizations* of *global maximal monotonicity* for general set-valued operators in Hilbert spaces. These results present far-going nonsmooth extensions of the classical criterion for monotonicity of smooth functions in terms of the positive-semidefiniteness of their derivatives. Note that the *mixed coderivative* (1.65) is used in Theorem 5.6 for the limiting characterization (5.11) in infinite dimensions.

Implication (i)⇒(ii) of Theorem 5.6 was first obtained by Poliquin and Rockafellar [642] in finite dimensions and then was extended in [152, 551] to Hilbert spaces; we follow here the proof given in [551]. For single-valued continuous mappings in Hilbert spaces, the regular coderivative characterization (5.2) of global monotonicity was obtained by Chieu and Trang [154], while its limiting version of Theorem 5.6 was given in [154] in finite-dimensional spaces.

To the best of our knowledge, the concept of *hypomonotonicity* was introduced by Rockafellar [673] who utilized its semilocal (in our terminology) version for certain subdifferential operators; see also [641, 678]. Local hypomonotonicity was employed by Levy and Poliquin [455] in the study of Lipschitzian stability, by Pennanen [633] in developing the proximal point and related methods of numerical optimization, and by the author and Nghia [555] in characterizing the strong local maximal monotonicity property of general operators (see Section 5.2) with applications to full stability of parametric variational systems. Global hypomonotonicity was implemented, e.g., in the book by Burachik and Iusem [130] (see also the references therein) to study enlargements of monotone operators.

It is important to emphasize that all the three classes of hypomonotone operators considered in Definition 5.1 are *sufficiently broad* and contain, in particular, Lipschitzian single-valued mappings and set-valued subdifferential mappings generated by "nice" functions, which are *prox-regular* and *subdifferentially continuous* (for local hypomonotonicity), *lower*–$\mathcal{C}^2$ on open sets (for the semilocal version), etc.; see [641, 678] for more details. That is, the hypomonotonicity properties are not restrictive for numerous applications in variational analysis and optimization.

The notion of *global strong monotonicity*, which maximality is characterized in Corollary 5.9, goes back to Zarantonello [759, 760] who used it for justifying the convergence of some numerical algorithms to solve functional equations.

Section 5.2. The results presented in this section are taken from the paper by Mordukhovich and Nghia [555], which contains also their applications to second-order (particularly *full stability* of subdifferential variational systems) in finite and infinite dimensions. The main emphasis here is on verifiable characterizations of *local strong maximal* monotonicity of set-valued mappings given in Theorem 5.13 and Theorem 5.14 and the pointbased coderivative criterion of Theorem 5.16

in the case of single-valued Lipschitzian ones, while the condition obtained therein is necessary for such a monotonicity in the general set-valued setting; see also Remark 5.17. Observe that the proofs of Theorems 5.13 and 5.14 work in any Hilbert space, but Theorem 5.16 seems to be finite-dimensional.

It follows from Theorem 5.16 that the positive-definiteness coderivative condition (5.24) is sufficient for the *strong metric regularity* of T around $\bar{x}$, i.e., for the existence of a single-valued Lipschitzian localization of the inverse mapping T^{-1} around $(T(\bar{x}), \bar{x})$. This notion is an abstract version of Robinson's strong regularity originally introduced [662] for solution maps to generalized equations via their linearization; see [227] for more discussions. A necessary and sufficient condition for the latter property was established by Kummer [443] in terms of Thibault's *strict derivative* [705] of T at $\bar{x}$; see also [678] for more details on this construction. Recall to this end that the coderivative condition (5.24) provides a complete characterization of the local strong maximal monotonicity of T around $\bar{x}$, which is a *much weaker* property than strong metric regularity around this point.

Observe that there are different motivations and formalizations of local maximal monotonicity (compare, e.g., [112, 455, 633, 641, 642]). This book and the preceding paper [555] adopt in the setting of strong local monotonicity the one defined by Poliquin and Rockafellar [642, p. 290]. In this way, besides complete *coderivative characterizations* of strong local maximal monotonicity, we obtain verifiable sufficient conditions for strong metric regularity. Note that for locally *monotone* mappings, we have in fact the *equivalence* between metric regularity and strong metric regularity around the reference point; see [227, Theorem 3G.5], which is a particular case of the fundamental result by Kenderov [408] already used in Chapter 3.

Section 5.3. The *coderivative approach* to global and local maximal monotonicity and the results presented in this chapter open new perspectives in the study and applications of maximal monotonicity in both finite and infinite dimensions. Some of the challenging *open questions* are formulated as "exercises" in Section 5.3. To this end we mention Exercises 5.25, 5.26, and 5.39 and the related Remarks 5.7 and 5.17.

Second-order growth conditions for *strong metric regularity* of the convex and basic subdifferentials presented in Exercises 5.27 and 5.28 are mainly taken from Mordukhovich and Nghia [551], with the quantitative relations therein; see also Aragón and Geoffroy [20, 21] and Drusvyatskiy and Lewis [230] for some related results in this direction. Note that the *monotonicity* issues play a significant role in the proofs; cf. the papers by Rockafellar and his collaborators [456, 642] on stability of local minimizers, which have also been behind the motivations and results of [551].

The results of Exercise 5.30 are obtained by Drusvyatskiy, Mordukhovich, and Nghia [232], where the authors formulated the statement of Exercise 5.31 as a *conjecture*. Besides the $\mathcal{C}^2$-smooth cases presented in Exercise 5.29, this conjecture is known to be true for broad classes of continuously prox-regular functions $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$; in particular, for convex ones due to Kenderov's theorem [408] and the maximal monotonicity of the convex subdifferential, for functions of the type $\varphi(x) = \varphi_0 + \delta_\Omega(x)$ with $\varphi_0 \in \mathcal{C}^2$ and a polyhedral convex set Ω due to Dontchev and Rockafellar [225], and for functions of the latter type with Ω being a second-order/Lorentz cone due to Outrata and Ramírez [625], as well as in the setting of Exercise 5.30. However, in the general case of prox-regular and subdifferentially continuous functions, this conjecture remains a challenging and very important *open question*.

The notion of *strong metric subregularity* and the equivalent notion of *isolated calmness* for inverse mappings have been long time studied in the literature under different names (or without giving a name); see, e.g., [94, 220, 226, 417, 420, 453, 632, 677] for early publications. The aforementioned terminology was suggested by Dontchev and Rockafellar and by now has been widely used; see [227] along with the recent publications [20, 21, 23, 151, 160, 228, 232, 286, 433, 561, 562, 577, 724].

Although the explicit proof of sufficiency of the graphical derivative condition for isolated calmness in Exercise 5.33(i) is given by King and Rockafellar [417] and the necessity of this condition is proved by Levy [453], this criterion is actually goes back to the earlier paper by Rockafellar [677]. It also can be deduced from Zhang and Treiman [771]; see Exercise 5.33(ii).

The quadratic growth and strong monotonicity characterizations of strong metric subregularity of the convex subdifferential from Exercise 5.34 are due to Aragón and and Geoffroy [20, 21], while the results for the basic subdifferential with the quantitative interplay between constraints in Exercise 5.35 are taken from Drusvyatskiy, Mordukhovich, and Nghia [232]. It is easy to find functions $\varphi\colon \mathbb{R} \to \mathbb{R}$ for which the strong metric subregularity of the basic subdifferential at *local minimizers* cannot be characterized by the quadratic growth of φ; see [232]. It is shown lately by Drusvyatskiy and Ioffe [228] that such a characterization holds for the class of subdifferentially continuous and *semialgebraic* functions $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$, which is different from both classes of $\mathcal{C}^2$-smooth and convex functions while being nevertheless important for some applications in variational analysis and nonsmooth optimization.

Quite recently [209], a robust version of isolated calmness for a set-valued mapping F at $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ with the additional requirement that $F(x) \cap U \neq \emptyset$ for all $x \in V$ has been labeled as the *robust isolated calmness*. Note that this property was actually employed earlier in particular settings without naming it; see [94, 226, 420, 632]. If the set-valued mapping F is lower semicontinuous at $(\bar{x}, \bar{y})$ in the standard topological sense, then isolated calmness implies its robust counterpart. However, it doesn't hold in general as shown, e.g., in [562, Example 6.4]. In fact, the usage of robust isolated calmness in numerical optimization has been recognized in the literature starting with the 1990s. In particular, the sharpest result for the sequential quadratic programming (SQP) method for solving NLPs, obtained by Bonnans [94], imposes the strict Mangasarian-Fromovitz constraint qualification together with the conventional second-order sufficient condition for NLPs. As later proved by Dontchev and Rockafellar [226, Theorem 2.6], the simultaneous validity of these conditions characterizes robust isolated calmness of solutions maps of canonically perturbed KKT systems in NLPs. Recently this result has been extended by Ding, Sun, and Zhang [209] to some nonpolyhedral problems of constrained optimization under the so-called strict Robinson constraint qualification . Mordukhovich and Sarabi [577] characterized robust isolated calmness for generalized KKT systems in problems of composite optimization relating this notion to noncriticality of Lagrange multipliers associated with local minimizers for such problems. The latter notion extends the one introduced by Izmailov and Solodov [383] for the classical KKT systems in NLPs. Furthermore, it is shown in [577] that the Lipschitz-like property of solution maps to generalized KKT systems for composite optimization problems yields their robust isolated calmness at the corresponding points.

Chapter 6
Nondifferentiable and Bilevel Optimization

It is not accidental that we unify the exposition of these two areas of optimization theory in one chapter. It has been widely recognized that problems of *nondifferentiable/nonsmooth optimization* (i.e., those containing nondifferentiable functions and/or sets with nonsmooth boundaries in their objectives and/or constraints) naturally and frequently appear in different aspects of variational analysis and numerous applications while being very challenging from both theoretical and algorithmic viewpoints. On the other hand, problems of *bilevel optimization* are *intrinsically nonsmooth*, even in the case of fully smooth data at their lower and upper levels. In fact, they can be reduced to single-level optimization problems, but the price to pay is the unavoidable presence of nonsmooth functions as a result of such reductions, regardless of smoothness assumptions imposed on the given data.

The main emphasis of this chapter is obtaining efficient first-order *necessary optimality conditions* for problems of *nondifferentiable programming* and then applying them to *bilevel programs* with smooth and nonsmooth functions at both levels of optimization. To proceed in these directions, we rely on the constructions and results of variational analysis and generalized differentiation developed in the previous chapters of the book.

6.1 Problems of Nondifferentiable Programming

We start with deriving necessary optimality conditions for problems of nonsmooth *minimization* with geometric constraints given by closed sets and then extend them to general problems of nondifferentiable programming with functional constraints described by finitely many inequalities and equalities.

B. S. Mordukhovich, *Variational Analysis and Applications*, Springer Monographs in Mathematics, https://doi.org/10.1007/978-3-319-92775-6_6

6.1.1 Lower and Upper Subdifferential Conditions

Given $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ and $\Omega \subset \mathbb{R}^n$, consider the problem:

$$\text{minimize} \;\; \varphi(x) \;\; \text{subject to} \;\; x \in \Omega. \tag{6.1}$$

Our goal here is to obtain necessary conditions for (feasible) local minimizers $\bar{x} \in \operatorname{dom}\varphi \cap \Omega$ in (6.1). We derive two different types of necessary optimality conditions. Conditions of the first type, called the *lower subdifferential optimality conditions*, are expressed in terms of the basic subdifferential (1.24) under appropriate qualification conditions formulated in terms of the singular subdifferential (1.25). Conditions of the second type, called the *upper subdifferential optimality conditions*, make use of the upper regular subdifferential (1.76) of the cost function φ that is equivalently described as

$$\widehat{\partial}^+\varphi(\bar{x}) = -\widehat{\partial}(-\varphi)(\bar{x}), \quad |\varphi(\bar{x})| < \infty. \tag{6.2}$$

Note that (6.2) may be empty for broad classes of nonsmooth functions (e.g., for convex functions nondifferentiable at $\bar{x}$) while giving more selective necessary conditions for minimization than the lower subdifferential ones in certain "upper regular" settings; see the results, examples, and discussions below.

As before, we always assume without loss of generality that cost functions are *l.s.c.* around the reference points (although it is not needed for upper subdifferential conditions) and constraint sets are locally *closed* around them.

The following theorem contains necessary optimality conditions of both types for problem (6.1). Observe that both of them are derived from the *variational/extremal principles*. Indeed, the upper subdifferential conditions are induced by the smooth variational description of regular subgradients. To establish the lower subdifferential optimality conditions, we employ the basic subdifferential sum rule, which follows from the extremal principle. In fact, the extremal principle can be used directly; see, e.g., the proof of Theorem 6.5 below for problems involving functional and geometric constraints.

Theorem 6.1 (Optimality Conditions for Problems with a Single Geometric Constraint). *Let $\bar{x} \in \operatorname{dom}\varphi \cap \Omega$ be a local optimal solution to the minimization problem* (6.1). *The following assertions hold:*

(i) *The entire set of upper regular subgradients satisfies the inclusions*

$$-\widehat{\partial}^+\varphi(\bar{x}) \subset \widehat{N}(\bar{x};\Omega), \quad -\widehat{\partial}^+\varphi(\bar{x}) \subset N(\bar{x};\Omega). \tag{6.3}$$

(ii) *Under the qualification condition*

$$\partial^\infty\varphi(\bar{x}) \cap \big(-N(\bar{x};\Omega)\big) = \{0\} \tag{6.4}$$

, we have the lower subdifferential relationships

$$\partial\varphi(\bar{x}) \cap \big(-N(\bar{x};\Omega)\big) \neq \emptyset, \;\; \textit{i.e.,} \;\; 0 \in \partial\varphi(\bar{x}) + N(\bar{x};\Omega). \tag{6.5}$$

Proof. To justify assertion (i), it suffices to verify only the first inclusion in (6.3) since $\widehat{N}(\bar{x};\Omega) \subset N(\bar{x};\Omega)$ by Theorem 1.6. To proceed with this task, suppose that $\widehat{\partial}^+\varphi(\bar{x}) \neq \emptyset$ (there is nothing to prove otherwise), and pick any $v \in \widehat{\partial}^+\varphi(\bar{x})$. Using (6.2) and applying the first part of Theorem 1.27 (which holds without the l.s.c. assumption on φ), we find a function $\psi\colon \mathbb{R}^n \to \mathbb{R}$ with $\psi(\bar{x}) = \varphi(\bar{x})$ and $\psi(x) \geq \varphi(x)$ whenever $x \in \mathbb{R}^n$ such that ψ is (Fréchet) differentiable at $\bar{x}$ and $\nabla\psi(\bar{x}) = v$. It gives us

$$\psi(\bar{x}) = \varphi(\bar{x}) \leq \varphi(x) \leq \psi(x) \ \text{ for all } \ x \in \Omega \ \text{ close to } \ \bar{x}$$

showing therefore that $\bar{x}$ is a local minimizer of the constrained problem:

$$\text{minimize } \ \psi(x) \ \text{ subject to } \ x \in \Omega,$$

where the cost function is differentiable at $\bar{x}$. This problem can be equivalently written in the form of unconstrained optimization:

$$\text{minimize } \ \psi(x) + \delta(x;\Omega), \quad x \in \mathbb{R}^n.$$

Employing in the latter setting the generalized Fermat rule from Proposition 1.30(i) and then the regular subdifferential sum rule from Proposition 1.30(ii) with taking into account that $\nabla\psi(\bar{x}) = v$, we get

$$0 \in \widehat{\partial}\big(\psi + \delta(\cdot;\Omega)\big)(\bar{x}) = \nabla\psi(\bar{x}) + \widehat{N}(\bar{x};\Omega) = v + \widehat{N}(\bar{x};\Omega).$$

This yields $-v \in \widehat{N}(\bar{x};\Omega)$ for any $v \in \widehat{\partial}^+\varphi(\bar{x})$ and thus verifies (i).

To prove assertion (ii), we apply the generalized Fermat rule to the local optimal solution $\bar{x}$ of problem (6.1) written in the unconstrained form:

$$\text{minimize } \ \varphi(x) + \delta(x;\Omega), \quad x \in \mathbb{R}^n,$$

and then deduce from the basic subdifferential sum rule of Theorem 2.19 that

$$0 \in \partial\big(\varphi + \delta(\cdot;\Omega)\big)(\bar{x}) \subset \partial\varphi(\bar{x}) + N(\bar{x};\Omega)$$

provided the validity of the qualification condition (6.4) due to Proposition 1.19. This verifies (6.5) and completes the proof of the theorem. △

Let us discuss some particular features of the lower and upper subdifferential conditions from Theorem 6.1 and relationships between them.

Remark 6.2 (Upper vs. Lower Subdifferential Optimality Conditions).

(i) Note first that in the case where φ is (Fréchet) differentiable at $\bar{x}$, the optimality conditions in (6.3) reduce to

$$-\nabla\varphi(\bar{x}) \in \widehat{N}(\bar{x};\Omega), \quad -\nabla\varphi(\bar{x}) \in N(\bar{x};\Omega),$$

while only the second inclusion can be derived from (6.5) provided that φ is *strictly* differentiable at $\bar{x}$. On the other hand, the upper subdifferential conditions in (6.3) are trivial when $\widehat{\partial}^+\varphi(\bar{x}) = \emptyset$, which is the case of, e.g., convex continuous functions nondifferentiable at $\bar{x}$. In contrast, the lower subdifferential condition (6.5) is nontrivial for broad collections of nonsmooth functions including, e.g., every locally Lipschitzian function φ for which $\partial\varphi(\bar{x}) \neq \emptyset$ and the qualification condition (6.4) holds due to $\partial^\infty\varphi(\bar{x}) = \{0\}$ by Theorem 1.22.

(ii) Note also that the *triviality condition* $\widehat{\partial}^+\varphi(\bar{x}) = \emptyset$ itself is an easy checkable *necessary condition for optimality* in (6.1) provided that φ is nondifferentiable at $\bar{x}$ and $\Omega = \mathbb{R}^n$. Indeed, in this case, we have the inclusion $0 \in \widehat{\partial}\varphi(\bar{x}) \neq \emptyset$ by the generalized Fermat rule and hence $\widehat{\partial}^+\varphi(\bar{x}) = \emptyset$ by the simple observation from Exercise 1.76(ii).

(iii) Recall that φ is *upper regular* at $\bar{x}$ if $\widehat{\partial}^+\varphi(\bar{x}) = \partial^+\varphi(\bar{x})$. Note that, besides concave functions and differentiable ones, this class includes, e.g., a rather large class of *semiconcave* functions important in various applications to optimization and control; see, e.g., [136, 523]. If φ is upper regular at $\bar{x}$ and locally Lipschitzian around this point, we have $\widehat{\partial}^+\varphi(\bar{x}) = -\partial(-\varphi)(\bar{x}) \neq \emptyset$ by Theorem 1.22, i.e., the upper subdifferential conditions in (6.3) definitely give us a nontrivial information. Furthermore, in this case, we also have $\overline{\partial}\varphi(\bar{x}) = \widehat{\partial}^+\varphi(\bar{x})$ for Clarke's generalized gradient due to its plus-minus symmetry (1.79). Taking into account that the inclusions in (6.3) are valid for the *entire set* of upper subgradients, these observations show that the upper subdifferential optimality conditions may have sizable advantages over the lower subdifferential ones from Theorem 6.1(ii).

(iv) Let us consider in more detail problems of *concave minimization*, i.e., when the cost function φ is concave in (6.1). This class is of significant interest for various aspects of optimization theory and applications; in particular, from the viewpoints of *global* optimization; see, e.g., [355]. When φ is concave and continuous around $\bar{x}$, it follows from Exercise 1.77 that

$$\partial\varphi(\bar{x}) \subset \partial^+\varphi(\bar{x}) = \widehat{\partial}^+\varphi(\bar{x}) \neq \emptyset.$$

Then comparing the second inclusion in (6.3) (which is even weaker than the first inclusion therein) with the lower subdifferential condition in (6.5), we see that the necessary condition of Theorem 6.1(i) requires that *every* element v of the set $\widehat{\partial}^+\varphi(\bar{x})$ must belong to $-N(\bar{x};\Omega)$, instead of that *some* element v from the *smaller* set $\partial\varphi(\bar{x})$ belongs to $-N(\bar{x};\Omega)$ in Theorem 6.1(ii). Let us illustrate it by the following simple *example*:

$$\text{minimize } \ \varphi(x) := -|x| \ \text{ subject to } \ x \in \Omega := [-1, 0] \subset \mathbb{R}.$$

Obviously $\bar{x} = 0$ is not an optimal solution to this problem. However, it cannot be taken away by the lower subdifferential condition (6.5) due to

$$\partial\varphi(0) = \{-1, 1\}, \quad N(0;\Omega) = [0, \infty), \quad \text{and} \quad -1 \in -N(0;\Omega).$$

On the other hand, checking the upper subdifferential condition (6.3) gives us

$$\widehat{\partial}^+\varphi(0) = [-1, 1] \quad \text{and} \quad [-1, 1] \not\subset N(0; \Omega),$$

which confirms that $\bar{x} = 0$ is not optimal in (6.1), and thus (6.3) is a more selective necessary condition for optimality in the problem under consideration.

Observe further that minimization problems for *differences of two convex* (*DC*) *functions* can be equivalently reduced to minimizing concave functions subject to convex constraints. This allows us to deduce necessary conditions for such problems from the upper subdifferential conditions of Theorem 6.1(i).

Proposition 6.3 (DC Optimization Problems). *Consider the problem:*

$$\textit{minimize} \;\; \varphi_1(x) - \varphi_2(x), \quad x \in \mathbb{R}^n, \tag{6.6}$$

where $\varphi_1, \varphi_2 \colon \mathbb{R}^n \to \overline{\mathbb{R}}$ *are convex under the convention that* $\infty - \infty := \infty$. *Then* $\bar{x}$ *is a local minimizer of* (6.6) *if and only if the pair* $(\bar{x}, \varphi_1(\bar{x}))$ *gives a local minimum to the following problem on minimizing a concave function subject to convex geometric constraints:*

$$\textit{minimize} \;\; \psi(x, \alpha) := \alpha - \varphi_2(x) \;\; \textit{subject to} \;\; (x, \alpha) \in \operatorname{epi} \varphi_1. \tag{6.7}$$

Moreover, the upper subdifferential condition (6.3) *for* (6.7) *reduces to the* (*lower*) *subdifferential inclusion* $\partial\varphi_2(\bar{x}) \subset \partial\varphi_1(\bar{x})$.

Proof. If $\bar{x}$ solves (6.6) locally, i.e., there is a neighborhood U of $\bar{x}$ such that

$$\varphi_1(\bar{x}) - \varphi_2(\bar{x}) \le \varphi_1(x) - \varphi_2(x) \;\; \text{for all} \;\; x \in U,$$

then for $\bar{\alpha} := \varphi_1(\bar{x})$, we obviously have

$$\bar{\alpha} - \varphi_2(\bar{x}) \le \alpha - \varphi_2(x) \;\; \text{whenever} \;\; (x, \alpha) \in (U \times \mathbb{R}) \cap \operatorname{epi} \varphi_1,$$

which means that $(\bar{x}, \bar{\alpha})$ locally solves problem (6.7). Conversely, suppose that there exist $\varepsilon > 0$ and a neighborhood U of $\bar{x}$ such that

$$\varphi_1(\bar{x}) - \varphi_2(\bar{x}) \le \alpha - \varphi_2(x) \;\; \text{for all} \;\; \alpha \ge \varphi_1(x), \;\; x \in U, \;\; |\alpha - \varphi_1(\bar{x})| < \varepsilon.$$

Since φ_1 is convex and finite around $\bar{x}$ by the above, it is (Lipschitz) continuous around this point. Thus there is a neighborhood $\widetilde{U}$ of $\bar{x}$ on which

$$|\varphi_1(x) - \varphi_1(\bar{x})| < \varepsilon, \;\; \text{and so} \;\; \varphi_1(\bar{x}) - \varphi_2(\bar{x}) \le \varphi_1(x) - \varphi_2(x), \quad x \in \widetilde{U}.$$

This verifies that $\bar{x}$ is a local solution to (6.6).

It remains to show that the upper subdifferential optimality condition

$$-\widehat{\partial}^+\psi\big(\bar{x}, \varphi_1(\bar{x})\big) \subset N\big((\bar{x}, \varphi_1(\bar{x})); \operatorname{epi} \varphi_1\big) \tag{6.8}$$

for (6.7) reduces to the subdifferential inclusion claimed in the proposition. Indeed, we get by the direct calculations that

$$-\widehat{\partial}^+\psi\big(\bar{x},\varphi_1(\bar{x})\big)=\widehat{\partial}\big(\varphi_2-\alpha\big)\big(\bar{x},\varphi_1(\bar{x})\big)=\partial\varphi_2(\bar{x})\times\{0\}+\{0\}\times\{-1\}$$
$$=\partial\varphi_2(\bar{x})\times\{-1\}.$$

Hence the upper subdifferential inclusion (6.8) implies that

$$(v,-1)\in N\big((\bar{x},\varphi_1(\bar{x}));\operatorname{epi}\varphi_1\big)\;\text{ for all }\;v\in\partial\varphi_2(\bar{x}),$$

which is equivalent to $v\in\partial\varphi_1(\bar{x})$ for all $v\in\partial\varphi_2(\bar{x})$ and thus justifies the claimed necessary optimality condition $\partial\varphi_2(\bar{x})\subset\partial\varphi_1(\bar{x})$ in (6.6). △

The crucial advantage of the second upper subdifferential inclusion in (6.3) in comparison with the first one and also a strong feature of the lower subdifferential qualification and optimality conditions are well-developed *calculus rules* available for basic normals and subgradients in contrast to their regular counterparts. In particular, calculus results obtained in Chapter 2 allow us to derive various consequences of both assertions (i) and (ii) of Theorem 6.1 in cases where Ω is represented as a product and a sum of finitely many sets, as an inverse image of another set under a set-valued mapping, as a system of inequalities and/or equalities, etc. Qualification conditions that ensure the validity of the obtained representations of $N(\bar{x};\Omega)$ are transferred in this way into *constraint qualifications* under which the corresponding necessary optimality conditions hold in the *qualified/normal/KKT* (*Karush-Kuhn-Tucker*) *form*, i.e., with no (=1) multiplier associated with the cost function; see below.

Next we present both upper and lower subdifferential optimality conditions obtained in this scheme for problems with finitely many geometric constraints.

Proposition 6.4 (Optimality Conditions for Problems with Many Geometric Constraints). *Consider the problem:*

$$\text{minimize }\;\varphi(x)\;\text{ subject to }\;x\in\Omega_i\;\text{ for }\;i=1,\ldots,s, \tag{6.9}$$

and suppose that $\bar{x}\in\operatorname{dom}\varphi\cap\Omega_1\cap\ldots\cap\Omega_s$ *is a local minimizer for* (6.9). *Then the following upper subdifferential and lower subdifferential necessary optimality conditions hold at* $\bar{x}$*:*

(i) *Under the validity of the constraint qualification*

$$\big[v_1+\ldots+v_s=0,\;v_i\in N(\bar{x};\Omega_i)\big]\Longrightarrow v_i=0\;\text{ for all }\;i=1,\ldots,s, \tag{6.10}$$

we have the upper subdifferential inclusion

$$-\widehat{\partial}^+\varphi(\bar{x})\subset N(\bar{x};\Omega_1)+\ldots+N(\bar{x};\Omega_s).$$

(ii) *Under the validity of the qualification condition*

$$\Big[v+\sum_{i=1}^{s}v_i=0\ \textit{for}\ \ v\in\partial^{\infty}\varphi(\bar{x}),\ \ v_i\in N(\bar{x};\Omega_i)\Big]\Longrightarrow v=v_1=\ldots=v_s=0$$

stronger than (6.10)*, we have the lower subdifferential inclusion*

$$0\in\partial\varphi(\bar{x})+N(\bar{x};\Omega_1)+\ldots+N(\bar{x};\Omega_s).$$

Proof. Necessary optimality conditions in both assertions (i) and (ii) follow directly from the corresponding results of Theorem 6.1 and the normal intersection rule for finitely many sets given in Corollary 2.17. △

6.1.2 Finitely Many Inequality and Equality Constraints

Let us consider here the problem of *nondifferentiable programming*:

$$\begin{cases}\text{minimize}\ \ \varphi_0(x)\ \ \text{subject to}\\ \varphi_i(x)\le 0,\quad i=1,\ldots,m,\\ \varphi_i(x)=0,\quad i=m+1,\ldots,m+r,\\ x\in\Omega\subset\mathbb{R}^n\end{cases}\tag{6.11}$$

with finitely many inequality and equality constraints while keeping geometric constraints as well. In what follows we derive various necessary optimality conditions of both lower subdifferential and upper subdifferential types for local solutions to program (6.11) depending on assumptions imposed on their initial data and proof techniques. Our first theorem presents general necessary optimality conditions of the lower subdifferential type expressed via normals and subgradients of each function and set in (6.11) *separately*. The proof is based on the direct application of the *extremal principle* from Theorem 2.3. Recall that, unless otherwise stated, all the functions in question are assumed to be *lower semicontinuous* around the reference points.

Theorem 6.5 (Lower Subdifferential Conditions via Normals and Subgradients of Separate Constraints). *Let $\bar{x}$ be a feasible solution to* (6.11)*, that is, a local minimizer for this problem. The following necessary optimality conditions hold at $\bar{x}$:*

(i) *Assume that the equality constraint functions φ_i are continuous around $\bar{x}$ for all $i=m+1,\ldots,m+r$. Then there are elements $(v_i,\lambda_i)\in\mathbb{R}^{n+1}$ for $i=0,\ldots,m+r$, not equal to zero simultaneously, and a vector $v\in\mathbb{R}^n$ such that $\lambda_i\ge 0$ for $i=0,\ldots,m$ and*

$$(v_0,-\lambda_0)\in N\big((\bar{x},\varphi_0(\bar{x}));\operatorname{epi}\varphi_i\big),\quad v\in N(\bar{x};\Omega),\tag{6.12}$$

$$(v_i,-\lambda_i)\in N\big((\bar{x},0);\operatorname{epi}\varphi_i\big),\quad i=1,\ldots,m,\tag{6.13}$$

$$(v_i, -\lambda_i) \in N\big((\bar{x}, 0); \operatorname{gph} \varphi_i\big), \quad i = m+1, \ldots, m+r, \tag{6.14}$$

$$v + \sum_{i=0}^{m+r} v_i = 0. \tag{6.15}$$

If in addition the function φ_i is u.s.c. at $\bar{x}$ for some $i \in \{1, \ldots, m\}$ with $\varphi_i(\bar{x}) < 0$, then $\lambda_i = 0$. If this happens for all $i = 1, \ldots, m$, then we have the complementary slackness conditions for the inequality constraints

$$\lambda_i \varphi_i(\bar{x}) = 0, \quad i = 1, \ldots, m. \tag{6.16}$$

(ii) *Assume that the functions φ_i are Lipschitz continuous around $\bar{x}$ for all $i = 0, \ldots, m+r$. Then there are multipliers $(\lambda_0, \ldots, \lambda_{m+r}) \neq 0$ such that*

$$0 \in \sum_{i=0}^{m} \lambda_i \partial\varphi_i(\bar{x}) + \sum_{i=m+1}^{m+r} \lambda_i \Big[\partial\varphi_i(\bar{x}) \cup \partial(-\varphi_i)(\bar{x})\Big] + N(\bar{x}; \Omega), \tag{6.17}$$

$$\lambda_i \geq 0,\ i = 0, \ldots, m+r, \ \ \textit{and}\ \ \lambda_i \varphi_i(\bar{x}) = 0,\ i = 1, \ldots, m. \tag{6.18}$$

Proof. To justify (i), assume without loss of generality that $\varphi_0(\bar{x}) = 0$. Then it is easy to check that $(\bar{x}, 0)$ is a *locally extremal point* of the following system of locally closed sets in the product space $\mathbb{R}^n \times \mathbb{R}^{m+r+1}$:

$$\Omega_i := \Big\{(x, \alpha_0, \ldots, \alpha_{m+r})\Big|\ \alpha_i \geq \varphi_i(x)\Big\}, \quad i = 0, \ldots, m,$$

$$\Omega_i := \Big\{(x, \alpha_0, \ldots, \alpha_{m+r})\Big|\ \alpha_i = \varphi_i(x)\Big\}, \quad i = m+1, \ldots, m+r,$$

$$\Omega_{m+r+1} := \Omega \times \{0\}.$$

Applying the extremal principle of Theorem 2.3 immediately gives us the relationships in (6.12)–(6.15). It follows from Proposition 1.17 that $\lambda_i \geq 0$ for $i = 0, \ldots, m$. To finish the proof of (i), it remains to show that the *complementary slackness* conditions in (6.16) hold for each $i \in \{1, \ldots, m\}$ with $\varphi_i(\bar{x}) < 0$ provided that φ_i is u.s.c. at $\bar{x}$. Indeed, we get from this assumption that $\varphi_i(x) < 0$ for all x around $\bar{x}$, and so $(\bar{x}, 0)$ is an *interior point* of the epigraph of φ_i. Thus $N((\bar{x}, 0); \operatorname{epi} \varphi_i) = \{0\}$ and $(v_i, \lambda_i) = (0, 0)$ for such i.

Assertion (ii) easily follows from (i) due to Theorem 1.22, which shows that the normal cone to the epigraph of a locally Lipschitzian function φ_i is fully determined by the (basic) subdifferential of φ_i. In the case of $\operatorname{gph} \varphi_i$ for the equality constraints, we deal with the epigraph of either φ_i or $-\varphi_i$ scaled by the corresponding *nonnegative* multiplier λ_i due to Proposition 1.17. △

The necessary optimality conditions of Theorem 6.5 are given in the *nonqualified/Fritz John* form, which doesn't ensure that $\lambda_0 \neq 0$ for the multiplier associated with the cost function. However, it is not hard to deduce from them (or from the qualification conditions in the calculus rules employed in the proofs) appropriate *constraint qualifications* of the generalized Mangasarian-Fromovitz and other types, which yield $\lambda_0 = 1$; see, e.g., [523, Chapter 5] with the commentaries and references therein as well as the exercises in Section 6.4.

Observe that for standard nonlinear programs (6.11) with smooth functions φ_i and $\Omega = \mathbb{R}^n$, the necessary optimality conditions of Theorem 6.5(ii) agree with the classical *Lagrange multiplier rule*. However, it is not the case for problems with nonsmooth *equality constraints*. Indeed, in the latter case, the result obtained in Theorem 6.5(ii) involves *nonnegative multipliers* λ_i associated with the the unions $\partial\varphi_i(\bar{x}) \cup \partial(-\varphi_i)(\bar{x})$ as $i = m+1, \ldots, m+r$, which are $\{\nabla\varphi_i(\bar{x}), -\nabla\varphi_i(\bar{x})\}$ for smooth functions. It is not hard to deduce from (6.17) and (6.18) a more conventional form of the generalized Lagrange multiplier rule with no sign condition for the equality multipliers, but in this way we arrive at a weaker necessary optimality condition as shown in Example 6.7 below. To proceed, recall the two-sided version of the basic subdifferential

$$\partial^0\varphi(\bar{x}) = \partial\varphi(x) \cup \partial^+\varphi(\bar{x}),$$

which is the *symmetric subdifferential* (1.75) already used in the book.

Corollary 6.6 (Equality Constraints via Symmetric Subgradients). *Let $\bar{x}$ be a local minimizer of* (6.11) *under the assumptions of Theorem* 6.5(ii). *Then there exists a nonzero collection of multipliers $(\lambda_0, \ldots, \lambda_{m+r}) \in \mathbb{R}^{m+r+1}$ satisfying the sign conditions $\lambda_i \geq 0$ for $i = 0, \ldots, m$, the complementary slackness condition* (6.16)*, and the symmetric Lagrangian inclusion*

$$0 \in \sum_{i=0}^{m} \lambda_i \partial\varphi_i(\bar{x}) + \sum_{i=m+1}^{m+r} \lambda_i \partial^0\varphi_i(\bar{x}) + N(\bar{x}; \Omega). \tag{6.19}$$

Proof. Follows directly from Theorem 6.5(ii) due to the (proper) inclusion

$$|\lambda|\big[\partial\varphi(\bar{x}) \cup \partial(-\varphi)(\bar{x})\big] \subset \lambda\big[\partial^0\varphi(\bar{x}) \cup \big(-\partial^0\varphi(\bar{x})\big)\big], \quad \lambda \in \mathbb{R},$$

applied to the functions φ_i, $i = m+1, \ldots, m+r$, in (6.17). △

6.1.3 Examples and Discussions on Optimality Conditions

Now we present several examples illustrating the difference between the obtained versions of the generalized Lagrange multiplier rule and compare them with other major versions known in nonsmooth optimization.

Example 6.7 (Nonnegative Sign vs. Symmetric Lagrangian Inclusions). As shown above, inclusion (6.17) with all the nonnegative multipliers always implies the symmetric one (6.19) with $\lambda_i \in \mathbb{R}$ as $i = m+1, \ldots, m+r$. The following two-dimensional problem with a single equality constraint confirms that the converse implication doesn't hold. Consider the problem:

$$\text{minimize } \ x_1 \ \text{ subject to } \ \varphi_1(x_1, x_2) := \varphi(x_1, x_2) + x_1 = 0, \tag{6.20}$$

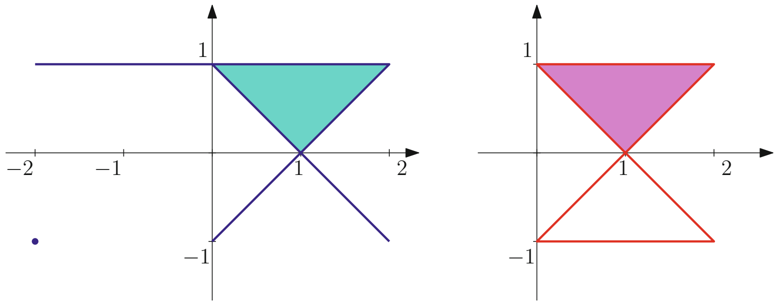
where φ is taken from Example 1.31(ii). It follows from the subdifferential calculation therein that the set $\partial\varphi_1(0,0) \cup \partial(-\varphi_1)(0,0)$ in (6.17) is

$$\big\{(v_1, v_2) \in \mathbb{R}^2 \big|\ |v_1 - 1| \le v_2 \le 1\big\} \cup \big\{(v_1, -|v_1 - 1|)\big| 0 \le v_1 \le 2\big\}$$
$$\cup\{(v_1, 1)\big| -2 \le v_1 \le 0\big\} \cup \big\{(-2, -1)\big\}$$

as depicted on Fig. 6.1(a). The symmetric subdifferential of φ_1 is

$$\partial^0\varphi_1(0,0) = \partial\varphi(0,0) \cup \big\{(v, -1)\big| -1 \le v \le 1\big\} + (1, 0)$$

with $\partial\varphi(0,0)$ calculated in Example 1.31(ii); see Fig. 6.1(b). It is now easy to check that the nonnegative sign inclusion (6.17) allows us to exclude the feasible solution $\bar{x} = (0,0)$ from the candidates for optimality, while the symmetric one (6.19) is satisfied at the nonoptimal point $\bar{x}$.



(a) $\partial\varphi_1(0,0) \cup \partial(-\varphi_1)(0,0)$

(b) $\partial^0\varphi_1(0,0)$

Fig. 6.1 Subdifferentials of $\varphi_1(x_1, x_2) = \big||x_1| + x_2\big| + x_1$ at (0, 0).

Example 6.8 (Comparison with the Convexified/Clarke Version of the Lagrange Multiplier Rule). Clarke's version [164, 165] of the Lagrange multiplier rule for nondifferentiable programming (6.11) with Lipschitzian data is given in the form of Corollary 6.6 where the nonconvex subdifferentials $\partial\varphi_i(\bar{x})$ for $i = 0, \ldots, m$ and $\partial^0\varphi_i(\bar{x})$ for $i = m+1, \ldots, m+r$, as well as the normal cone $N(\bar{x}; \Omega)$, are replaced by their convexified counterparts:

$$0 \in \sum_{i=0}^{m+r} \lambda_i \overline{\partial}\varphi_i(\bar{x}) + \overline{N}(\bar{x}; \Omega). \tag{6.21}$$

This version is obviously weaker than (6.6) and doesn't allow us to exclude the nonoptimal solution $\bar{x}$ in problem (6.20) of the preceding Example 6.7. Moreover, Clarke's version (6.21) fails to recognize nonoptimal solutions even in much less sophisticated examples from unconstrained nonsmooth optimization and also for problems with only inequality constraints. One of the reasons for this is that, due to the plus-minus symmetry of $\overline{\partial}\varphi$, condition (6.21) does *not* distinguish between *minima* and *maxima* and also between *inequality* constraints of the "$\le$" and "$\ge$" types. It makes an easy task to construct examples for which (6.21) is satisfied at clearly nonoptimal points.

(i) First consider the simplest *unconstrained* minimization problem:

$$\text{minimize} \;\; \varphi(x) := -|x| \;\; \text{over all} \;\; x \in \mathbb{R},$$

where $\bar{x} = 0$ is a point of maximum, not minimum. Nevertheless, we have $0 \in \overline{\partial}\varphi(0) = [-1, 1]$ while $0 \notin \partial\varphi(0) = \{-1, 1\}$.

(ii) The second example in this direction concerns the following two-dimensional problem with a single nonsmooth *inequality constraint*:

$$\text{minimize} \;\; x_1 \;\; \text{subject to} \;\; \varphi(x_1, x_2) := |x_1| - |x_2| \le 0.$$

We have here $\partial\varphi(0,0) = \{(v_1, v_2) |\, -1 \le v_1 \le 1, v_2 = 1, \text{ or } v_2 = -1\}$ by Example 1.31(i), and hence the point $\bar{x} = (0,0)$ is ruled out from optimality by Corollary 6.6, while the usage of the generalized gradient $\overline{\partial}\varphi(0,0) = \{(v_1, v_2) |\, -1 \le v_1 \le 1, -1 \le v_2 \le 1\}$ doesn't allow us to do it by (6.21).

Example 6.9 (Comparison with Warga's Version of the Lagrange Multiplier Rule). Another extension of the Lagrange multiplier rule to problems of nondifferentiable programming (6.11) with $\Omega = \mathbb{R}^n$ and Lipschitzian functions φ_i was obtained by Warga [736, 737] in terms of his *derivate containers* $\Lambda^0\varphi_i(\bar{x})$ in the form of Corollary 6.6 with the Lagrangian inclusion

$$0 \in \sum_{i=0}^{m+r} \lambda_i \Lambda^0 \partial\varphi_i(\bar{x}). \tag{6.22}$$

Note that the set $\Lambda^0\varphi(\bar{x})$ is generally nonconvex, possesses the classical plus-minus symmetry, and may be strictly smaller than Clarke's generalized gradient $\overline{\partial}\varphi(\bar{x})$. As shown in [522, Corollary 2.48], we always have $\partial^0\varphi(\bar{x}) \subset \Lambda^0\varphi(\bar{x})$. Hence the necessary optimality conditions of Theorem 6.5(ii) and Corollary 6.6 definitely yield the result of (6.22). Let us illustrate that the improvement is *strict* in both cases of equality and inequality constraints.

(i) For the case of only *equality* constraints in (6.11), the claimed strict inclusion follows from Example 6.7 with the constraint function φ_1 defined in (6.20). Indeed, condition (6.22) is satisfied at the nonoptimal point $\bar{x} = (0,0)$, while (6.19) confirms its nonoptimality. Recall that the derivative container $\Lambda^0\varphi(\bar{x})$ for the function φ in this example is depicted on Fig. 1.13.

(ii) To demonstrate the advantage of (6.17) for nondifferentiable programs with *inequality* constraints, consider the problem

$$\text{minimize} \;\; x_2 \;\; \text{subject to} \;\; \varphi_1(x_1, x_2) := \varphi(x_1, x_2) + x_2 \le 0,$$

where φ is taken from Example 1.31(ii) and its subdifferential $\partial\varphi(0,0)$ is calculated therein. Hence we have

$$\partial\varphi_1(0,0) = \left\{(v_1, v_2)\,\middle|\; |v_1| + 1 \le v_2 \le 2\right\} \cup \left\{(v_1, v_2)\,\middle|\; 0 \le v_2 = -|v_1| + 1\right\}$$

as depicted on Fig. 6.2. This shows that the result of Theorem 6.5(ii) (same in Corollary 6.6) allows us to rule out the nonoptimal point $\bar{x} = (0,0)$, while it cannot be done by using Warga's condition (6.22).

Next we derive yet another type of lower subdifferential optimality conditions for problem (6.11) with Lipschitzian data that are expressed in the *condensed form* via the basic subdifferential (1.24) of Lagrangian combinations of the initial data. Consider the standard *Lagrangian*

$$\mathcal{L}(x, \lambda_0, \ldots, \lambda_{m+r}) := \lambda_0\varphi_0(x) + \ldots + \lambda_{m+r}\varphi_{m+r}(x)$$

involving the cost function and all the functional (while not geometric) constraints and also the *extended Lagrangian*

$$\mathcal{L}_\Omega(x; \lambda_0, \ldots, \lambda_{m+r}) := \lambda_0\varphi_0(x) + \ldots + \lambda_{m+r}\varphi_{m+r}(x) + \delta(x; \Omega)$$

involving also the set geometric constraint via its indicator function.

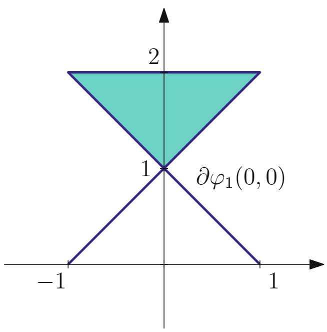


Fig. 6.2 Basic subdifferential of $\varphi_1(x_1, x_2) = \big||x_1| + x_2\big| + x_2$ at $(0,0)$.

Theorem 6.10 (Condensed Lower Subdifferential Optimality Conditions). *Let $\bar{x}$ be a local minimizer of problem* (6.11) *under the assumptions of Theorem* 6.5(ii). *Then there are multipliers $\lambda_0, \ldots, \lambda_{m+r}$, not equal to zero simultaneously, satisfying* (6.16) *and the condensed Lagrangian inclusions*

$$0 \in \partial_x\mathcal{L}_\Omega(\bar{x}, \lambda_0, \ldots, \lambda_{m+r}) \subset \partial_x\mathcal{L}(\bar{x}, \lambda_0, \ldots, \lambda_{m+r}) + N(\bar{x}; \Omega). \tag{6.23}$$

Proof. Note that the second inclusion in (6.23) follows from the first one due to the subdifferential sum rule from Corollary 2.20. To justify the first inclusion therein, consider the set

$$\mathcal{E}(\varphi_0,\ldots,\varphi_{m+r},\Omega) := \Big\{ (x,\alpha_0,\ldots,\alpha_{m+r}) \in \mathbb{R}^{n+m+r+1} \Big| x \in \Omega, \varphi_i(x) \le \alpha_i, \; i=0,\ldots,m;\; \varphi_i(x)=\alpha_i,\; i=m+1,\ldots,m+r \Big\}$$

and suppose without loss of generality that $\varphi_0(\bar{x}) = 0$. Denoting now by U a neighborhood of the local minimizer $\bar{x}$ in (6.11), we claim that the pair $(\bar{x},0) \in \mathbb{R}^n \times \mathbb{R}^{m+r+1}$ is an *extremal point* of the closed set system

$$\Omega_1 := \mathcal{E}(\varphi_0,\ldots,\varphi_{m+r},\Omega) \;\text{ and }\; \Omega_2 := \operatorname{cl} U \times \{0\}. \tag{6.24}$$

Indeed, we obviously have $(\bar{x},0) \in \Omega_1 \cap \Omega_2$ and $(\Omega_1 - (0,\nu_k,0,\ldots,0)) \cap \Omega_2 = \emptyset$, $k \in \mathbb{N}$, for any sequence of negative numbers $\nu_k \uparrow 0$ by the local optimality of $\bar{x}$ in (6.11). Applying to this system the basic *extremal principle* from Theorem 2.3 gives us multipliers $(\lambda_0,\ldots,\lambda_{m+r}) \neq 0$ satisfying the inclusion

$$(0,-\lambda_0,\ldots,-\lambda_{m+r}) \in N\big((\bar{x},0);\mathcal{E}(\varphi_0,\ldots,\varphi_{m+r},\Omega)\big), \tag{6.25}$$

which implies the conditions in (6.16) due to the structure of the set Ω_1 in (6.24). Furthermore, it follows from the scalarization formula of Theorem 1.32 and its proof that (6.25) can be equivalently rewritten as the first inclusion in (6.23) under the assumed local Lipschitz continuity of φ_i. △

If the geometric constraint set Ω is *convex*, the second inclusion in (6.23) can be written in the form of the *abstract maximum principle*.

Corollary 6.11 (Abstract Maximum Principle in Nondifferentiable Programming). *Suppose that the set Ω is convex in the assumptions of Theorem 6.10. Then there are multipliers $(\lambda_0,\ldots,\lambda_{m+r}) \neq 0$ such that*

$$\langle v,\bar{x}\rangle = \max_{x\in\Omega}\langle v,x\rangle \;\; \textit{for some} \;\; v \in -\partial_x \mathcal{L}(\bar{x},\lambda_0,\ldots,\lambda_{m_r}).$$

Proof. It follows from Theorem 6.10 by the representation of the normal cone to convex sets given in Proposition 1.7. △

We conclude this section by deriving *upper subdifferential* necessary optimality conditions for (6.11) that are independent of the obtained lower subdifferential conditions; see more discussions in Remark 6.2.

Theorem 6.12 (Upper Subdifferential Optimality Conditions in Nondifferentiable Programming). *Let $\bar{x}$ be a local minimizer of problem (6.11). Assume that the functions φ_i are locally Lipschitzian around $\bar{x}$ for the equality indices $i = m+1,\ldots,m+r$. Then for any $v_i \in \widehat{\partial}^+\varphi_i(\bar{x})$, $i=0,\ldots,m$, there are multipliers $(\lambda_0,\ldots,\lambda_{m+r}) \neq 0$ satisfying (6.16) and the inclusion*

$$-\sum_{i=0}^{m} \lambda_i v_i \in \partial\Big(\sum_{i=m+1}^{m+r} \lambda_i \varphi_i\Big)(\bar{x}) + N(\bar{x};\Omega). \tag{6.26}$$

Proof. Suppose without loss of generality that $\widehat{\partial}^+\varphi_i(\bar{x}) \neq \emptyset$ for $i = 0,\ldots,m$. Applying the second part of Theorem 1.27 to $-v_i \in \widehat{\partial}(-\varphi_i)(\bar{x})$ (we can always assume that the functions $-\varphi_i$ are bounded from below, which is actually not needed for the localized version of Theorem 1.27 used in what follows) allows us to find functions $\psi_i\colon \mathbb{R}^n \to \mathbb{R}$ for $i = 0,\ldots,m$ satisfying

$$\psi_i(\bar{x}) = \varphi_i(\bar{x}) \;\text{ and }\; \psi_i(x) \geq \varphi_i(x) \;\text{ around }\; \bar{x}$$

and such that each $\psi_i(x)$ is continuously differentiable around $\bar{x}$ with the gradient $\nabla\psi_i(\bar{x}) = v_i$. It is easy to check that $\bar{x}$ is a local solution to the following optimization problem of type (6.11) but with the cost and inequality constraint functions continuously differentiable around $\bar{x}$:

$$\begin{cases} \text{minimize } \; \psi_0(x) \quad \text{subject to} \\ \psi_i(x) \leq 0, \quad i = 1,\ldots,m, \\ \varphi_i(x) = 0, \quad i = m+1,\ldots,m+r, \\ x \in \Omega \subset \mathbb{R}^n. \end{cases} \tag{6.27}$$

To arrive finally at (6.26), it remains to apply to the solution $\bar{x}$ of (6.27) the second Lagrangian inclusion in (6.23) of Theorem 6.10 and then to use therein the elementary subdifferential sum rule from Proposition 1.30(ii). △

Employing further in (6.26) the subdifferential sum rule for Lipschitzian functions from Corollary 2.20 and weakening in this way the necessary optimality conditions for the case of equality constraints, we can express them in forms (6.17) and (6.19) via the corresponding subdifferential constructions for the separate functions φ_i, $i = m+1,\ldots,m+r$.

6.2 Problems of Bilevel Programming

In this section we begin considering a remarkable class of problems in hierarchical optimization known as *bilevel programming* and also as *Stackelberg games*. Such problems are highly interesting and challenging in optimization theory and important for numerous applications. There is an enormous bibliography on bilevel programming and related topics; see commentaries and references in Section 6.5 for more discussions on major approaches and results.

Our primary goal here is to reduce bilevel programs to those in nondifferentiable programming considered above and derive in this way several types of necessary optimality conditions in terms of the initial bilevel data by using the results of Section 6.1 together with subdifferentiation of marginal functions and other machinery of variational analysis.

6.2.1 Optimistic and Pessimistic Versions

Bilevel programming deals with problems of hierarchical optimization that address minimizing a given *upper-level/leader's* objective function $f(x, y)$ from $\mathbb{R}^n \times \mathbb{R}^m$ to $\mathbb{R}$ subject to the upper-level constraints $x \in \Omega \subset \mathbb{R}^n$ along an optimal solution $y = y(x)$ to the parametric *lower-level/follower's* problem

$$\text{minimize}_y \ \varphi(x, y) \ \text{ subject to } \ y \in G(x) \tag{6.28}$$

with the objective/cost $\varphi\colon \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ and the constraint set-valued mapping $G\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$. For simplicity we confine ourselves to the case where the lower-level constraints are given by the parameterized inequality systems

$$G(x) := \big\{ y \in \mathbb{R}^m \big|\ g(x, y) \le 0 \big\}, \tag{6.29}$$

where $g = (g_1, \ldots, g_p)\colon \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p$ and the vector inequality for g are understood componentwise. As follows from the proofs below, appropriately modified similar results can be derived for other types of constraints in (6.28).

Note that the bilevel optimization problem formulated above is not fully determined when the *solution/argminimum map*

$$S(x) := \text{argmin}\big\{\varphi(x, y) \big|\ y \in G(x)\big\}, \quad x \in \mathbb{R}^n, \tag{6.30}$$

for the lower-level problem is set-valued, since in this case we did not specify how to choose a single-valued decision function $y(x)$. To deal with such a typical situation, the two major versions, known as optimistic and pessimistic models, have been designated in bilevel programming. We always suppose that the argminimum sets $S(x)$ are nonempty around the reference point.

The *optimistic* version in bilevel programming is formulated as follows:

$$\begin{aligned} &\text{minimize } f_{opt}(x) \ \text{ subject to } \ x \in \Omega, \\ &\text{where } \ f_{opt}(x) := \inf\big\{ f(x, y) \big|\ y \in S(x) \big\}, \end{aligned} \tag{6.31}$$

which means that the decision $y(x)$ is chosen in $S(x)$ to benefit the objective f_{opt}. As usual, a point $\bar{x} \in \Omega$ is called a global (local) optimistic solution to (6.31) if $f_{opt}(\bar{x}) \le f_{opt}(x)$ for all $x \in \Omega$ (sufficiently close to $\bar{x}$). From the economics viewpoint , this corresponds to a situation where the follower participates in the profit of the leader, i.e., there exists some cooperation between both players on the upper and lower levels.

However, it would not always be possible for the leader to convince the follower to make choices that are favorable for him or her. Hence it is necessary for the upper-level player to reduce damages resulting from undesirable selections on the lower level. This brings us to the *pessimistic* version in bilevel programming formulated in the following way:

$$\begin{aligned} &\text{minimize } f_{pes}(x) \text{ subject to } x \in \Omega, \\ &\text{where } f_{pes}(x) := \sup\big\{f(x,y)\big|\ y \in S(x)\big\}. \end{aligned} \tag{6.32}$$

We can see that (6.32) is a special type of *minimax* problems, which challenges come from the complicated structure of the moving set $S(x)$ as the solution set to the lower-level optimization problem.

Our main attention in this chapter is paid to the *optimistic version*, although we'll present some comments on the pessimistic version as well. Further, we'll discuss in the exercise and commentary sections of this chapter a *multiobjective* approach to problems of bilevel programming that can be applied to both optimistic and pessimistic versions by reducing them to constrained multiobjective optimization problems studied in Chapter 9.

6.2.2 Value Function Approach

There are several approaches to optimistic bilevel programs known in the literature; see Section 6.5 for more discussions and references. We concentrate here on the so-called *value function approach*, which involves the optimal value function of the lower-level problem (6.28) defined by

$$\mu(x) := \inf\big\{\varphi(x,y)\big|\ y \in G(x)\big\}, \quad x \in \mathbb{R}^n, \tag{6.33}$$

and provides a reformulation of the bilevel problem (6.31) in the form

$$\begin{aligned} &\text{minimize } f(x,y) \text{ subject to } x \in \Omega, \\ &g(x,y) \le 0, \text{ and } \varphi(x,y) \le \mu(x). \end{aligned} \tag{6.34}$$

It is easy to see that problem (6.34) is *globally* equivalent to the original optimistic bilevel program (6.31). The next proposition reveals relationships between *local* solutions to these problems. To give its exact formulation and proof, we introduce the *two-level value function*

$$\eta(x) := \inf\big\{f(x,y)\big|\ g(x,y) \le 0,\ \varphi(x,y) \le \mu(x)\big\}, \quad x \in \mathbb{R}^n, \tag{6.35}$$

and then define the corresponding modification of the solution map (6.30) by

$$\widetilde{S}(x) := \operatorname{argmin}\big\{\varphi(x,y)\big|\ g(x,y) \le 0,\ f(x,y) \le \eta(x)\big\}. \tag{6.36}$$

We obviously have $\widetilde{S}(x) \subset S(x)$ for all $x \in \mathbb{R}^n$.

Proposition 6.13 (Local Optimal Solutions to Optimistic Bilevel Programs). *Let $\widetilde{S}(x)$ be defined in* (6.36). *The following assertions hold:*

(i) *If $\bar{x}$ is a local optimal solution to* (6.31)*, then for any $\bar{y} \in \widetilde{S}(\bar{x})$, the pair $(\bar{x},\bar{y})$ is a local optimal solution to problem* (6.34).

(ii) *Conversely, let* $(\bar{x}, \bar{y})$ *be a local optimal solution to* (6.34) *for some* $\bar{y} \in \widetilde{S}(\bar{x})$, *and let the mapping* $\widetilde{S}$ *be inner semicontinuous at* $(\bar{x}, \bar{y})$. *Then* $\bar{x}$ *is a local optimal solution to the original optimistic bilevel problem* (6.31).

Proof. We verify (i) arguing by contradiction. Suppose that $(\bar{x}, \bar{y})$ with some $\bar{y} \in \widetilde{S}(\bar{x})$ is not a local optimal solution to (6.34). Then we find a sequence of (x_k, y_k) with $x_k \to \bar{x}$, $y_k \to \bar{y}$ so that $x_k \in \Omega$, $g(x_k, y_k) \le 0$, $\varphi(x_k, y_k) \le \mu(x_k)$, and $f(x_k, y_k) < f(\bar{x}, \bar{y}) = \eta(\bar{x})$ for all $k \in \mathbb{N}$. It follows from the construction of $\eta(\cdot)$ in (6.35) that $\eta(x_k) \le f(x_k, y_k)$. This shows that $f_{opt}(x_k) < f_{opt}(\bar{x})$, which contradicts the local optimality of $\bar{x}$ in (6.31).

To justify (ii), suppose that $\bar{x}$ is not a local optimal solution to (6.31) while the assumptions in (ii) are satisfied. Then we find a sequence $x_k \to \bar{x}$ with $x_k \in \Omega$ such that $f_{opt}(x_k) < f_{opt}(\bar{x})$ for all k. Since $\widetilde{S}$ is inner semicontinuous at $(\bar{x}, \bar{y})$, there is a sequence of $y_k \in \widetilde{S}(x_k)$ with $y_k \to \bar{y}$. This implies by (6.36) that $\varphi(x_k, y_k) = \mu(x_k)$, $g(x_k, y_k) \le 0$, and $f(x_k, y_k) < f(\bar{x}, \bar{y})$, which contradicts the local optimality of $(\bar{x}, \bar{y})$ in (6.34). △

The obtained results (see also Exercise 6.36) allow us to adequately replace the original optimistic bilevel problem (6.31) by the problem of constrained optimization (6.34) of the type considered in Section 6.1 and derive necessary optimality conditions for (6.31) from those for (6.34). Observe to this end that problem (6.34) is written in form (6.11) of nonlinear programming without equality constraints, where the inequality constraint

$$\varphi(x, y) - \mu(x) \le 0 \tag{6.37}$$

unavoidably involves the *nondifferentiable* function $\mu(x)$ of the marginal type (4.1) the generalized differential properties of which were studied in Section 4.1. Note however that the designated constraint (6.37) contains the term $-\mu(x)$, different from $\mu(x)$ in generalized differentiation, and that the constraint mapping G in (6.33) is given in the particular form (6.29).

It turns out that, even in the case where the upper-level constraint set Ω reduces to the whole space $\mathbb{R}^n$ or it is described by smooth inequalities, the usual Mangasarian-Fromovitz and other standard constraint qualifications as well as their natural extensions are *violated*; see more in Section 6.5.

6.2.3 Partial Calmness and Weak Sharp Minima

To overcome these difficulties, we present a qualification condition of another type that allows us to incorporate the troublesome constraint (6.37) into a *penalized* cost function and deal with it by using appropriate calculus rules of generalized differentiation. Consider a perturbed version of (6.34) with the linear parameterization of constraint (6.37) defined as follows:

$$\begin{aligned}&\text{minimize } f(x, y) \ \text{ subject to } x \in \Omega, \ \ g(x, y) \le 0,\\ &\text{and } \ \varphi(x, y) - \mu(x) + \vartheta = 0, \quad \ \vartheta \in \mathbb{R}.\end{aligned} \tag{6.38}$$

Definition 6.14 (Partial Calmness). *The unperturbed problem* (6.34) *is* PARTIALLY CALM *at its feasible solution* $(\bar{x}, \bar{y})$ *if there exist a constant* $\kappa > 0$ *and a neighborhood* U *of the triple* $(\bar{x}, \bar{y}, 0) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}$ *such that*

$$f(x, y) - f(\bar{x}, \bar{y}) + \kappa|\vartheta| \geq 0 \tag{6.39}$$

for all $(x, y, \vartheta) \in U$ *feasible to* (6.38).

The next result reveals the role of partial calmness in bilevel programming.

Proposition 6.15 (Penalization via Partial Calmness). *Let* $(\bar{x}, \bar{y})$ *be a partially calm feasible solution to problem* (6.34)*, and let* f *be continuous at this point. Then* $(\bar{x}, \bar{y})$ *is a local optimal solution to the penalized problem*

$$\begin{aligned} &\textit{minimize} \;\; f(x, y) + \kappa\big(\varphi(x, y) - \mu(x)\big) \\ &\textit{subject to} \;\; x \in \Omega \;\; \textit{and} \;\; g(x, y) \leq 0, \end{aligned} \tag{6.40}$$

where the constant κ *is taken from* (6.39)*. Conversely, any local optimal solution* $(\bar{x}, \bar{y})$ *to* (6.40) *with some number* $\kappa > 0$ *is partially calm in* (6.34).

Proof. By the assumed partial calmness, we get κ and U for which (6.39) holds. It follows from the continuity of f at $(\bar{x}, \bar{y})$ that there are numbers $\gamma > 0$ and $\eta > 0$ such that $V := [(\bar{x}, \bar{y}) + \eta\mathbb{B}] \times (-\gamma, \gamma) \subset U$ and that

$$|f(x, y) - f(\bar{x}, \bar{y})| \leq \kappa\gamma \;\; \text{whenever} \;\; (x, y) - (\bar{x}, \bar{y}) \in \eta\mathbb{B}.$$

This allows us to establish the relationship

$$f(x, y) - f(\bar{x}, \bar{y}) + \kappa\big(\varphi(x, y) - \mu(x)\big) \geq 0 \tag{6.41}$$

whenever $(x, y) \in [(\bar{x}, \bar{y}) + \eta\mathbb{B}] \cap \operatorname{gph} G$ with G defined in (6.29) and $x \in \Omega$. Indeed, for $(x, y, \mu(x) - \varphi(x, y)) \in V$, we deduce (6.41) directly from (6.39). If otherwise $(x, y, \mu(x) - \varphi(x, y)) \notin V$, it follows that

$$\varphi(x, y) - \mu(x) \geq \gamma \;\; \text{and so} \;\; \kappa\big(\varphi(x, y) - \mu(x)\big) \geq \kappa\gamma.$$

This also implies (6.41) due to $f(x, y) - f(\bar{x}, \bar{y}) \geq -\kappa\gamma$. To complete the proof of the first assertion of the proposition, observe that $\varphi(\bar{x}, \bar{y}) - \mu(\bar{x}) = 0$ since $(\bar{x}, \bar{y})$ is a feasible solution to (6.34). The converse statement follows directly from the definitions while arguing by contradiction. △

It is easy to see that a verifiable sufficient condition for the desired partial calmness is provided by the following notion of local weak sharp minima, which has been well recognized in qualitative and numerical aspects of optimization.

Definition 6.16 (Local Weak Sharp Minima). *Given* $Q \subset \mathbb{R}^s$*, we say that* $P \subset Q$ *is a set of* (LOCAL) WEAK SHARP MINIMA *for a function* $\phi\colon \mathbb{R}^s \to \mathbb{R}$ *over* Q *at* $\bar{z} \in P$ *with modulus* $\alpha > 0$ *if*

$$\phi(z) \geq \phi(\bar{z}) + \alpha \operatorname{dist}(z; P) \;\; \textit{for all} \;\; z \in Q \;\; \textit{near} \;\; \bar{z}. \tag{6.42}$$

The next proposition presents the precise formulation and provides a simple proof of the result needed in what follows with some *uniformity* in (6.42).

Proposition 6.17 (Partial Calmness from Uniform Weak Sharp Minima). *Let $(\bar{x}, \bar{y})$ be a local optimal solution to the bilevel program (6.34) such that we have the* UNIFORM WEAK SHARP MINIMUM *condition:*

$$\varphi(x, y) - \mu(x) \geq \alpha \operatorname{dist}\big(y; S(x)\big) \ \textit{with some} \ \alpha > 0 \tag{6.43}$$

for all (x, y) near $(\bar{x}, \bar{y})$ with $x \in \Omega$ and $y \in G(x)$. Assume that f is locally Lipschitzian around $(\bar{x}, \bar{y})$. Then problem (6.34) is partially calm at $(\bar{x}, \bar{y})$.

Proof. Picking any triple (x, y, ϑ) feasible to problem (6.38) and sufficiently close to $(\bar{x}, \bar{y}, 0)$, we have $x \in \Omega$, $y \in G(x)$, and $\varphi(x, y) - \mu(x) + \vartheta = 0$, where $|\vartheta|$ is small enough. Using assumption (6.43) gives us some $v \in S(x)$ with

$$\varphi(x, y) - \mu(x) \geq \frac{\alpha}{2}\|y - v\| \geq 0.$$

Since $(\bar{x}, \bar{y})$ is a local optimal solution to (6.34), we get

$$\begin{aligned} & f(x, y) - f(\bar{x}, \bar{y}) \geq f(x, y) - f(x, v) \geq -\ell\|y - v\| \\ & \geq -\frac{2\ell}{\alpha}\Big(\varphi(x, y) - \mu(x)\Big) = -\kappa|\vartheta| \end{aligned}$$

with $\kappa := 2\ell/\alpha$, where $\ell > 0$ is a Lipschitz constant of f around $(\bar{x}, \bar{y})$. This justifies the partial calmness condition (6.39). △

Note that assumption (6.43) corresponds to the local weak sharp minimum condition of Definition 6.16 at $\bar{z} = (\bar{x}, \bar{y})$ with respect to y for any fixed feasible vector x with the following data:

$$z := (x, y), \ \phi(z) := \varphi(x, y), \ P := S(x), \ \text{ and } \ Q := G(x). \tag{6.44}$$

Observe also that the uniform weak sharpness in (6.43) requires that the constant $\alpha > 0$ therein can be selected *uniformly* in x. Proceeding in this way and deriving, in particular, sufficient conditions for (6.42) that being applied to (6.43) are independent of x, would allow us to decrease serious difficulties in dealing with nonsmooth marginal function (6.33) in the value function approach to optimistic bilevel programs.

Let us now present an easily verifiable condition of this type for weak sharp minimizers in nonlinear programming, which is of its own interest while being useful bilevel optimization; see below more discussions in this vein.

Proposition 6.18 (Sufficient Conditions for Weak Sharp Minima). *Let $\bar{z}$ be a local optimal solution to the nonlinear program:*

$$\textit{minimize} \ \ \phi(z) \ \ \textit{subject to} \ \ \psi_i(z) \leq 0 \ \textit{for} \ i = 1, \ldots, p, \tag{6.45}$$

where the functions $\phi, \psi_i \colon \mathbb{R}^s \to \mathbb{R}$ *as* $i \in I(\bar z) := \{i \mid \psi_i(\bar z) = 0\}$ *are Fréchet differentiable at* $\bar z$*. Suppose that necessary optimality conditions for* $\bar z$ *hold in the qualified Karush-Kuhn-Tucker form*

$$\nabla\phi(\bar z) + \sum_{i \in I(\bar z)} \lambda_i \nabla\psi_i(\bar z) = 0 \text{ for some } \lambda_i \ge 0$$

and that the following kernel condition

$$\bigcap_{i \in J} \ker \nabla\psi_i(\bar z) = \{0\} \text{ with } J := \big\{i \,\big|\, \lambda_i > 0\big\} \tag{6.46}$$

is satisfied. Then there exists a positive constant α *such that*

$$\phi(z) - \phi(\bar z) \ge \alpha\|z - \bar z\| \text{ if } \psi_i(z) \le 0 \text{ for } i = 1, \ldots, p \tag{6.47}$$

whenever z *is sufficiently close to* $\bar z$*. Consequently,* ϕ *admits a set of local weak sharp minima over* $Q := \{z \in \mathbb{R}^s \mid \psi_i(z) \le 0,\ i = 1, \ldots, p\big\}$ *at* $\bar z$*.*

Proof. To justify (6.47) with some $\alpha > 0$, suppose on the contrary that there exists a sequence $\{z_k\} \subset Q$ with $z_k \ne \bar z$ and $z_k \to \bar z$ such that

$$\phi(z_k) - \phi(\bar z) \le \frac{1}{k}\|z_k - \bar z\| \text{ for all } k \in \mathbb{N}. \tag{6.48}$$

Let $d_k := \frac{z_k - \bar z}{\|z_k - \bar z\|}$ and without loss of generality assume that $d_k \to d$ as $k \to \infty$ with $\|d\| = 1$. It follows from (6.48) by the (Fréchet) differentiability of ϕ at $\bar z$ that $\langle \nabla\phi(\bar z), d\rangle \le 0$. On the other hand, the assumed differentiability of the active constraint functions at $\bar z$ ensures that

$$\langle \nabla\psi_i(\bar z), d\rangle \le 0 \text{ for all } i \in I(\bar z).$$

Using the last two inequalities and the imposed KKT condition tells us that

$$0 \le -\langle \nabla\phi(\bar z), d\rangle = \sum_{i \in J} \lambda_i \langle \nabla\psi_i(\bar z), d\rangle \le 0,$$

which yields $\langle \nabla\psi_i(\bar z), d\rangle = 0$ for all $i \in J$. Thus we have $d = 0$ by the kernel condition (6.46), a contradiction that completes the proof. △

Observe that the *kernel condition* (6.46) is *essential* for Proposition 6.18 to hold. Indeed, consider problem (6.45) with $\phi, \psi \colon \mathbb{R}^2 \to \mathbb{R}$ defined by

$$\phi(z_1, z_2) := z_1^2 - z_2 \text{ and } \psi(z_1, z_2) := z_2.$$

Then $Q = \mathbb{R} \times \mathbb{R}_-$ and $\bar z := (0, 0)$ is the only solution to this problem. Since

$$\ker \nabla\psi(\bar z) = \mathbb{R} \times \{0\},$$

the kernel condition (6.46) is violated. It is easy to see that for any vector $z = (\gamma, 0) \in Q$ with $\gamma \neq 0$, we have the equalities

$$\phi(z) - \phi(\bar{z}) = \gamma^2 \text{ and } \|z - \bar{z}\| = \gamma,$$

which immediately imply that the conclusion in (6.47) doesn't hold, since the number $\gamma > 0$ can be chosen arbitrarily small.

Besides the presented conditions for weak sharp minima and their uniform counterparts, there are other sufficient conditions for these properties with various applications to partial calmness in bilevel programming and related topics; see more details in Sections 6.4 and 6.5. In particular, partial calmness is always satisfied for bilevel programs where lower-level problems are *linear* with respect to their *lower-level decision variables*; see Exercise 6.37(i).

The following examples illustrate some possibilities of verifying partial calmness in bilevel program via the results established above.

Example 6.19 (Verification of Partial Calmness via Penalization). Let us show that the penalty function characterization of partial calmness in Proposition 6.15 is a convenient tool to verify the validity or failure of partial calmness in bilevel programming. Consider first the *fully nonlinear*, at both lower and upper levels, bilevel program (6.34) with $(x, y) \in \mathbb{R}^2$, $\Omega = \mathbb{R}$, and

$$f(x, y) := \frac{(x-1)^2}{2} + \frac{y^2}{2}, \quad S(x) = \operatorname{argmin}\left\{\frac{x^2}{2} + \frac{y^2}{2}\right\}.$$

It is easy to see that $S(x) = \{0\}$ for all $x \in \mathbb{R}$ and that $\mu(x) = x^2/2$ for the lower-level value function in (6.33). Furthermore, the pair $(\bar{x}, \bar{y}) = (1, 0)$ is the only solution to the upper-level problem, and so it is an optimal solution to the bilevel program under consideration. We have $\varphi(x, y) - \mu(x) = y^2/2$, and hence the corresponding unconstrained penalized problem (6.40) is

$$\text{minimize } \frac{(x-1)^2}{2} + \frac{y^2}{2} + \kappa \frac{y^2}{2}$$

with no constraints on (x, y). Observe that for any $\kappa > 0$, the latter problem is smooth and strictly convex and has the unique optimal solution $(\bar{x}, \bar{y}) = (1, 0)$. Thus the initial bilevel program is partially calm at this point.

On the other hand, replacing the upper-level cost function $f(x, y)$ by

$$\frac{(x-1)^2}{2} + \frac{(y-1)^2}{2}$$

and keeping the same lower-level problem gives us the bilevel program (6.34) with the optimal solution $(\bar{x}, \bar{y}) = (1, 1)$, which fails to satisfy the partial calmness condition. Indeed, it is easy to see that the corresponding penalized problem (6.40) has the only optimal solution

$$\Big(1, \frac{1}{1+\kappa}\Big) \neq (1,1) \ \text{ whenever } \kappa > 0.$$

Example 6.20 (Verification of Partial Calmness via Uniform Weak Sharp Minima). Consider the constrained optimization problem in $\mathbb{R}^3$:

$$\text{minimize } \ \frac{x_1^2}{2} + \frac{x_2^2}{2} \ \text{ subject to } \ a_i \le x_i \le b_i, \quad i = 1, 2, 3. \tag{6.49}$$

It is not hard to check that optimal solutions to this problem constitute the set of weak sharp minima if either $a_i > 0$ or $b_i < 0$ for $i = 1, 2$; see Exercise 6.38(ii). Thus Proposition 6.17 tells us that any bilevel program having (6.49) as its lower-level problem with the above parameters a_i, b_i is partially calm at each of its local optimal solutions.

Note that Example 6.19 shows that partial calmness in bilevel programs may significantly *depend* on the structure of upper-level objectives. On the contrary, Example 6.20 describes a class of multidimensional bilevel programs where partial calmness holds *independently* of the upper level.

6.3 Bilevel Programs with Smooth and Lipschitzian Data

In this section we develop the value function approach to bilevel programming discussed above to obtain necessary optimality conditions in optimistic bilevel programs first with smooth and then with Lipschitzian initial data. For simplicity, consider here the bilevel program (6.31) in the value/marginal function form (6.34) with the upper-level constraint set Ω given by the inequalities

$$\Omega := \big\{x \in \mathbb{R}^n \,\big|\, h(x) \le 0\big\} \ \text{ with } \ h(x) = \big(h_1(x), \ldots, h_q(x)\big), \tag{6.50}$$

which are described by the real-valued functions h_j. Our major results are derived under the inner semicontinuity of the argminimum map S in (6.30) at the reference local optimal solution $(\bar{x}, \bar{y})$ by passing to problem (6.34) via Proposition 6.13. Imposing further the partial calmness of (6.34) at the given local solution $(\bar{x}, \bar{y})$ and using Proposition 6.15, we reduce (6.34) to the single-level programming form (6.40), which is essentially used in our proofs.

Observe that problem (6.40) with constraints (6.50) is written as

$$\begin{aligned} &\text{minimize } \ f(x,y) + \kappa\big(\varphi(x,y) - \mu(x)\big) \\ &\text{subject to } \ g(x,y) \le 0 \ \text{ and } \ h(x) \le 0 \end{aligned} \tag{6.51}$$

for some $\kappa > 0$. Hence it can be treated as a particular case of the mathematical program (6.11) with inequality constraints. The most essential specific features of (6.51) are *intrinsic nonsmoothness* of the marginal function $\mu(x)$ from (6.33), regardless of smoothness of the initial data, and the presence of function (6.33) in the objective of (6.51) with the *negative sign*. Nevertheless, the above subdifferential results for marginal functions and explicit representations of normals to sets described by inequality constraints allow us to efficiently proceed in deriving necessary optimality conditions for (6.34).

6.3.1 Optimality Conditions for Smooth Bilevel Programs

Given a feasible solution $(\bar{x}, \bar{y})$ to the original optimistic bilevel program (6.34) with the constraint set Ω defined in (6.50), denote by

$$I(\bar{x}, \bar{y}) := \big\{i \in \{1, \dots, p\}\big|\ g_i(\bar{x}, \bar{y}) = 0\big\},\ \ J(\bar{x}) := \big\{j \in \{1, \dots, q\}\big|\ h_j(\bar{x}) = 0\big\}$$

the collections of the corresponding active constraint indices. Considering first problem (6.34) with smooth initial data and following the traditional terminology in bilevel programming, we say that $(\bar{x}, \bar{y})$ is *lower-level regular* if for any nonnegative numbers λ_i the implication

$$\Big[\sum_{i \in I(\bar{x}, \bar{y})} \lambda_i \nabla_y g_i(\bar{x}, \bar{y}) = 0\Big] \Longrightarrow \Big[\lambda_i = 0 \ \text{ whenever } \ i \in I(\bar{x}, \bar{y})\Big] \tag{6.52}$$

holds. Similarly, $\bar{x}$ is *upper-level regular* if

$$\Big[\lambda_j \geq 0,\ \sum_{j \in J(\bar{x})} \lambda_j \nabla h_j(\bar{x}) = 0\Big] \Longrightarrow \Big[\lambda_j = 0 \ \text{ whenever } \ j \in J(\bar{x})\Big]. \tag{6.53}$$

Now we are ready to present our first result on necessary optimality conditions for the original optimistic version of bilevel programming (6.31) with the upper-level constraint set Ω given in (6.50).

Theorem 6.21 (Optimality Conditions for Smooth Bilevel Programs, I). *Let $(\bar{x}, \bar{y})$ be a local optimal solution to the bilevel program (6.31) with Ω from (6.50). Assume that all the functions therein are smooth around $(\bar{x}, \bar{y})$ and $\bar{x}$, respectively, and that the bilevel program is partially calm at $(\bar{x}, \bar{y})$. Suppose further that $(\bar{x}, \bar{y})$ is lower-level regular, that $\bar{x}$ is upper-level regular, and that the solution map S in (6.30) is inner semicontinuous at $(\bar{x}, \bar{y})$. Then there are numbers $\kappa > 0$, $\lambda_1, \dots, \lambda_p$, $\beta_1, \dots, \beta_p$, and $\alpha_1, \dots, \alpha_q$ such that*

$$\nabla_x f(\bar{x}, \bar{y}) + \sum_{i=1}^{p} (\beta_i - \kappa \lambda_i) \nabla_x g_i(\bar{x}, \bar{y}) + \sum_{j=1}^{q} \alpha_j \nabla h_j(\bar{x}) = 0, \tag{6.54}$$

$$\nabla_y f(\bar{x}, \bar{y}) + \kappa \nabla_y \varphi(\bar{x}, \bar{y}) + \sum_{i=1}^{p} \beta_i \nabla_y g_i(\bar{x}, \bar{y}) = 0, \tag{6.55}$$

$$\nabla_y \varphi(\bar{x}, \bar{y}) + \sum_{i=1}^{p} \lambda_i \nabla_y g_i(\bar{x}, \bar{y}) = 0 \tag{6.56}$$

with the following sign and complementary slackness conditions:

$$\lambda_i \geq 0, \quad \lambda_i g_i(\bar{x}, \bar{y}) = 0 \ \textit{for all} \ \ i = 1, \ldots, p, \tag{6.57}$$

$$\beta_i \geq 0, \quad \beta_i g_i(\bar{x}, \bar{y}) = 0 \ \textit{for all} \ \ i = 1, \ldots, p, \tag{6.58}$$

$$\alpha_j \geq 0, \quad \alpha_j h_j(\bar{x}) = 0 \ \textit{for all} \ \ j = 1, \ldots, q. \tag{6.59}$$

Proof. Proposition 6.13(i) tells us that $(\bar{x}, \bar{y})$ is a local optimal solution to (6.34), even without the lower semicontinuity of S at $(\bar{x}, \bar{y})$. Furthermore, the imposed partial calmness ensures that $(\bar{x}, \bar{y})$ is a local minimizer of the penalized problem (6.51) with some fixed $\kappa > 0$. As mentioned above, the latter problem is a particular case of the nondifferentiable program (6.11) with only the inequality constraints therein. To apply to it the results of Theorem 6.5(ii), we need to check first that the marginal function $\mu(x)$ from (6.33), where $G(x)$ defined in (6.29) is locally Lipschitzian around $\bar{x}$ under the assumed lower-level regularity of $(\bar{x}, \bar{y})$ in the bilevel program under consideration.

Indeed, it is easy to see that the function $\mu(x)$ is l.s.c. around $\bar{x}$. Since $\bar{y} \in S(\bar{x})$, the mapping M in (4.2) obviously reduces in this case to S that is assumed to be inner semicontinuous at $(\bar{x}, \bar{y})$, we deduce from formula (4.5) of Theorem 4.1(i) the following inclusion:

$$\partial^{\infty}\mu(\bar{x}) \subset D^*G(\bar{x}, \bar{y})(0) \ \text{ with } \ G(x) = \big\{ y \in \mathbb{R}^m \big| \ g(x, y) \leq 0 \big\}. \tag{6.60}$$

The result of Exercise 2.51(ii) on representing the normal cone to the set

$$\operatorname{gph} G = \big\{ (x, y) \in \mathbb{R}^n \times \mathbb{R}^m \big| \ g_i(x, y) \leq 0, \quad i = 1, \ldots, p \big\}$$

at $(\bar{x}, \bar{y})$ tells us that $D^*G(\bar{x}, \bar{y})(0) = \{0\}$ under the imposed lower-level regularity. Thus we have $\partial^{\infty}\mu(\bar{x}) = \{0\}$ from (6.60), which ensures that $\mu(\cdot)$ is locally Lipschitzian around $\bar{x}$ due to Theorem 1.22; see also Exercise 4.25(iv).

Applying now the necessary optimality conditions of Theorem 6.5(ii) to problem (6.51) at $(\bar{x}, \bar{y})$ and then using the subdifferential sum rule from Proposition 1.30(ii) give us multipliers $\lambda \geq 0$, $\beta_1, \ldots, \beta_p$, and $\alpha_1, \ldots, \alpha_q$, not all zero, satisfying the sign and complementary slackness conditions in (6.58) and (6.59) and the generalized Lagrangian inclusion

$$\begin{aligned} 0 \in \lambda \nabla f(\bar{x}, \bar{y}) + \kappa \lambda \nabla \varphi(\bar{x}, \bar{y}) + \big(\kappa \lambda \partial(-\mu)(\bar{x}), 0 \big) \\ + \sum_{i=1}^{p} \beta_i \nabla g_i(\bar{x}, \bar{y}) + \sum_{j=1}^{q} \alpha_j \big(\nabla h_j(\bar{x}), 0 \big). \end{aligned} \tag{6.61}$$

It follows from the assumed lower-level regularity of $(\bar{x}, \bar{y})$ and upper-level regularity of $\bar{x}$, combined with the sign and complementarity slackness conditions, that $\lambda \neq 0$ and hence $\lambda = 1$ without loss of generality. Since

$$\partial(-\mu)(\bar{x}) \subset \overline{\partial}(-\mu)(\bar{x}) = -\overline{\partial}\mu(\bar{x}) = -\mathrm{co}\,\partial\mu(\bar{x})$$

by (1.83) and (1.79) due to the Lipschitz continuity of $\mu(x)$, we can incorporate into (6.61) with $\lambda = 1$ the basic subdifferential estimate (4.4) for the marginal function with the smooth constraints (6.29) under the imposed inner semicontinuity assumption on S at $(\bar{x}, \bar{y})$. This gives us multipliers $\lambda_1, \ldots, \lambda_p$ satisfying (6.56) and (6.57) such that the conditions in (6.55) and

$$\nabla_x f(\bar{x}, \bar{y}) + \kappa \nabla_x \varphi(\bar{x}, \bar{y}) - \kappa \Big[\nabla_x \varphi(\bar{x}, \bar{y}) + \sum_{i=1}^{p} \lambda_i \nabla_x g_i(\bar{x}, \bar{y})\Big]$$
$$+ \sum_{i=1}^{p} \beta_i \nabla_x g_i(\bar{x}, \bar{y}) + \sum_{j=1}^{q} \alpha_j \nabla h_j(\bar{x}) = 0$$

hold. Collecting the like terms in the latter equation, we arrive at the remaining equality (6.54) and thus complete the proof of the theorem. △

Now we develop a different device of necessary optimality conditions for bilevel programs, which brings us to results significantly different from Theorem 6.21 in both assumptions and conclusions. To proceed, let us first present a lemma of its own interest that is crucial in the device below. It concerns calculus of regular subgradients, which is pretty limited in general (e.g., no sum rule, etc.) while happens to contain a nice *difference rule* particularly important in applications to bilevel programs via the value function approach. Note that the proof of the following lemma is based on the *smooth variational* description of regular subgradients taken from Theorem 1.27. Observe also that the necessary optimality condition in this lemma has been already deduced in Proposition 6.3 from the upper subdifferential one.

Lemma 6.22 (Difference Rule for Regular Subgradients). *Let both functions $\varphi_1, \varphi_2\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ be finite at $\bar{x}$, and let $\widehat{\partial}\varphi_2(\bar{x}) \neq \emptyset$. Then we have*

$$\widehat{\partial}(\varphi_1 - \varphi_2)(\bar{x}) \subset \bigcap_{v \in \widehat{\partial}\varphi_2(\bar{x})} \Big[\widehat{\partial}\varphi_1(\bar{x}) - v\Big] \subset \widehat{\partial}\varphi_1(\bar{x}) - \widehat{\partial}\varphi_2(\bar{x}). \qquad (6.62)$$

This implies that any local minimizer $\bar{x}$ of the difference function $\varphi_1 - \varphi_2$ satisfies the necessary optimality condition

$$\widehat{\partial}\varphi_2(\bar{x}) \subset \widehat{\partial}\varphi_1(\bar{x}). \qquad (6.63)$$

Proof. To verify the first inclusion in (6.62), fix any $u \in \widehat{\partial}(\varphi_1 - \varphi_2)(\bar{x})$ and $v \in \widehat{\partial}\varphi_2(\bar{x})$. Employing the first assertion of Theorem 1.27, find a real-valued function $s(\cdot)$ defined on a neighborhood U of $\bar{x}$ such that it is (Fréchet) differentiable at $\bar{x}$ satisfying the relationships

$$s(\bar{x}) = \varphi_2(\bar{x}), \quad \nabla s(\bar{x}) = v, \quad \text{and} \quad s(x) \le \varphi_2(x) \quad \text{for all} \quad x \in U.$$

This yields due to definition (1.33) of the regular subgradient $u \in \widehat{\partial}(\varphi_1 - \varphi_2)(\bar{x})$ that for any $\varepsilon > 0$ there exists $\gamma > 0$ such that

$$\begin{aligned}\langle u, x - \bar{x}\rangle &\le \varphi_1(x) - \varphi_2(x) - \big(\varphi_1(\bar{x}) - \varphi_2(\bar{x})\big) + \varepsilon\|x - \bar{x}\|\\ &\le \varphi_1(x) - s(x) - \big(\varphi_1(\bar{x}) - s(\bar{x})\big) + \varepsilon\|x - \bar{x}\|\end{aligned}$$

whenever $\|x - \bar{x}\| \le \gamma$. The latter ensures by Proposition 1.30(ii) that

$$u \in \widehat{\partial}(\varphi_1 - s)(\bar{x}) = \widehat{\partial}\varphi_1(\bar{x}) - \nabla s(\bar{x}) = \widehat{\partial}\varphi_1(\bar{x}) - v,$$

which justifies the first inclusion in (6.62) and obviously yields the second one.

To verify (6.63), observe that it is trivial if $\widehat{\partial}\varphi_2(\bar{x}) = \emptyset$. Otherwise, pick $v \in \widehat{\partial}\varphi_2(\bar{x})$ and deduce from (6.62) by the generalized Fermat rule that

$$0 \in \widehat{\partial}(\varphi_1 - \varphi_2)(\bar{x}) \subset \widehat{\partial}\varphi_1(\bar{x}) - v,$$

which shows that $v \in \widehat{\partial}\varphi_1(\bar{x})$ and thus justifies the set inclusion (6.63). △

For simplicity we consider in the next theorem the optimistic bilevel problem (6.31) without upper-level constraints.

Theorem 6.23 (Optimality Conditions for Smooth Bilevel Programs, II). *Let $(\bar{x}, \bar{y})$ be a local optimal solution to problem (6.31) with $\Omega = \mathbb{R}^n$ and with the functions $f, g_1, \ldots, g_p, \varphi$ continuously differentiable around $(\bar{x}, \bar{y})$. Assume that this problem is partially calm at the point $(\bar{x}, \bar{y})$, which is lower-level regular for (6.34), and also that $\widehat{\partial}\mu(\bar{x}) \ne \emptyset$ for lower-level value function (6.33). Then there exist multipliers ν_i and β_i as $i = 1, \ldots, p$ such that β_i satisfy the sign and complementarity slackness conditions in (6.58), that ν_i satisfy the complementarity slackness conditions*

$$\nu_i g_i(\bar{x}, \bar{y}) = 0 \ \textit{for all} \ \ i = 1, \ldots, p,$$

and that the following equalities hold:

$$\begin{aligned}&\nabla f(\bar{x}, \bar{y}) + \sum_{i=1}^{p} \nu_i \nabla g_i(\bar{x}, \bar{y}) = 0,\\ &\nabla_y\varphi(\bar{x}, \bar{y}) + \sum_{i=1}^{p} \beta_i \nabla_y g_i(\bar{x}, \bar{y}) = 0.\end{aligned}$$

Proof. By Proposition 6.13(i) we get that $(\bar{x}, \bar{y})$ is a local optimal solution to the nondifferentiable program (6.34). It follows from Proposition 6.15 and the infinite constraint penalization via the indicator function $\delta(\cdot; \operatorname{gph} G)$ that $(\bar{x}, \bar{y})$ a local optimal solution to the unconstrained problem:

$$\text{minimize} \ \ f(x, y) + \kappa\big(\varphi(x, y) - \mu(x)\big) + \delta\big((x, y); \operatorname{gph} G\big), \tag{6.64}$$

where the constant $\kappa > 0$ is taken from the definition of partial calmness. Applying now the necessary optimality condition (6.63) from Lemma 6.22 to the difference function in (6.64), we get

$$\big(\kappa\widehat{\partial}\mu(\bar{x}),0\big)\subset\widehat{\partial}\big(f(\cdot)+\kappa\varphi(\cdot)+\delta(\cdot;\operatorname{gph}G)\big)(\bar{x},\bar{y}). \tag{6.65}$$

It is not hard to observe (cf. the proof of Theorem 4.1) that

$$\big(\widehat{\partial}\mu(\bar{x}),0\big)\subset\widehat{\partial}\big(\varphi(\cdot)+\delta(\cdot;\operatorname{gph}G)\big)(\bar{x},\bar{y}). \tag{6.66}$$

Passing to the larger limiting subdifferential on the right-hand sides of (6.65) and (6.66) and employing the elementary subdifferential sum rule, we have

$$\begin{aligned}&\big(\kappa\widehat{\partial}\mu(\bar{x}),0\big)\subset\nabla f(\bar{x},\bar{y})+\kappa\nabla\varphi(\bar{x},\bar{y})+N\big((\bar{x},\bar{y});\operatorname{gph}G\big),\\&\big(\widehat{\partial}\mu(\bar{x}),0\big)\subset\nabla\varphi(\bar{x},\bar{y})+N\big((\bar{x},\bar{y});\operatorname{gph}G\big).\end{aligned}$$

Then the description of basic normals from Exercise 2.51 for sets given by inequality constraints under the imposed lower-level regularity ensures the existence of multipliers λ_i and β_i satisfying the sign and complementarity slackness conditions in (6.57) and (6.58) as well as a vector $v\in\widehat{\partial}\mu(\bar{x})$ with

$$\begin{aligned}&(v,0)=\nabla\varphi(\bar{x},\bar{y})+\sum_{i=1}^{p}\lambda_i\nabla g_i(\bar{x},\bar{y})\ \text{ and}\\&\kappa(v,0)=\nabla f(\bar{x},\bar{y})+\kappa\nabla\varphi(\bar{x},\bar{y})+\sum_{i=1}^{p}\beta_i\nabla g_i(\bar{x},\bar{y}).\end{aligned}$$

Dividing the latter inclusion by $\kappa > 0$ and denoting $\nu := \kappa^{-1}$ while keeping the same notation for the modified multipliers β_i and collecting the like terms, we arrive at the equalities claimed in the theorem. △

The following example, consisting of two parts, illustrates the possibility to solve bilevel programs by using necessary optimality conditions obtained in Theorem 6.21 and Theorem 6.23, respectively.

Example 6.24 (Solving Bilevel Programs via Optimality Conditions).
(i) *Applying the conditions of Theorem* 6.21. Consider the bilevel program:

$$\text{minimize }\ f(x,y):=-y\ \text{ subject to }\ y\in S(x),$$

where $S\colon\mathbb{R}\rightrightarrows\mathbb{R}$ is the solution map of the nonlinear lower-level problem:

$$\begin{aligned}&\text{minimize }\ \varphi(x,y):=-y^2+x^4-3x^2+1\ \text{ subject to}\\&y\in G(x):=\big\{y\in\mathbb{R}\big|\ y+x^2-1\le 0,\ -y+x^2-1\le 0\big\}.\end{aligned}$$

It is easy to check that the bilevel program in this example admits an optimal solution with x belonging to the interval $[-1,1]$. Furthermore, we have

$$S(x) = \big\{ -x^2 + 1, x^2 - 1 \big\} \text{ and } \mu(x) = -x^2 \text{ for } x \in [-1, 1].$$

This shows that S is inner semicontinuous at any point $(x, y) \in \operatorname{gph} S$ and the lower-regularity assumption (6.52) is satisfied everywhere but $(-1, 0)$ and $(1, 0)$; the upper regularity is automatic due to the absence of inequality constraints on the upper level. Applying Theorem 6.21, we calculate

$$\begin{aligned} &\nabla f(x, y) = (0, -1), \quad \nabla \varphi(x, y) = (4x^3 - 6x, -2y),\\ &\nabla g_1(x, y) = (2x, 1), \quad \nabla g_2(x, y) = (2x, -1) \end{aligned}$$

and hence obtain the following relationships:

$$\begin{aligned} &0 = (\beta_1 - \kappa\lambda_1)2x + (\beta_2 - \kappa\lambda_2)2x, \quad 0 = -1 + \kappa(-2y) + \beta_1(1) + \beta_2(-1),\\ &0 = -2y + \lambda_1(1) + \lambda_2(-1), \quad 0 = \lambda_1(y + x^2 - 1) = \lambda_2(-y + x^2 - 1),\\ &0 = \beta_1(y + x^2 - 1) = \beta_2(-y + x^2 - 1) \end{aligned}$$

with $\kappa > 0$ and all the nonnegative multipliers. Solving the above system gives us the points $(x, y) \in \{(0, 1), (0, -1), (1, 0), (-1, 0)\}$ suspicious for optimality. Comparing the value of the upper-level objective at these points, we arrive at the pair $(\bar{x}, \bar{y}) = (0, 1)$ and check finally that the given bilevel program is partially calm at $(0, 1)$. Thus this pair is the unique optimal solution to the bilevel program under consideration by Theorem 6.21.

(ii) *Applying the conditions of Theorem* 6.23. Consider the program:

$$\text{minimize } f(x, y) := -y \text{ subject to } y \in S(x),$$

where $S\colon \mathbb{R} \rightrightarrows \mathbb{R}$ is the solution map for the lower-level problem:

$$\begin{aligned} &\text{minimize } \varphi(x, y) := -y^2 \text{ subject to}\\ &y \in G(x) := \big\{ y \in \mathbb{R} \big| -x + y^4 - 1 \le 0,\ x + y^4 - 1 \le 0 \big\}. \end{aligned}$$

It is easy to see that this bilevel program admits an optimal solution. Then we calculate the lower-level solution map by $S(x) = \{\pm\sqrt{1 - |x|}\}$ and the marginal function by $\mu(x) = -\sqrt{1 - |x|}$ for which $\widehat{\partial}\mu(x) \neq \emptyset$ on $\mathbb{R}$. Applying the necessary optimality conditions of Theorem 6.23 gives us the relationships

$$\begin{gathered} -\nu_1 + \nu_2 = 0,\ -1 + 4y^3\nu_1 + 4y^3\nu_2 = 0,\\ -2y + 4y^3\beta_1 + 4y^3\beta_2 = 0,\ \nu_1(-x + y^4 - 1) = \nu_2(x + y^4 - 1) = 0,\\ \beta_1(-x + y^4 - 1) = \beta_1(x + y^4 - 1) = 0,\ \beta_1 \ge 0,\ \beta_2 \ge 0. \end{gathered}$$

Solving this system of equations, we obtain the points $(x, y) = (0, \pm 1)$. Comparing the upper-level objective selects the point $(0, 1)$. Since the bilevel program under consideration is partially calm at $(0, 1)$, we conclude that $(0, 1)$ is the unique optimal solution to this problem.

6.3.2 Optimality Conditions for Lipschitzian Problems

Analyzing the proofs of Theorem 6.21 and Theorem 6.23, it is not difficult to observe that these proofs and the results used therein lead us to necessary optimality conditions for bilevel programs with *Lipschitzian data*. In the following Lipschitzian versions of necessary optimality conditions, we replace the gradients of the Lipschitzian functions involved by their basic subgradients and reformulate the upper-level regularity condition (6.53) as satisfied for all the subgradients of h_j at $\bar{x}$ and the lower-level regularity condition (6.52) as satisfied for all (u_i, v_i) with $(u_i, v_i) \in \partial g_i(\bar{x}, \bar{y})$. In this way we have:

Theorem 6.25 (Optimality Conditions for Lipschitzian Bilevel Programs, I). *Let $(\bar{x}, \bar{y})$ be a local optimal solution to the bilevel program (6.31) with Ω from (6.50). Suppose that all the functions therein are locally Lipschitzian around $(\bar{x}, \bar{y})$ and $\bar{x}$, respectively, under the validity of the other assumptions of Theorem 6.21. Then there exist a number $\nu > 0$, multipliers $\lambda_1, \ldots, \lambda_p$, $\beta_1, \ldots, \beta_p$, and $\alpha_1, \ldots, \alpha_q$ as well as a vector $u \in \mathbb{R}^n$ such that conditions (6.57)–(6.59) are satisfied together with*

$$(u, 0) \in \operatorname{co} \partial\varphi(\bar{x}, \bar{y}) + \sum_{i=1}^{p} \lambda_i \operatorname{co} \partial g_i(\bar{x}, \bar{y}) \quad \textit{and}$$

$$(u, 0) \in \partial\varphi(\bar{x}, \bar{y}) + \nu\partial f(\bar{x}, \bar{y}) + \sum_{i=1}^{p} \beta_i \partial g_i(\bar{x}, \bar{y}) + \sum_{j=1}^{q} \alpha_j \Big(\partial h_j(\bar{x}), 0\Big).$$

Theorem 6.26 (Optimality Conditions for Lipschitzian Bilevel Programs, II). *Let $(\bar{x}, \bar{y})$ be a local optimal solution to problem (6.31) with $\Omega = \mathbb{R}^n$. Suppose that all the functions therein are locally Lipschitzian around $(\bar{x}, \bar{y})$ under the validity of the other assumptions of Theorem 6.23. Then there exist a number $\nu > 0$, nonnegative multipliers λ_i and β_i satisfying the complementary slackness condition (6.57) and (6.58) as $i = 1, \ldots, p$, and a vector $u \in \mathbb{R}^n$ such that we have the inclusions*

$$(u, 0) \in \partial\varphi(\bar{x}, \bar{y}) + \sum_{i=1}^{p} \lambda_i \partial g_i(\bar{x}, \bar{y}) \quad \textit{and}$$

$$(u, 0) \in \partial\varphi(\bar{x}, \bar{y}) + \nu\partial f(\bar{x}, \bar{y}) + \sum_{i=1}^{p} \beta_i \partial g_i(\bar{x}, \bar{y}).$$

The proofs of these results as well as their several extensions are assigned in the exercises of Section 3.4.

Remark 6.27 (Inner Semicompactness vs. Inner Semicontinuity of Solution Maps). Observe that the necessary optimality conditions of Theorems 6.23 and 6.26 hold, in contrast to those in Theorems 6.21 and 6.25, without the inner semicontinuity assumption on the solution map S (6.30). While the latter assumption is satisfied in rather broad settings (e.g., when S is Lipschitz-like around $(\bar{x}, \bar{y})$ and also when

$S(\bar{x})$ is a singleton but $S(x)$ may not be for x close to $\bar{x}$), it definitely doesn't hold in generality.

A significantly less restrictive assumption in the frameworks of Theorems 6.21 and 6.25 is provided by the *inner semicompactness* property of S at the domain point $\bar{x}$ defined in Exercise 2.46. In finite dimensions this property is rather close to the *local boundedness* of S around $\bar{x}$. The results obtained under the lower semicompactness of S are different from their inner semicontinuity counterparts in that they require considering all the vectors $\bar{y}$ from the set $S(\bar{x})$. The proofs go in the same direction with replacing the results on the subdifferentiation of marginal functions from Theorem 4.1(i) by their "union" versions from assertion (ii) therein.

Some consequences and specifications of the necessary optimality conditions for bilevel programs with fully and partially convex (smooth and nonsmooth) structures can be derived from Theorems 6.25 and 6.26. However, significantly stronger results for problems of these type will be obtained as consequences of those given in Subsection 7.5.4. Hence we omit here formulating the corresponding consequences of Theorems 6.25 and 6.26 while leaving this as exercises for the reader; see more hints in Exercise 6.46.

6.4 Exercises for Chapter 6

Exercise 6.28 (Optimization Problems with Geometric Constraints).

(i) Derive necessary optimality conditions of Theorem 6.1(ii) and Proposition 6.4(ii) directly from the extremal principle.

(ii) Extend the necessary optimality conditions of Theorem 6.1 and Proposition 6.4 to appropriate Banach spaces. Which assumption should be added to (6.4) to ensure the validity of Theorem 6.1(ii) in infinite dimensions? *Hint:* Compare this with [523, Propositions 5.2 and 5.3 and Theorem 5.5].

(iii) Construct an example of the optimization problem (6.1) with a Lipschitz continuous objective function φ defined on a Banach space X such that condition (6.5) doesn't hold at a local minimizer $\bar{x}$ of this problem.

Exercise 6.29 (Problems of DC Programming).

(i) Extend the results of Proposition 6.3 to problems with convex geometric constraints of the type $x \in \Omega$.

(ii) Show that the convexity of the function φ_1 in Proposition 6.3 can be replaced by the more general property of *quasiconvexity* in the sense that

$$\varphi\big(\lambda x_1 + (1-\lambda)x_2\big) \le \max\big\{\varphi(x_1), \varphi(x_2)\big\} \ \text{ for all } \ x_1, x_2 \in \mathbb{R}^n, \ \lambda \in [0,1].$$

(iii) Do all the results of Proposition 6.3 and the parts (i) and (ii) of this exercise hold in arbitrary Banach spaces?

Exercise 6.30 (Necessary Conditions in Nondifferentiable Programming).

(i) Extend necessary optimality conditions of Theorem 6.5 for problems of nondifferentiable programming of type (6.11) with finitely many geometric constraints.

(ii) Derive appropriate versions of Theorem 6.5 in Asplund spaces. *Hint:* Proceed as in the proof of Theorem 6.5 with applying the corresponding results in infinite dimensions taken from Exercises 2.31 and 1.69; compare with [523, Theorem 5.5].

Exercise 6.31 (Extended Lagrangian Conditions for Lipschitzian Nondifferentiable Programs in Asplund Spaces). Consider the nondifferentiable program (6.11) described by locally Lipschitzian functions on an Asplund space. Show that the necessary optimality conditions of Theorem 6.10 hold true in this case.

Hint: Proceed by using the exact extremal principle from Exercise 2.31 and the subdifferential sum from Exercise 2.54 in Asplund spaces; cf. [523, Theorem 5.24].

Exercise 6.32 (Constraint Qualifications in Nondifferentiable Programming).

(i) Based on Theorem 6.5(ii) and the normal cone representations from Exercises 2.51 and 2.52 in the case of locally Lipschitzian functions, derive constraint qualifications ensuring that $\lambda_0 = 1$ in the optimality conditions of Theorem 6.5.

(ii) Which constraint qualifications correspond to those obtained in (i) in the case of smooth constraint functions φ_i and $\Omega = \mathbb{R}^n$?

(iii) Derive extensions of the results in (i) to problems in Asplund spaces.

Exercise 6.33 (Necessary Optimality Conditions for Problems with Inclusion/Operator Constraints). Given $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$, $f\colon \mathbb{R}^n \to \mathbb{R}^m$, $\Omega \subset \mathbb{R}^n$, and $\Theta \subset \mathbb{R}^m$, consider the optimization problem:

$$\text{minimize } \varphi(x) \text{ subject to } f(x) \in \Theta,\ x \in \Omega, \tag{6.67}$$

where f is strictly differentiable at the reference local minimizer $\bar{x} \in f^{-1}(\Theta) \cap \Omega$ and its Jacobian matrix $\nabla f(\bar{x})$ has full row rank.

(i) Prove that the upper subdifferential optimality condition

$$-\widehat{\partial}^+\varphi(\bar{x}) \subset \nabla f(\bar{x})^* N\big(f(\bar{x}); \Theta\big) + N(\bar{x}; \Omega)$$

holds under the validity of the constraint qualification

$$\nabla f(\bar{x})^* N\big(f(\bar{x}); \Theta\big) \cap \big(-N(\bar{x}; \Omega)\big) = \{0\}.$$

(ii) Derive lower subdifferential optimality conditions for $\bar{x}$ in both qualified/KKT and nonqualified/Fritz John forms.

(iii) Extend the results of (i) and (ii) to appropriate infinite-dimensional spaces and specify the results for the operator constraints $f(x) = 0 \in Y$ with $\dim Y = \infty$.

Hint: Employ the corresponding calculus rules in the framework of Proposition 6.4 with $\Omega_1 := f^{-1}(\Theta)$, $\Omega_2 := \Omega$; compare it with [523, Theorems 5.7,5.8,5.11].

Exercise 6.34 (Optimization Problems with Inverse Image Constraints via the Extremal Principle). Given $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$, $F\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, $\Omega \subset \mathbb{R}^n$, and $\Theta \subset \mathbb{R}^m$, consider the optimization problem:

$$\text{minimize } \varphi(x) \text{ subject to } F^{-1}(\Theta) \cap \Omega. \tag{6.68}$$

(i) Let $\bar{x}$ be a local minimizer for problem (6.68). Show that the point $(\bar{x}, \varphi(\bar{x}))$ is locally extremal for the system of three sets in $\mathbb{R}^n \times \mathbb{R}$:

$$\Omega_1 := \operatorname{epi}\varphi, \quad \Omega_2 := F^{-1}(\Theta), \quad \Omega_3 := \Omega \times \mathbb{R}.$$

(ii) Derive necessary optimality conditions for problem (6.68) by applying the extremal principle to the set system in (i).

(iii) Extend the result of (ii) to problem (6.68) in Asplund spaces.

Exercise 6.35 (Suboptimality Conditions in Nonlinear and Nondifferentiable Programming). Consider problem (6.11), fix $\varepsilon > 0$, and recall that x_ε is an ε-*optimal* (suboptimal) solution to this problem if it is feasible to (6.11) and satisfies the inequality $\varphi_0(x_\varepsilon) \le \inf \varphi_0(x) + \varepsilon$, where the infimum of φ_0 is taken over all the feasible solutions to problem (6.11).

(i) Assume that $\Omega = \mathbb{R}^n$ and that the functions $\varphi_0, \ldots, \varphi_{m+r}$ are strictly differentiable on the set of ε-optimal solutions to (6.11) while $\varphi_1, \ldots, \varphi_{m+r}$ satisfy the Mangasarian-Fromovitz constraint qualifications (see Exercise 2.53) on this set. Then for any ε-optimal solution x_ε to (6.11) and any $\gamma > 0$, there exist an ε-optimal solution $\bar{x}$ to this problem and multipliers $\lambda_1, \ldots, \lambda_{m+r}$ such that

$$\|\bar{x} - x_\varepsilon\| \le \gamma,\ \ \lambda_i \ge 0,\ \ \lambda_i \varphi_i(\bar{x}) = 0 \ \text{ for } \ i = 1, \ldots, m,$$
$$\Big\|\nabla\varphi_0(\bar{x}) + \sum_{i=1}^{m+r} \lambda_i \nabla\varphi_i(\bar{x})\Big\| \le \frac{\varepsilon}{\gamma}.$$

(ii) Assume that $\varphi_i, i = 0, \ldots, m+r$, are locally Lipschitzian on the set of ε-optimal solutions to (6.11) and that Ω is closed therein. Then for any ε-optimal solution x_ε to (6.11) and any $\gamma > 0$, there exist an ε-optimal solution $\bar{x}$ to this problem and multipliers $\lambda_0, \ldots, \lambda_{m+r}$ such that $\|\bar{x} - x_\varepsilon\| \le \gamma$,

$$\Big\| \sum_{i \in I(\bar{x}) \cup \{0\}} \lambda_i x_i^* + x^* \Big\| \le \frac{\varepsilon}{\gamma}, \qquad \sum_{i \in I(\bar{x}) \cup \{0\}} \lambda_i = 1$$

with some $\lambda_i \ge 0$ for $i \in I(\bar{x}) \cup \{0\}$, $x^* \in N(\bar{x}; \Omega)$, $x_0^* \in \partial\varphi_0(\bar{x})$,

$$x_i^* \in \partial\varphi_i(\bar{x}) \ \text{ for } \ i \in \big\{1, \ldots, m\big\} \cap I(\bar{x}), \quad \text{ and}$$
$$x_i^* \in \partial\varphi_i(\bar{x}) \cup \partial(-\varphi_i)(\bar{x}) \ \text{ for } \ i = m+1, \ldots, m+r.$$

(iii) Extend the results in (i) and (ii) to problems (6.11) in Asplund spaces.

Hint: Employ the subdifferential variational principle from Exercise 2.39 and then the subdifferential sum rule from Corollary 2.20; cf. [523, Theorem 5.30].

Exercise 6.36 (Single-Level Reduction of Optimistic Bilevel Programs)

(i) Show that standard constraint qualifications (of the Mangasarian-Fromovitz type, etc.) fail for the nondifferentiable program (6.34). *Hint:* Compare it with the results and proofs in [194, 745, 748].

(ii) Show that the inner semicontinuity assumption on the mapping $\widetilde{S}$ from (6.36) at $(\bar{x}, \bar{y})$ is essential for the validity of Proposition 6.13(ii).

(iii) Prove that assertion (i) of Proposition 6.13 holds for *some* $\bar{y} \in S(\bar{x})$ if the latter set from (6.30) is assumed to be bounded and the upper-level cost function $f(\bar{x}, \cdot)$ is assumed to be l.s.c. on $S(\bar{x})$. Give examples showing that both of these assumptions are essential for the validity of the result in question. Does it follow from the presented version of Proposition 6.13(i)?

Exercise 6.37 (Partial Calmness and Uniform Weak Sharp Minima in Bilevel Programming). Consider the class of optimistic bilevel programs in form (6.34).

(i) Let $\Omega = \mathbb{R}^n$, and let g_i in (6.34) be *linear* with respect y with $\operatorname{dom} G = \mathbb{R}^n$. Prove that any local optimal solution $(\bar{x}, \bar{y})$ to (6.34) is partially calm provided that f is locally Lipschitzian around this point; compare with the results in [201, 748].

(ii) Construct an example of a bilevel program partially calm at its local optimal solution without the validity of the uniform weak sharp minimum condition (6.43).

(iii) Construct an example of a bilevel program where the partial calmness condition fails at a local optimal solution.

Exercise 6.38 (Sufficient Conditions for Uniform Weak Sharp Minima).

(i) Under which assumptions on problem (6.34) the pointwise local weak sharp minima as in (6.42) with data (6.44) yields the uniform one as in (6.43)?

(ii) Show that the set of optimal solutions to problem (6.49) consists of weak sharp minimizers if either $a_i > 0$ or $b_i < 0$ for $i = 1, 2$. *Hint:* Compare it with [133].

(iii) Derive sufficient conditions for uniform sharp minima in the case of quadratic lower-level problems. *Hint:* Compare it with [748, 749].

Exercise 6.39 (Kernel Condition for Weak Sharp Minima)

(i) Is the kernel condition (6.46) equivalent to a full rank property of a matrix?

(ii) Apply the kernel condition (6.46) in the parametric version of (6.45) with data (6.44) to ensure the uniform weak sharp minima in bilevel programming.

(iii) Extend the result of Proposition 6.18 to Lipschitzian nonlinear programs, and apply it to bilevel programs with nonsmooth data.

Exercise 6.40 (Inf-Differentiability and Dual Characterizations of Weak Sharp Minimizers). Considering a function $\varphi\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ and a set $\Omega \subset \mathbb{R}^n$, we say as in [785] that φ is *inf-differentiable* at $\bar{x} \in \operatorname{dom}\varphi$ *relative* to Ω if

$$\liminf_{x \xrightarrow{\Omega} \bar{x},\, u \to \bar{x}} \frac{\varphi(u) - \varphi(x) - d\varphi(x; u - x)}{\|u - x\|} = 0, \tag{6.69}$$

where the contingent directional derivative $d\varphi$ is taken from (1.42). In particular, if (6.69) holds with $\Omega = \mathbb{R}^n$ and with $\Omega = \{\bar{x}\}$, then φ is called to be *inf-differentiable at* $\bar{x}$ and *single inf-differentiable* at $\bar{x}$, respectively.

(i) Verify that if φ is locally Lipschitzian around $\bar{x}$, then it is single inf-differentiable at this point. Could the later property hold for non-Lipschitzian functions?

(ii) Show that every convex function is inf-differentiable on any closed and bounded subset of the interior of its domain.

(iii) Let φ be locally Lipschitzian around $\bar{x}$, subdifferentially regular on the set $L_\varphi(\bar{x}) := \{x \in \mathbb{R}^n \,|\, \varphi(x) = \varphi(\bar{x})\}$ and inf-differentiable at $\bar{x}$ relative to $L_\varphi(\bar{x})$. Prove that the existence of $\eta, r > 0$ such that the inclusion

$$N\big(x; L_\varphi(\bar{x})\big) \cap \eta\mathbb{B} \subset \partial\varphi(x)$$

holds for any $x \in L_\varphi(\bar{x}) \cap B_r(\bar{x})$ is necessary and sufficient for the following specification of local weak sharp minima in Definition 6.16:

$$\eta \operatorname{dist}\big(x; L_\varphi(\bar{x})\big) \le \varphi(x) - \varphi(\bar{x}) \quad \text{whenever} \quad x \in B_r(\bar{x}).$$

Hint: Compare (i)–(iii) with the corresponding statements and proofs in [785].

(iv) Clarify possible counterparts of (iii) for the study of uniform weak sharp minima in parametric optimization and bilevel programs.

Exercise 6.41 (Regular Subgradients of Value Functions in Lower-Level Problems). Let $\mu(x)$ be the optimal value function of the lower-level problem in (6.34).

(i) Give a detailed proof of inclusion (6.65) in general Banach spaces.

(ii) Show that the equalities don't hold in (6.64) and (6.65) for problems with smooth data in finite dimensions.

Exercise 6.42 (Comparing Necessary Optimality Conditions for Bilevel Programs with Smooth Data). Construct examples in which all the assumptions of both Theorem 6.21 and Theorem 6.23 are satisfied while the necessary optimality conditions obtained in these theorems are independent.

Exercise 6.43 (Necessary Optimality Conditions in Lipschitzian Bilevel Programming). Consider local optimal solutions to the optimistic model (6.31).

(i) Give a detailed proof of Theorem 6.25. *Hint:* Compare it with [195, 540].

(ii) Give a detailed proof of Theorem 6.26. *Hint:* Compare it with [540].

(iii) Extend these theorems to bilevel programs with Lipschitzian (and smooth) data in the presence of equality constraints.

(iv) Derive versions of these results for bilevel problems in Asplund spaces. *Hint:* Apply the calculus rules used in the proofs of Theorems 6.21 and 6.23, their equality constraint versions presented in Chapters 2 and 4, and their infinite-dimensional extensions discussed therein in the commentaries and exercises.

(v) Investigate the possibility to improve the necessary optimality conditions in Theorems 6.21 and 6.25 by using the symmetric subdifferential $\partial^0\mu(\bar{x})$ of the value function (6.33) instead of the convexified one in their proofs.

Exercise 6.44 (Extended Inner Semicontinuity in Bilevel Programming). Obtain finite-dimensional and Asplund space extensions of Theorems 6.21 and 6.25 with replacing the inner semicontinuity of the solution map $S(x)$ by its μ-inner semicontinuity defined in Exercise 4.21. *Hint:* Proceed similarly the proofs of these theorems and compare it with [540].

Exercise 6.45 (Bilevel Programs with Inner Semicompact Solution Maps for Lower-level Problems). Considering the solution map $S(x)$ of the lower-level problem in (6.34), verify the following assertions:

(i) $S(x)$ may not be inner semicontinuous at $(\bar{x}, \bar{y})$ as in Theorems 6.21 and 6.25.

(ii) Show that the necessary optimality conditions of Theorems 6.21 and 6.25 may fail without the inner semicontinuity requirement imposed on $S(x)$ at $(\bar{x}, \bar{y})$ under the validity of the other assumptions therein.

(iii) Derive the corresponding version of Theorems 6.21 and 6.25 with replacing the inner semicontinuity of $S(x)$ by its inner semicompactness as well as the more general μ-semicompactness property. *Hint:* Proceed as discussed in Remark 6.27 in the case of finite-dimensional and Asplund spaces.

Exercise 6.46 (Convex Bilevel Programs). Consider the bilevel program (6.31) and its partially calm local optimal solution. Suppose that the lower-level cost and constraint functions are convex jointly with respect to all their variables.

(i) Prove the convexity of the optimal value function (6.33).

(ii) Assuming that the upper-level cost and constraint functions are also fully convex, derive a specification of Theorem 6.21 by using the decomposition property

$$\partial\psi(\bar{x}, \bar{y}) \subset \partial_x\psi(\bar{x}, \bar{y}) \times \partial_y\psi(\bar{x}, \bar{y})$$

valued for full and partial subdifferentials of convex continuous functions ψ and the symmetric property $\partial(-\varphi)(\bar{x}) \subset -\partial\varphi(\bar{x})$. *Hint:* Compare this with [195].

(iii) Assuming that all the functions involved in (6.31) are continuously differentiable in addition to full convexity at the lower level, derive a further specification of Theorem 6.21 by using the equality formula

$$\partial\mu(\bar{x}) = \bigcup_{(\lambda_1,\ldots,\lambda_p)\in\Lambda(\bar{x},\bar{y})} \Big\{\nabla_x\varphi(\bar{x}, \bar{y}) + \sum_{i=1}^{p} \nabla_x g_i(\bar{x}, \bar{y})\Big\} \tag{6.70}$$

for the subdifferential of the optimal value function, where

$$\Lambda(\bar{x}, \bar{y}) := \Big\{(\lambda_1, \ldots, \lambda_p) \in \mathbb{R}^p \Big|\ \nabla_y\varphi(\bar{x}, \bar{y}) + \sum_{i=1}^{p} \lambda_i\nabla_y g_i(\bar{x}, \bar{y}) = 0,$$
$$\lambda_i \geq 0,\ \lambda_i g_i(\bar{x}, \bar{y}) = 0,\ i = 1, \ldots, p\Big\}.$$

Hint: Deduce (6.70) from the equality representation

$$\partial\mu(\bar{x}) = \bigcup_{(u,v)\in\partial\varphi(\bar{x},\bar{y})} \Big\{u + D^*G(\bar{x}, \bar{y})(v)\Big\}$$

for the subdifferential of the marginal function (6.33) with a convex function φ and a convex-graph mapping G given in [537, Theorem 2.61] and the normal cone representations from Exercises 2.51

and 2.52 with the equalities therein for convex functions due to the equality statement of Theorem 2.26. Compare this with another approach to justify (6.70) with convex differentiable data originated in [703].

(iv) Derive the corresponding consequences of Theorem 6.23 for convex bilevel programs with continuous data cost functions and inequality constraints.

Exercise 6.47 (Hölder Subgradients in Bilevel Programming). Given a Banach space X, we say as in [108] that $x^* \in X^*$ is a *Hölder subgradient* of order $s \geq 0$ for $\varphi\colon X \to \overline{\mathbb{R}}$ at $\bar{x} \in \operatorname{dom}\varphi$ if there are constants $C \geq 0$ and $r > 0$ such that

$$\langle x^*, x - \bar{x}\rangle \leq \varphi(x) - \varphi(\bar{x}) + C\|x - \bar{x}\|^{1+s} \quad \text{for all} \quad x \in \bar{x} + r\mathbb{B}. \tag{6.71}$$

The collection of all x^* satisfying (6.71) is called the *s-Hölder subdifferential* of φ at $\bar{x}$ and is denoted by $\widehat{\partial}_{H(s)}(\bar{x})$. The case of $s = 0$ in (6.71) reduces to the regular/Fréchet subdifferential, while the case of $s = 1$ corresponds to the proximal subdifferential $\partial_p\varphi(\bar{x})$ defined above. We also consider the *upper s-Hölder subdifferentials* of φ at $\bar{x} \in \operatorname{dom}\varphi$ defined symmetrically by

$$\widehat{\partial}^+_{H(s)}\varphi(\bar{x}) := -\widehat{\partial}_{H(s)}(-\varphi)(\bar{x}).$$

Similarly to our basic subdifferential, let us introduce the *limiting s-Hölder subdifferential* $\partial_{H(s)}(x)$ of φ at $\bar{x}$ by taking the outer limit of $\widehat{\partial}_{H(s)}(x)$ as $x \xrightarrow{\varphi} \bar{x}$.

(i) Show that the regular subgradient difference rule given in Lemma 6.22 can be extended to the s-Hölder subdifferentials of any real order $s \geq 0$. *Hint:* Proceed as in the proof of Lemma 6.22 and compare it with [540].

(ii) For each $s \geq 0$, determine the classes of Banach spaces, where the limiting s-Hölder subdifferential $\partial_{H(s)}(\bar{x})$ agrees with our basic limiting construction $\partial\varphi(\bar{x})$, and where these constructions may be different.

(iii) Derive counterparts of the necessary optimality conditions from Theorem 6.26 in terms of the corresponding s-Hölder subdifferentials, and clarify whether they are different, in appropriate Banach spaces, from those given in the theorem. *Hint:* For the latter part, apply the tools of analysis developed in [108].

Exercise 6.48 (Mathematical Programs with Equilibrium Constraints). This class of optimization problems (abbr. *MPECs*) is written in the form:

$$\text{minimize} \;\; f(x, y) \;\; \text{subject to} \;\; y \in S(x), \;\; x \in \Omega, \tag{6.72}$$

where $f\colon X \times Y \to \overline{\mathbb{R}}$ is defined on finite-dimensional or infinite-dimensional spaces and where $S\colon X \rightrightarrows Y$ is given, with $q\colon X \times Y \to Z$ and $Q\colon X \times Y \rightrightarrows Z$, by

$$S(x) := \big\{ y \in Y \,\big|\, 0 \in q(x, y) + Q(x, y) \big\}, \tag{6.73}$$

i.e., $x \mapsto S(x)$ is the solution map to the parametric variational system in (6.73). The latter is often labeled as the parameterized generalized equation (GE) if $Q(x, y) = N(y; G(x))$ for some $G\colon X \rightrightarrows Y$; cf. Section 3.3 with a bit different notation.

(i) Derive necessary optimality conditions for abstract MPECs given in form (6.72) under the most general assumptions on $f(x, y)$ and $S(x)$, and then deduce from them optimality conditions for (6.73) entirely via the initial data q, Q, G. Provide specifications of the obtained results in the case of finite-dimensional spaces. *Hint:* Reduce the models under consideration to those studied in Section 6.1 and , then apply to the necessary optimality conditions therein the corresponding results of generalized differential calculus. Compare it with [523, Section 5.2].

(ii) Under which assumptions the solution map $S(x)$ for the lower-level problem (6.30) can be equivalently written in the MPEC form (6.73)?

(iii) Investigate relationships between global and local solutions to optimistic bilevel programs and to MPECs in (6.72), (6.73) for the case where the lower-level program in (6.28) is convex. *Hint:* Consult [194] for problems with smooth data.

Exercise 6.49 (Value Function Constraint Qualification). Consider the class of optimistic bilevel programs defined by (6.72) with $\Omega = \mathbb{R}^n$ and the solution map to the lower-level problem given as

$$S(x) := \operatorname{argmin}\big\{\varphi(x, y)\big|\, y \in G\big\} \;\text{ for }\; G := \big\{y \in \mathbb{R}^m\big|\, g_i(y) \le 0,\; i = 1, \ldots, p\big\},$$

where $\varphi\colon \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is convex and continuously differentiable in y together with the functions $g_i : \mathbb{R}^m \to \mathbb{R}$. Following [341], introduce the parameterized sets

$$C(\nu) := \big\{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m\big|\, \varphi(x, y) - \mu(x) \le \nu\big\}, \quad \nu \in \mathbb{R},$$

involving the value function $\mu(x)$ from (6.33) for $G(x) = G$, and say that the *value function constraint qualification* (VFCQ) is satisfied at $(\bar{x}, \bar{y}) \in \operatorname{gph} S$ if the mapping $C\colon \mathbb{R} \rightrightarrows \mathbb{R}^n \times \mathbb{R}^m$ is calm at $(0, \bar{x}, \bar{y})$ as defined in Exercise 3.51.

(i) Verify that $S(x)$ can be equivalently written, under the assumptions made, in the MPEC form (6.73) with $q(x, y) = \nabla_y \varphi(x, y)$ and $Q(x, y) = N(y; G)$. *Hint:* Use the classical necessary and sufficient conditions in convex programming.

(ii) Show that if the bilevel program defined in this way has the uniform weak sharp minima (6.43) around the local solution pair $(\bar{x}, \bar{y})$, then VFCQ is satisfied at $(\bar{x}, \bar{y})$. Give an example that the reverse implication fails.

(iii) Verify that the validity of VFCQ at $(\bar{x}, \bar{y})$ ensures that the partial calmness property holds at this point, but not vice versa.

(iv) Assuming that the set G is bounded and that VFCQ is satisfied at $(\bar{x}, \bar{y})$, prove that the perturbation mapping

$$M(\nu) := \big\{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m\big|\, \nu \in \nabla_y \varphi(x, y) + N(y; G)\big\}, \quad \nu \in \mathbb{R}, \tag{6.74}$$

is *calm* at $(0, \bar{x}, \bar{y})$ as defined in Exercise 3.51, while the latter property is strictly weaker than VFCQ in the setting under consideration.

Hint: Consult [341] for the proofs of the results stated in (ii)–(iv).

Exercise 6.50 (Necessary Optimality Conditions for Optimistic Bilevel Programs Without Imposing Partial Calmness).

(i) Investigate the possibility of deriving necessary optimality conditions for the optimistic bilevel program (6.31) by applying the corresponding results of the Fritz John type from Section 6.1 to the equivalent nondifferentiable program (6.34).

(ii) With the usage of the necessary optimality conditions for problems in the general form (6.72) obtained in [523, Subsection 5.2.1] and expressed via the basic coderivative $S(x)$, while normal and mixed versions of it in Asplund spaces, derive their specifications for bilevel programming by evaluating the coderivatives of the solution map (6.30) to the lower-level problem. *Hint:* Consult [198] and the references therein for evaluating the basic coderivative of $S(x)$ in finite dimensions.

(iii) Following the approach of [341] developed for the optimistic bilevel programs with convex lower-level problems and the MPEC solution maps described in Exercise 6.49, derive necessary optimality conditions for *nonconvex* bilevel programs with replacing the partial calmness as in Theorem 6.21 by the *calmness property* of the perturbation mapping (6.74) at $(0, \bar{x}, \bar{y})$ in the sense of Exercise 3.51.

(iv) Compare the results of [341] with those presented in Section 6.3 in the same smooth and convex setting, and then investigate the possibility of extending the approach of [341] to the more general frameworks studied above.

Exercise 6.51 (Two-Level Value Function in Bilevel Programming). Consider the cost function $f_{opt}(x)$ in (6.31) with $S(x)$ taken from (6.30); $f_{opt}(x)$ is labeled as the *two-level optimal value function* in bilevel programming [198].

(i) Evaluate the basic and singular subdifferentials of f_{opt}, and then establish verifiable conditions for the local Lipschitz continuity of this function around a local solution to the optimistic bilevel program (6.31) by using Corollary 4.3 and the Lipschitz-like property of $S(x)$ via the coderivative criterion from Theorem 3.3.

(ii) Apply (i) to deriving necessary optimality conditions in the original optimistic model (6.31), which may not be locally equivalent to model (6.34) studied above; see Proposition 6.13. Compare it with the results presented in Section 6.3.

(iii) Implement this approach to justifying various types of *stationarity* in optimistic bilevel programming as formulated, e.g., in [198].

Hint: Consult [198] for the results, proofs, and additional material.

Exercise 6.52 (Necessary Optimality Conditions in Pessimistic Bilevel Programming). Consider the class of pessimistic bilevel programs (6.32) with the cost function $f_{pes}(x)$ under the same constraints as in (6.31).

(i) Employing the results of Exercise 6.51(i) on the local Lipschitz continuity Lipschitz continuity of f_{opt} with taking into account that $f_{pes} = -f_{opt}$, derive necessary optimality and stationarity conditions for (6.32) from those in Exercise 6.51(ii,iii).

(ii) Derive upper subdifferential conditions for pessimistic bilevel programs from the corresponding results of Section 6.1.

Hint: Consult [199] for more details on both (i) and (ii).

Exercise 6.53 (Multiobjective Approach to Bilevel Programming). Given an upper-level objective function $f\colon X \times Y \to \mathbb{R}$ and the solution map $S\colon X \rightrightarrows Y$ to the lower-level problem as described in (6.30) in the cases of finite-dimensional or infinite-dimensional spaces X and Y, consider the set-valued mapping $F\colon X \rightrightarrows \mathbb{R}$ given in the composition form $F(x) := f(x, S(x))$ for $x \in X$, and rewrite the upper-level problem of bilevel programming as follows:

$$\text{minimize} \quad F(x) \quad \text{subject to } x \in \Omega \tag{6.75}$$

with respect to the standard order on $\mathbb{R}$, where the upper-level constraint set Ω in (6.75) can be represented as or added by some other types of constraints (functional, operator, complementarity, equilibrium, etc.).

(i) Applying to (6.75) the coderivative and subdifferential types of necessary optimality conditions obtained in Section 9.4 for multiobjective optimization together with the coderivatice/subdifferential chain rules for the composition $f(x, S(x))$ and then evaluating the coderivative of S, derive necessary optimality conditions for bilevel programs in terms of their initial data.

(ii) Specify the results obtained in this way for optimistic and pessimistic models of bilevel programming, and compare it with those derived and discussed above.

6.5 Commentaries to Chapter 6

Section 6.1. Deriving *necessary optimality conditions* for optimization problems with *nonsmooth data* has been among early motivations to develop constructions and machinery of modern variational analysis and generalized differentiation. Nonsmoothness naturally appeared in the original framework of *optimal control* problems starting from the mid-1950s; see [645]. A simple albeit typical problem of this type was formulated as minimizing a cost function $\varphi(x(1))$ depending on the right endpoints of trajectories for the ODE control system

$$\frac{dx}{dt} = f(x, u), \;\; x(0) = x_0 \in \mathbb{R}^n, \;\; u(t) \in U, \;\; t \in T := [0, 1] \tag{6.76}$$

over measurable (or piecewise continuous) control functions $u(t)$ on T with values belonging to the prescribed closed set $U \subset \mathbb{R}^m$. Since the feasible control region U may be arbitrary (a typical case is when U consists of finitely many points as in systems of automatic control), the formulated optimal control problem can be treated as an optimization problem with irregular geometric constraints regardless of smoothness assumptions on the given functions φ and f. Furthermore, this problem can be equivalently rewritten in form (6.1) studied above, where $\Omega \subset \mathbb{R}^n$ is the reachable set of trajectory endpoints generated by feasible controls in (6.76). Optimal control theory from the very beginning, while revolving around different proofs and extensions of the Pontryagin maximum principle, has been seeking appropriate techniques to deal with this *intrinsic nonsmoothness*. It was a major driving force to develop modern forms of variational analysis that invoke generalized differentiation.

Another remarkable class of intrinsically nonsmooth optimization problems was also discovered in the mid-1950s and named *dynamic programming* by Bellman [77]. His "Principle of Optimality" led him to the so-called Bellman equation for the corresponding optimal value function while *assuming* the smoothness of the latter. Since this assumption fails even in simple examples, the Bellman equation plays just a heuristic role in some practical problems but generally may result in wrong conclusions; see, e.g., [645]. Comprehensive theories of the Hamilton-Jacobi-Bellman and related PDE equations with numerous applications have been developed in the frameworks of *viscosity* and *minimax* solutions by using tools of generalized differentiation; see the books [66, 136, 167, 268, 698] and the references therein.

In fact, intrinsic (often hidden) nonsmoothness already appears at the very fundamental level of modern optimization for problems with *inequality constraints*

$$\varphi_i(x) \le 0, \quad i \in I, \tag{6.77}$$

where the index set I may be finite (while fairly large as, i.e., in linear programming) or infinite as in semilinear programming studied below in Chapters 7 and 8. It is well recognized that the development of efficient machinery for studying and solving optimization problems with inequality constraints is probably the most monumental contribution of mathematical optimizers to society. Saying this, we observe that the inequality constraints (6.77) closely relate to nonsmoothness even in the case of finitely many linear functions φ_i. Geometrically it is manifested by the vertices of convex polyhedra that are described by (6.77) and play a crucial role in the groundbreaking simplex algorithm to solve linear programs. Analytically nonsmoothness is revealed via the equivalent replacement of (possibly great many) inequality constraints in (6.77) by the *single one*

$$\phi(x) := \max\left\{\varphi_i(x)\middle|\ i \in I\right\} \le 0$$

given by the *maximum function* $\phi(x)$, which is nondifferentiable even in the case of two linear functions on the real line: $\phi(x) = \max\{x, -x\} = |x|$. As the reader can see in this book, among other numerous publications, maximum/supremum functions and their generalized differentiation are highly important for the study and applications of various types of optimization and equilibrium problems.

To complete these discussions on the role of nonsmoothness in optimization, observe that nondifferentiable functions unavoidably arise while applying *perturbation* and *approximation techniques*, which are central in modern variational analysis, to problems with smooth initial data. Also powerful *variational principles* (notably the Ekeland one) lead us to considering nonsmooth optimization problems.

Now we comment on some specific results presented in Section 6.1 and the corresponding exercises from Section 6.4. *Lower subdifferential optimality conditions* in terms of basic normals and subgradients were derived by using the method of metric approximations in the original publications by the author [502, 503, 504, 507] and those joint with Kruger [439, 440, 528]. Their infinite-dimensional extensions were given in [441, 426, 430] for problems in Fréchet smooth spaces under certain Lipschitzian assumptions and in [516, 523] for the case of Asplund spaces under SNC-type requirements imposed on the sets, mappings, and functions in question. Clarke's version (6.21) of the generalized Lagrange multiplier rule was obtained in [164, 165], and Warga's rule (6.22) was derived in [736, 737]. Other results in this direction in both Fritz John and KKT forms under various qualification conditions can be found in, e.g., [16, 84, 273, 326, 328, 366, 523, 678, 685] and the references therein.

Applications of necessary optimality conditions presume that optimal solutions exist. This is not always the case, especially in infinite dimensions. One of the primary motivations for developing of Ekeland's variational principle was to obtain the "almost stationarity" condition for "almost optimal" (suboptimal) solutions formulated in (2.24). More general (lower) *necessary suboptimality conditions* for problems of nonlinear and nondifferentiable programming presented in Exercise 6.35 are based on the *lower subdifferential variational principle* formulated in Exercise 2.39 and are taken from [523, 587], where the reader can find more discussions and references.

Upper subdifferential optimality conditions for *minimization* problems were initiated by the author [519] who obtained the results presented in Section 6.1 and their counterparts for other optimization problems in general Banach spaces; see also [523, Chapter 5]. As discussed in Remark 6.2, upper subdifferential conditions may have serious advantages over lower subdifferential ones provided that $\widehat{\partial}^+\varphi(\bar{x}) \neq \emptyset$. Various classes of such functions were discussed in [523, Subsection 5.5.4].

It is interesting to observe as in Proposition 6.3 that for problems of minimizing the *DC* (*difference of convex*) functions $\varphi_1(x) - \varphi_2(x)$, the upper subdifferential condition (6.3) reduces to the well-known one $\partial\varphi_2(\bar{x}) \subset \partial\varphi_1(\bar{x})$ as in [350]. Note that the class of DC functions as well as its specifications and modifications play an important role in various qualitative and quantitative issues of optimization including its global aspects and numerical algorithms; see, e.g., [203, 302, 311, 327, 329, 350, 487, 355] among many other publications. Problems of this type will be also studied in Chapter 7 below in the framework of semi-infinite programming.

Sections 6.2 and 6.3. *Bilevel programs* constitute a broad class of problems in hierarchical optimization that is very interesting mathematically and important in applications. We refer the reader to the book by Dempe [193] and more recent publications [177, 194, 195, 196, 197, 198, 199, 200, 201, 202, 341, 469, 540, 750, 763, 769, 764] for various versions in bilevel programming, different approaches to their study, and numerous applications. A characteristic feature of bilevel programs, which can be seen in all of their versions, reformulations, and transformations, is *intrinsic nonsmoothness* that creates serious theoretical and algorithmic challenges. Furthermore, it has been well recognized that standard constraint qualifications in nonlinear and nondifferentiable programming fail to fulfill in bilevel optimization.

The *optimistic* version is by far the most investigated one in bilevel programming, while there are many unsolved theoretical questions therein, not even mentioning numerical algorithms. Among several approaches to deriving necessary optimality conditions in optimistic bilevel programs, we present in Sections 6.2–6.3 the *value function approach*, which was initiated by Outrata [619] for a particular bilevel optimization model. This approach explicitly manifests nonsmoothness in bilevel programming via the nondifferentiable lower-level value function (6.33).

The value function approach to optimistic bilevel programs was greatly developed by Ye and Zhu [748], who introduced the *partial calmness* condition that allowed them to reduce bilevel programs to nonsmooth single-level ones via penalization. Combining it with Clarke's generalized differentiation, they derived in [748] necessary optimality conditions for bilevel programs in terms of their initial data.

In this book we mainly follow the papers [195, 540] and further develop the value function approach by employing our basic tools of generalized differentiation to express optimality conditions for nondifferentiable programs from Section 6.1 and then to evaluate basic subgradients of marginal/optimal value functions via the results of Section 4.1. Such a device allows us to essentially improve necessary optimality conditions for optimistic bilevel programs obtained in [748] and other publications. Note the importance of the rather surprising *difference rule* for regular subgradients from Lemma 6.22 established in [546] by using the smooth variational description of regular subgradients in Theorem 1.27.

The partial calmness assumption from Definition 6.14 plays an essential role in the value function approach to bilevel programming. Although it is satisfied in many important settings, it may fail in rather simple nonlinear examples; see the discussions above as well as the results in [133, 201, 198, 748, 749]. A sufficient (while far from being necessary) condition for the validity of partial calmness was introduced by Ye and Zhu [748] under the name of "uniformly weak sharp minima," which could be seen as a version of sharp minima by Polyak [643, 644] and weak sharp minima by Ferris [264]. In contrast to the latter two notions, which have been well investigated and applied in the literature (see, e.g., [132, 133, 237, 335, 462, 495, 546, 608, 697, 744, 782, 785]), uniform weak sharp minima have drawn much less attention. We refer the reader to [133, 327, 744, 748, 749] for some efficient conditions ensuring the validity of the uniform weak sharp minimum estimate (6.43) and also to the discussions right before Proposition 6.18, which seems to be new.

There are several approaches to deriving necessary optimality and stationary conditions that don't employ partial calmness; see [51, 198, 199, 200, 201, 341, 750, 763]. We particularly emphasize remarkable developments by Henrion and Surowiec [341] for the class of optimistic bilevel programs with $\mathcal{C}^2$-smooth data and *convex* lower-level problems, where the solution map to the lower-level problem can be equivalently rewritten in the *MPEC form* (6.73) with $q(x, y) = \nabla_y \varphi(x, y)$ and $Q(x, y) = N(y; G)$; see Exercise 6.49(i). They replace the partial calmness assumption by the weaker *calmness* property of the perturbation mapping (6.74) in the sense defined in Exercise 3.51. Imposing in addition the *constant rank constraint qualification* in the lower-level problem (see [477, 499] for more details about the latter notion), Henrion and Surowiec derive necessary optimality conditions (more precisely, M(ordukhovich)-stationarity conditions) for optimistic bilevel programs, which have serious advantages in comparison with the corresponding results of [195] in such settings. The reader may find more information about MPECs and their applications in the fundamental monographs [482, 624] and the subsequent publications [3, 78, 267, 314, 338, 341, 346, 290, 523, 620, 623, 684, 745, 746, 780] among other works with numerous references therein. See, in particular, the papers by Outrata [620] and Scheel and Scholtes [684] for introducing various notions of stationarity for MPECs, which have been similarly developed later in bilevel programming. Note to this end that, although MPECs [482, 624] and bilevel programs have many things in common, these two classes of optimization problems are essentially different in general; see the papers by Dempe and Dutta [194] and by Dempe and Zemkoho [202] for various results and comprehensive discussions.

Section 6.4. This section contains exercises of different levels of difficulties on necessary optimality conditions in nonsmooth optimization and bilevel programming with hints and references when needed. At the same time, we present here some *challenging* and largely *open questions* concerning various issues of bilevel optimization. They include Exercise 6.38(i), Exercise 6.39(ii,iii), and Exercise 6.40(iv) on uniform weak sharp minima, Exercise 6.43(v) on the usage of the symmetric subdifferential of marginal functions for deriving necessary optimality conditions for bilevel programs, Exercise 6.50 on deriving necessary optimality conditions for optimistic bilevel programs without the partial calmness assumption by using the approaches described therein, Exercise 6.52 and beyond on deriving necessary optimality and stationarity conditions for pessimistic bilevel programs that are considerably underinvestigated in the literature, and Exercise 6.53 on developing a new multiobjective optimization approach to bilevel programs by using the procedure described therein.

Chapter 7
Semi-infinite Programs with Some Convexity

This and the next chapters of the book contain mainly some recent applications of the constructions and results of variational analysis and generalized differentiation presented above, as well as new developments required for such applications, to a remarkable class of optimization problems unified under the name of *semi-infinite programming* (SIP). We also use the abbreviations "SIP" for a particular semi-infinite program and "SIPs" as plural. The SIP terminology comes from the fact that originally this class of optimization problems concerned minimizing real-valued functions on *finite*-dimensional spaces subject to *infinitely* many inequality constraints usually indexed by a compact set. Over the years, the theory and applications of SIP have been evolved to include optimization problems with noncompact index sets and on infinite-dimensional spaces. Sometimes SIPs with infinite-dimensional decision spaces are labeled as problems of "infinite programming," while here we prefer to use the conventional SIP terminology regardless of the decision space dimension. As seen, the underlying style in the previous chapters was to present major results in finite-dimensional spaces and then to discuss infinite-dimensional extensions only in exercise and commentary sections. In contrast, the standing framework of this and the next chapters is, unless otherwise stated, the general *Banach space* setting. The main reasons for it are as follows:

(1) Due to their essence, SIPs always contain an infinite-dimensional part and require the usage of infinite-dimensional analysis for their investigation.

(2) The major results obtained below are formulated exactly in the same way in both cases of finite-dimensional and Banach decision spaces.

(3) Many practically meaningful models can be described as SIPs with infinite-dimensional decision spaces. In particular, this is the case of the water resource optimization problem, which is formulated and solved in Section 7.2 by using the necessary optimality conditions obtained therein.

B. S. Mordukhovich, *Variational Analysis and Applications*, Springer Monographs in Mathematics, https://doi.org/10.1007/978-3-319-92775-6_7

7.1 Stability of Infinite Linear Inequality Systems

In this section, we study the sets of feasible solutions to SIPs described by the parameterized *infinite systems of linear inequalities*

$$\mathcal{F}(p) := \big\{x \in X \big|\, \langle a_t^*, x\rangle \le b_t + p_t,\ t \in T\big\}, \quad p = (p_t)_{t\in T}, \tag{7.1}$$

with an *arbitrary index set* T, where $x \in X$ is a *decision* variable belonging to a *Banach* space X and where $p = (p_t)_{t\in T} \in P$ is a functional *parameter* taking values in the prescribed Banach space P of perturbations specified below. The data of (7.1) are given as follows:

• $a_t^* \in X^*$ are fixed for all $t \in T$. We use the same notation for the given norm $\|\cdot\|$ on X and the corresponding dual norm on X^* defined by

$$\|x^*\| := \sup\big\{\langle x^*, x\rangle \big|\, \|x\| \le 1\big\}, \quad x^* \in X^*.$$

• $b_t \in \mathbb{R}$ are fixed for all $t \in T$. We identify the collection $\{b_t|\ t \in T\}$ with the real-valued function $b\colon T \to \mathbb{R}$.

• $p_t = p(t) \in \mathbb{R}$ for all $t \in T$. These functional parameters $p\colon T \to \mathbb{R}$ are our varying perturbations, which are taken from the Banach parameter space $P := l^\infty(T)$ of all bounded functions on T with the supremum norm $\|p\|_\infty := \sup\big\{|p(t)|\ \big|\ t \in T\big\}$. When T is compact and $p(\cdot)$ are restricted to be continuous on T, the parameter space P reduces to $\mathcal{C}(T)$.

It is obvious that the space $l^\infty(T)$ is never finite-dimensional when the index set T is infinite. Moreover, in the infinite-dimensional case, the space $l^\infty(T)$ is *never Asplund*; see [638, Example 1.21].

The primary goal of this section is to calculate the *coderivative* of the set-valued mapping $\mathcal{F}$ defined in (7.1) as well as the *coderivative norm* of $\mathcal{F}$ at the reference point entirely in terms of the initial data of (7.1). Based on this, we derive here a complete *coderivative characterization* of the *Lipschitz-like property* of $\mathcal{F}$ in the form identical to the finite-dimensional setting of Chapter 3. Furthermore, the obtained coderivative calculation is the key of deriving necessary optimality conditions for SIPs with linear inequality constraints of type (7.1) in Section 7.2 and then in turn becomes crucial to investigate SIPs described by convex inequalities and the like in the subsequent Section 7.3.

Recall that the coderivative of any mapping $F\colon X \rightrightarrows Y$ between Banach spaces studied in this and the next chapters is considered in the usual "normal" sense as in finite dimensions. This means that, given any $(\bar{x}, \bar{y}) \in \operatorname{gph} F$, the coderivative of F at $(\bar{x}, \bar{y})$ is the mapping $F\colon Y^* \rightrightarrows X^*$ defined by

$$D^*F(\bar{x}, \bar{y})(y^*) := \big\{x^* \in X^* \big|\, (x^*, -y^*) \in N\big((\bar{x}, \bar{y}); \operatorname{gph} F\big)\big\} \tag{7.2}$$

for $y^* \in Y^*$ via the corresponding normal cone to the graph of F at $(\bar{x}, \bar{y})$.

7.1.1 Lipschitz-Like Property and Strong Slater Condition

Since we are in the general Banach space setting, the symbol w^*-lim signifies here the weak* *topological* limit in the dual space in question. This corresponds to the convergence of *nets* denoted usually by $\{x^*_\nu\}_{\nu\in\mathcal{N}}$. In the case of sequences, we replace the symbol $\mathcal{N}$ by the standard natural series notion $\mathbb{N} = \{1, 2, \ldots\}$. For an arbitrary index set T, denote by $\mathbb{R}^T$ the *product space* of $\lambda = (\lambda_t|\ t \in T)$ with $\lambda_t \in \mathbb{R}$ for all $t \in T$. Finally, let $\mathbb{R}^{(T)}$ be the collection of multipliers $\lambda \in \mathbb{R}^T$ such that $\lambda_t \neq 0$ for *finitely many* $t \in T$, and let $\mathbb{R}^{(T)}_+$ be the *positive cone* in $\mathbb{R}^{(T)}$ defined by

$$\mathbb{R}^{(T)}_+ := \big\{\lambda \in \mathbb{R}^{(T)}\big|\ \lambda_t \geq 0 \ \text{ for all } \ t \in T\big\}. \tag{7.3}$$

Note also that throughout this chapter, the symbol "cone Ω" stands for the *convex* conic hull of the set in question.

Let us now recall a well-recognized *qualification condition* for SIPs with infinite linear inequality constraints and then show that it provides, along with other conditions, an equivalent description of the Lipschitz-like property of the constraint mapping $\mathcal{F}$ from (7.1).

Definition 7.1 (Strong Slater Condition). *We say that the infinite linear inequality system* (7.1) *satisfies the* STRONG SLATER CONDITION (SSC) *at* $p = (p_t)_{t\in T}$ *if there exists* $\widehat{x} \in X$ *such that*

$$\sup_{t\in T}\big[\langle a^*_t, \widehat{x}\rangle - b_t - p_t\big] < 0. \tag{7.4}$$

Furthermore, every point $\widehat{x} \in X$ *satisfying condition* (7.4) *is a* STRONG SLATER POINT *for system* (7.1) *at* $p = (p_t)_{t\in T}$.

Define further the parametric *characteristic sets*

$$C(p) := \text{co}\big\{(a^*_t, b_t + p_t)\big|\ t \in T\big\}, \quad p \in l^\infty(T), \tag{7.5}$$

and suppose without loss of generality that $\bar{p} = 0 \in l^\infty(T)$ is the designated *nominal parameter*. First, we verify the following equivalences.

Theorem 7.2 (Equivalent Descriptions of the Lipschitz-Like Property for Infinite Linear Systems). *Given* $p \in \text{dom}\,\mathcal{F}$ *for* (7.1) *in the Banach decision space* X, *the following properties are equivalent:*

(i) $\mathcal{F}$ *is Lipschitz-like around* (p, x) *for all* $x \in \mathcal{F}(p)$.

(ii) $p \in \text{int}(\text{dom}\,\mathcal{F})$.

(iii) $\mathcal{F}$ *satisfies the strong Slater condition at* p.

(iv) $(0, 0) \notin \text{cl}^*C(p)$ *via the characteristic set in* (7.5).

Finally, the boundedness of $\big\{a^*_t \mid t \in T\big\}$ *ensures the equivalence of* (i)–(iv) *to:*

(v) *there exists* $\widehat{x} \in X$ *such that* $(p, \widehat{x}) \in \text{int}(\text{gph}\,\mathcal{F})$.

Proof. The equivalence between (i) and (ii) is a consequence of the Robinson-Ursescu theorem and the equivalence between the Lipschitz-like property of the convex-graph mapping $\mathcal{F}$ and the metric regularity/covering properties of its inverse; see Theorem 3.2 and Corollary 3.6 together with the corresponding exercises and commentaries in Sections 3.4 and 3.5.

To verify implication (iii)$\Rightarrow$(ii), suppose that $\widehat{x}$ is a strong Slater point for system (7.1) at p and find $\vartheta > 0$ such that

$$\langle a_t^*, \widehat{x}\rangle - b_t - p_t \le -\vartheta \quad \text{for all} \quad t \in T.$$

Then it is obvious that for any $q \in l^\infty(T)$ with $\|q\| < \vartheta$, we have $\widehat{x} \in \mathcal{F}(p + q)$. Therefore $p + q \in \operatorname{dom}\mathcal{F}$, and thus (ii) holds. To justify further the converse implication (ii)$\Rightarrow$(iii), take $p \in \operatorname{int}(\operatorname{dom}\mathcal{F})$, and then get $p + q \in \operatorname{dom}\mathcal{F}$ provided that $q_t = -\vartheta$ as $t \in T$ and that $\vartheta > 0$ is sufficiently small. Thus every $\widehat{x} \in \mathcal{F}(p+q)$ is a strong Slater point for the infinite system (7.1) at p.

Next we show that (iii)$\Rightarrow$(iv). Arguing by contradiction, suppose that $(0, 0) \in \operatorname{cl}^* C(p)$. Then there exists a net $\{\lambda_\nu\}_{\nu\in\mathcal{N}} \in \mathbb{R}_+^{(T)}$ satisfying the equality $\sum_{t\in T} \lambda_{t\nu} = 1$ for all $\nu \in \mathcal{N}$ and the limiting condition

$$(0, 0) = w^*\text{-}\lim_{\nu\in\mathcal{N}} \sum_{t\in T} \lambda_{t\nu}\big(a_t^*, b_t + p_t\big). \tag{7.6}$$

If $\widehat{x}$ is a strong Slater point for system (7.1) at p, we find $\vartheta > 0$ such that

$$\langle a_t^*, \widehat{x}\rangle - b_t - p_t \le -\vartheta \quad \text{for all} \quad t \in T.$$

Then condition (7.6) leads us to the contradiction

$$0 = \langle 0, \widehat{x}\rangle + 0 \cdot (-1) = \lim_{\nu\in\mathcal{N}} \sum_{t\in T} \lambda_{t\nu}\big(\langle a_t^*, \widehat{x}\rangle + (b_t + p_t) \cdot (-1)\big) \le -\vartheta,$$

which thus justifies (iii)$\Rightarrow$(iv). To verify the converse implication (iv)$\Rightarrow$(iii), we employ the dual description of the *consistency* in (7.1) given by

$$p \in \operatorname{dom}\mathcal{F} \iff (0, -1) \notin \operatorname{cl}^*\operatorname{cone}\big\{(a_t^*, b_t + p_t)\big|\ t \in T\big\}, \tag{7.7}$$

which is discussed in Exercise 7.71 and the commentaries in Section 7.7. Then the classical strong separation theorem gives us $(0, 0) \ne (v, \alpha) \in X \times \mathbb{R}$ with

$$\big\langle a_t^*, v\big\rangle + \alpha(b_t + p_t) \le 0,\ \ t \in T, \quad \text{and} \quad \langle 0, v\rangle + (-1)\alpha = -\alpha > 0. \tag{7.8}$$

Using (iv), we get $(0, 0) \ne (z, \beta) \in X \times \mathbb{R}$ and $\gamma \in \mathbb{R}$ for which

$$\langle a_t^*, z\rangle + \beta(b_t + p_t) \le \gamma < 0 \quad \text{whenever} \quad t \in T. \tag{7.9}$$

Consider now the combination $(u, \eta) := (z, \beta) + \lambda(v, \alpha)$, and select $\lambda > 0$ such that $\eta < 0$. Defining $\widehat{x} := -\eta^{-1}u$, we deduce from (7.8) and (7.9) that

$$\langle a_t^*, \widehat{x}\rangle - b_t - p_t = -\eta^{-1}\big(\langle a_t^*, u\rangle + \eta(b_t + p_t)\big) \le -\eta^{-1}\gamma < 0.$$

Hence $\widehat{x}$ is a strong Slater point for system (7.1) at p, i.e., (iii) holds.

It remains to consider condition (v). It is easy to see that (v) always implies (iv) and so the other conditions of the theorem. Suppose now that the set $\{a_t^* \mid t \in T\}$ is bounded, and show that (iii) implies (v). Select $M \ge 0$ such that $\|a_t^*\| \le M$ for every $t \in T$, and take $\widehat{x} \in X$ satisfying (7.4). Denote

$$\gamma := -\sup_{t\in T}\big[\langle a_t^*, \widehat{x}\rangle - b_t - p_t\big] > 0$$

and consider any pair $(p', u) \in l^\infty(T) \times X$ such that

$$\|u\| \le \eta := \gamma/(M+1) > 0 \ \text{ and } \ \|p'\| \le \eta.$$

It is easy to see that for such (p', u) and every $t \in T$, we have

$$\langle a_t^*, \widehat{x} + u\rangle - b_t - p_t - p_t' \le -\gamma + M\|u\| + \|p'\| \le \eta(M+1) - \gamma = 0,$$

and so $(p + p', \widehat{x} + u) \in \operatorname{gph}\mathcal{F}$. Thus $(p, \widehat{x}) \in \operatorname{int}(\operatorname{gph}\mathcal{F})$, which verifies implication (iii)⇒(v) and completes the proof of the theorem. △

7.1.2 Coderivatives for Parametric Infinite Linear Systems

In this subsection, we calculate the coderivative $D^*\mathcal{F}(0, \bar{x})$ as in (7.2) of the parametric infinite system (7.1) at the reference point $(0, \bar{x})$ and also its norm $\|D^*\mathcal{F}(0, \bar{x})\|$ entirely via the initial data of (7.1). Recall that the dual space $l^\infty(T)^*$ to the parameter space in (7.1) is isometric to the space $ba(T)$ of all the bounded and additive measures $\mu(\cdot)$ on subsets of T with the norm

$$\|\mu\| := \sup_{A\subset T}\mu(A) - \inf_{B\subset T}\mu(B).$$

In what follows a dual element $p^* \in l^\infty(T)^*$ is identified with the corresponding measure $\mu \in ba(T)$ satisfying the canonical duality relationship

$$\langle \mu, p\rangle = \int_T p_t\,\mu(dt), \quad p = (p_t)_{t\in T}.$$

To proceed further, we need the following extension of the classical *Farkas lemma* to the case of infinite linear inequality systems; see Exercise 7.73 and the corresponding commentaries in Section 7.7.

Proposition 7.3 (Extended Farkas Lemma for Infinite Linear Inequalities). *Let $p \in \operatorname{dom}\mathcal{F}$ for the infinite system (7.1), and let $(x^*, \alpha) \in X^* \times \mathbb{R}$. The following assertions are equivalent:*

(i) *We have $\langle x^*, x\rangle \le \alpha$ whenever $x \in \mathcal{F}(p)$, i.e.,*

$$\big[\langle a_t^*, x\rangle \le b_t + p_t \;\; \textit{for all} \;\; t \in T\big] \Longrightarrow \big[\langle x^*, x\rangle \le \alpha\big].$$

(ii) *The pair* (x^*, α) *satisfies the inclusion*

$$(x^*, \alpha) \in \text{cl}^*\text{cone}\Big[\Big\{(a_t^*, b_t + p_t)\Big|\; t \in T\Big\} \cup \{(0, 1)\}\Big] \;\; \textit{with} \;\; 0 \in X^*.$$

Using Proposition 7.3, we first describe the normal cone to the graph

$$\text{gph}\,\mathcal{F} = \Big\{(p, x) \in l^\infty(T) \times X\Big|\; \langle a_t^*, x\rangle \le b_t + p_t \;\; \text{for all} \;\; t \in T\Big\}$$

at the reference point $(0, \bar{x}) \in \text{gph}\,\mathcal{F}$. Recall that δ_t stands for the classical *Dirac function/measure* at $t \in T$ satisfying

$$\langle \delta_t, p\rangle = p_t \;\; \text{as} \;\; t \in T \;\; \text{for} \;\; p = (p_t)_{t\in T} \in l^\infty(T). \tag{7.10}$$

Proposition 7.4 (Graphical Normals for Infinite Linear Systems). *Let* $(0, \bar{x}) \in \text{gph}\,\mathcal{F}$ *for the mapping* $\mathcal{F}$ *from* (7.1)*, and let* $(p^*, x^*) \in l^\infty(T)^* \times X^*$*. Then we have* $(p^*, x^*) \in N((0, \bar{x}); \text{gph}\,\mathcal{F})$ *if and only if*

$$\big(p^*, x^*, \langle x^*, \bar{x}\rangle\big) \in \text{cl}^*\text{cone}\Big[\Big\{(-\delta_t, a_t^*, b_t)\Big|\; t \in T\Big\} \cup \{(0, 0, 1)\}\Big], \tag{7.11}$$

where $0 \in l^\infty(T)^*$ *and* $0 \in X^*$ *stand for the first and second entry of the last triple, respectively. Furthermore, the inclusion* $(p^*, x^*) \in N\left((0, \bar{x})\,; \text{gph}\,\mathcal{F}\right)$ *implies that* $p^* \le 0$ *in the space* ba (T)*, i.e.,* $p^*(A) \le 0$ *for all* $A \subset T$*.*

Proof. It is easy to see that

$$\text{gph}\,\mathcal{F} = \Big\{(p, x) \in l^\infty(T) \times X\Big|\; \langle a_t^*, x\rangle - \langle \delta_t, p\rangle \le b_t \;\; \text{for all} \;\; t \in T\Big\},$$

and therefore we have $(p^*, x^*) \in N((0, \bar{x}); \text{gph}\,\mathcal{F})$ if and only if

$$\langle p^*, p\rangle + \langle x^*, x\rangle \le \langle x^*, \bar{x}\rangle \;\; \text{for every} \;\; (p, x) \in \text{gph}\,\mathcal{F}. \tag{7.12}$$

Employing now the equivalence between (i) and (ii) in Proposition 7.3, we conclude that $(p^*, x^*) \in N((0, \bar{x}); \text{gph}\,\mathcal{F})$ if and only if inclusion (7.11) holds.

To justify the last statement of the proposition, for every set $A \subset T$, consider its *characteristic function* $\chi_A\colon T \to \{0, 1\}$ defined by

$$\chi_A(t) := \begin{cases} 1 & \text{if} \;\; t \in A, \\ 0 & \text{if} \;\; t \notin A. \end{cases}$$

It is obvious that the inclusion $(p, x) \in \text{gph}\,\mathcal{F}$ implies that $(p + \lambda\chi_A, x) \in \text{gph}\,\mathcal{F}$ for each $\lambda > 0$. Replacing now in (7.12) the pair (p, x) by $(p + \lambda\chi_A, x)$, dividing both sides of the inequality by λ, and then letting $\lambda \to \infty$ give us

$$\big\langle p^*, \chi_A\big\rangle = \int_T \chi_A(t)\, p^*(dt) = p^*(A) \le 0,$$

which completes the proof of the proposition. △

The representation of graphical normals obtained in Proposition 7.4 is crucial to calculate the coderivative of $D^*\mathcal{F}(0,\bar{x})$ defined via the normal cone to the gph $\mathcal{F}$ at $(0,\bar{x})$ according to (7.2).

Theorem 7.5 (Coderivative Calculation). *Given $\bar{x} \in \mathcal{F}(0)$ for the infinite system (7.1), we have that $p^* \in D^*\mathcal{F}(0,\bar{x})(x^*)$ if and only if*

$$\big(p^*, -x^*, -\langle x^*, \bar{x}\rangle\big) \in \mathrm{cl}^*\mathrm{cone}\big\{(-\delta_t, a_t^*, b_t)\big|\ t \in T\big\}. \tag{7.13}$$

Proof. It follows from the coderivative definition and Proposition 7.4 that $p^* \in D^*\mathcal{F}(0,\bar{x})(x^*)$ if and only if

$$\big(p^*, -x^*, -\langle x^*, \bar{x}\rangle\big) \in \mathrm{cl}^*\mathrm{cone}\big[\big\{(-\delta_t, a_t^*, b_t)\big|\ t \in T\big\} \cup \{(0,0,1)\}\big]. \tag{7.14}$$

To justify the coderivative representation claimed in the theorem, we need to show that inclusion (7.14) yields in fact the "smaller" one in (7.13). Assuming indeed that (7.14) holds, we find by (7.14) some nets $\{\lambda_\nu\}_{\nu\in\mathcal{N}} \subset \mathbb{R}_+^{(T)}$ and $\{\gamma_\nu\}_{\nu\in\mathcal{N}} \subset \mathbb{R}_+$ satisfying the limiting relationship

$$\big(p^*, -x^*, -\langle x^*, \bar{x}\rangle\big) = w^*\text{-}\lim_{\nu\in\mathcal{N}}\Big(\sum_{t\in T}\lambda_{t\nu}(-\delta_t, a_t^*, b_t) + \gamma_\nu(0,0,1)\Big), \tag{7.15}$$

where $\lambda_{t\nu}$ stands for the t-entry of $\lambda_\nu = (\lambda_{t\nu})_{t\in T}$ as $\nu \in \mathcal{N}$. The component structure of (7.15) tells us that

$$0 = \langle p^*, 0\rangle + \langle -x^*, \bar{x}\rangle + (-\langle x^*, \bar{x}\rangle)(-1) = \lim_{\nu\in\mathcal{N}}\Big(\sum_{t\in T}\lambda_{t\nu}\big(\langle a_t^*, \bar{x}\rangle - b_t\big) - \gamma_\nu\Big).$$

Taking into account the definition (7.3) of the positive cone $\mathbb{R}_+^{(T)}$ and that $(0,\bar{x})$ satisfies the infinite inequality system in (7.1), we get $\lim_{\nu\in\mathcal{N}}\gamma_\nu = 0$. This justifies (7.13) and thus completes the proof of the theorem. △

The next consequence of Theorem 7.5 is useful in what follows.

Corollary 7.6 (Limiting Coderivative Description). *If $p^* \in D^*\mathcal{F}(0,\bar{x})(x^*)$ in the setting of Theorem 7.5, then there is a net $\{\lambda_\nu\}_{\nu\in\mathcal{N}} \subset \mathbb{R}_+^{(T)}$ with*

$$\sum_{t\in T}\lambda_{t\nu} \to \|p^*\| = -\langle p^*, e\rangle, \quad \sum_{t\in T}\lambda_{t\nu}a_t^* \xrightarrow{w^*} -x^*, \quad \textit{and} \quad \sum_{t\in T}\lambda_{t\nu}b_t \to -\langle x^*, \bar{x}\rangle.$$

Proof. Theorem 7.5 gives us a net $\{\lambda_\nu\}_{\nu\in\mathcal{N}} \subset \mathbb{R}_+^{(T)}$ such that

$$\sum_{t\in T}\lambda_{t\nu}\delta_t \xrightarrow{w^*} -p^*, \quad \sum_{t\in T}\lambda_{t\nu}a_t^* \xrightarrow{w^*} -x^*, \quad \text{and} \quad \sum_{t\in T}\lambda_{t\nu}b_t \to -\langle x^*, \bar{x}\rangle.$$

This readily implies the relationships

$$\left\langle \sum_{t\in T} \lambda_{t\nu}\delta_t, e \right\rangle = \sum_{t\in T} \lambda_{t\nu} \to \left\langle p^*, -e \right\rangle =: \lambda \in [0, \infty)\,.$$

Since the dual norm on X^* is w^*-lower semicontinuous, we have

$$\|p^*\| \le \liminf_{\nu\in\mathcal{N}} \left\| \sum_{t\in T} \lambda_{t\nu}\delta_t \right\| \le \liminf_{\nu\in\mathcal{N}} \sum_{t\in T} \lambda_{t\nu} = \lambda.$$

Furthermore, it follows from the norm definition that

$$\|p^*\| = \sup_{\|p\|\le 1} \left\langle p^*, p \right\rangle \ge \left\langle p^*, -e \right\rangle = \lambda,$$

which yields $\|p^*\| = -\langle p^*, e\rangle$ and thus completes the proof. △

Now we proceed with the exact calculation of the *coderivative norm*

$$\|D^*\mathcal{F}(0,\bar{x})\| := \sup\left\{\|p^*\| \;\middle|\; p^* \in D^*\mathcal{F}(0,\bar{x})(x^*),\ \|x^*\| \le 1\right\} \tag{7.16}$$

entirely via the initial data of the infinite linear inequality system (7.1). A part of our analysis is the following proposition on properties of the characteristic set (7.5) at $p = 0$ in connection with strong Slater points of (7.1).

Proposition 7.7 (Strong Slater Points Relative to the Characteristic Set). *Given $\bar{x} \in \mathcal{F}(0)$, consider the set*

$$S := \left\{x^* \in X^* \middle|\, \left(x^*, \langle x^*, \bar{x}\rangle\right) \in \mathrm{cl}^* C(0)\right\} \tag{7.17}$$

built upon $C(0)$ from (7.5). The following assertions hold:

(i) *Let $\bar{x}$ be not a strong Slater point of the infinite system (7.1) at $p = 0$, and let the coefficient collection $\{a_t^* |\ t \in T\}$ be bounded in X^*. Then the set S in (7.17) is nonempty and w^*-compact in X^*.*

(ii) *Let $\bar{x}$ be a strong Slater point of (7.1) at $p = 0$. Then $S = \emptyset$ in (7.17).*

Proof. To justify (i), assume that $\bar{x}$ is not a strong Slater point for the infinite system (7.1) at $p = 0$. Then there is a sequence $\{t_k\}_{k\in\mathbb{N}} \subset T$ such that $\lim_k(\langle a_{t_k}^*, \bar{x}\rangle - b_{t_k}) = 0$. The boundedness of $\{a_t^* |\ t \in T\}$ implies by the classical Alaoglu-Bourbaki theorem that this set is relatively w^*-compact in X^*, i.e., there is a subnet $\{a_{t_\nu}^*\}_{\nu\in\mathcal{N}}$ of the latter sequence that w^*-converges to some element $u^* \in \mathrm{cl}^*\{a_t^* |\ t \in T\}$. This yields $\lim_{\nu\in\mathcal{N}} b_{t_\nu} = \langle u^*, \bar{x}\rangle$ and

$$\left(u^*, \langle u^*, \bar{x}\rangle\right) = w^*\text{-}\lim_{\nu\in\mathcal{N}} \left(a_{t_\nu}^*, b_{t_\nu}\right) \in \mathrm{cl}^* C(0),$$

which justifies the nonemptiness of the set S in (7.17).

To verify the w^*-compactness of S, observe that the boundedness of the set $A := \{a_t^* |\ t \in T\}$ implies this property of $\mathrm{cl}^*\mathrm{co}A$; the latter set is actually w^*-compact due to its automatic w^*-closedness. Note further that the set S in (7.17) is a preimage of $\mathrm{cl}^*C(0)$ under the w^*-continuous mapping $x^* \mapsto (x^*, \langle x^*, \bar{x}\rangle)$, and thus it is w^*-closed in X^*. Since S is a subset of $\mathrm{cl}^*\mathrm{co}A$, it is also bounded and hence w^*-compact in X^*. We are done with (i).

To proceed with (ii), let $\bar{x}$ be a strong Slater point of system (7.1) at $p = 0$, and let $\gamma := -\sup_{t\in T}\left\{\left\langle a_t^*, \bar{x}\right\rangle - b_t\right\}$. Then we have the inequality

$$\left\langle x^*, \bar{x}\right\rangle \le \beta - \gamma \quad \text{whenever} \quad \left(x^*, \beta\right) \in \mathrm{cl}^*C\,(0)\,,$$

which justifies (ii) and thus completes the proof of the proposition. △

Now we are ready to calculate the coderivative norm $\|D^*\mathcal{F}(0, \bar{x})\|$ entirely in terms of the given data of the infinite system (7.1) in Banach spaces.

Theorem 7.8 (Calculating the Coderivative Norm). *Let $\bar{x} \in \mathrm{dom}\,\mathcal{F}$ for the infinite system (7.1), which satisfies the strong Slater condition at $p = 0$. Then the following assertions hold under the boundedness of $\{a_t^* |\ t \in T\}$:*

(i) *If $\bar{x}$ is a strong Slater point for $\mathcal{F}$ at $p = 0$, then $\|D^*\mathcal{F}(0, \bar{x})\| = 0$.*

(ii) *If $\bar{x}$ is not a strong Slater point for $\mathcal{F}$ at $p = 0$, then the coderivative norm (7.16) is positive and is calculated by*

$$\|D^*\mathcal{F}(0, \bar{x})\| = \max\left\{\|x^*\|^{-1} \,\Big|\, \left(x^*, \langle x^*, \bar{x}\rangle\right) \in \mathrm{cl}^*C(0)\right\}. \tag{7.18}$$

Proof. To verify assertion (i), suppose that $\bar{x}$ is a strong Slater point for the system $\mathcal{F}$ at $p = 0$. It follows from the proof of implication (iii)⇒(v) in Theorem 7.2 that $(0, \bar{x}) \in \mathrm{int}(\mathrm{gph}\,\mathcal{F})$ and hence $N((0, \bar{x}); \mathrm{gph}\,\mathcal{F}) = \{(0, 0)\}$. Thus (i) follows from definitions of the coderivative and its norm.

To prove assertion (ii), take $x^* \in X^*$ such that $(x^*, \langle x^*, \bar{x}\rangle) \in \mathrm{cl}^*C\,(0)$; the latter set is nonempty according to Proposition 7.7. Then there exists a net $\{\lambda_\nu\}_{\nu\in\mathcal{N}} \subset \mathbb{R}_+^{(T)}$ with $\sum_{t\in T}\lambda_{t\nu} = 1$ for all $\nu \in \mathcal{N}$ such that

$$\sum_{t\in T}\lambda_{t\nu}a_t^* \xrightarrow{w^*} x^* \quad \text{and} \quad \sum_{t\in T}\lambda_{t\nu}b_t \to \left\langle x^*, \bar{x}\right\rangle.$$

Form further the net elements $p_\nu^* \in l^\infty(T)^*$ by

$$p_\nu^* := -\sum_{t\in T}\lambda_{t\nu}\delta_t, \quad \text{with} \quad \left\|p_\nu^*\right\| = \left\langle p_\nu^*, -e\right\rangle = 1, \;\; \nu \in \mathcal{N},$$

and find by the Alaoglu-Bourbaki theorem a convergent subnet $p_\nu^* \xrightarrow{w^*} p^*$ for some $p^* \in l^\infty(T)^*$ with $\|p^*\| \le 1$. Employing the same arguments as in the proof of Corollary 7.6, we conclude that

$$1 = \lim_{\nu\in\mathcal{N}}\sum_{t\in T}\lambda_{t\nu} = \left\|p^*\right\| = \left\langle p^*, -e\right\rangle. \tag{7.19}$$

Furthermore, it follows by passing to the limit that

$$\big(p^*, x^*, \langle x^*, \bar{x}\rangle\big) \in \mathrm{cl}^*\mathrm{co}\left\{\big(-\delta_t, a_t^*, b_t\big) \,\big|\, t \in T\right\},$$

which implies by the coderivative calculation of Theorem 7.5 that

$$p^* \in D^*\mathcal{F}(0, \bar{x})(-x^*). \tag{7.20}$$

Suppose now that $x^* = 0$ in (7.20). Since $p^* \neq 0$ by (7.19), we get from (7.20) that $D^*\mathcal{F}(0, \bar{x})(0) \neq \{0\}$. It tells us by Exercise 3.35(i) and the graph convexity of $\mathcal{F}$ that $\mathcal{F}$ is *not Lipschitz-like* around $(0, \bar{x})$ and therefore it cannot satisfy the strong Slater condition by implication (iii)$\Rightarrow$(i) in Theorem 7.2. This contradicts the assumption imposed in the theorem.

Thus $x^* \neq 0$ in (7.20), and we derive from the latter relationship that

$$\left\|x^*\right\|^{-1} p^* \in D^*\mathcal{F}(0, \bar{x}) \left(- \left\|x^*\right\|^{-1} x^*\right),$$

which gives us in turn the estimate

$$\left\|D^*\mathcal{F}(0, \bar{x})\right\| \geq \left\| \left\|x^*\right\|^{-1} p^*\right\| = \left\|x^*\right\|^{-1}$$

and hence justifies the inequality "$\geq$" in (7.18).

It remains to prove the opposite inequality in (7.18). For the nonempty and w^*-compact set S in (7.17), we have $0 \notin S$ by Theorem 7.2, and the function $x^* \mapsto \|x^*\|^{-1}$ is w^*-upper semicontinuous of on S. Thus the supremum in the right-hand side of (7.18) is attained and belongs to $(0, \infty)$. Then condition (v) in Theorem 7.2 implies that $(0, \widehat{x}) \in \mathrm{int}(\mathrm{gph}\,\mathcal{F})$ for some $\widehat{x} \in X$ and so $0 \in \mathrm{int}(\mathrm{dom}\,\mathcal{F})$. Moreover, we have that $p^* \in D^*\mathcal{F}(0, \bar{x})\,(-x^*)$ if and only if $(p^*, x^*) \in N\left((0, \bar{x})\,; \mathrm{gph}\,\mathcal{F}\right)$, which is equivalent to

$$\langle p^*, p\rangle + \langle x^*, x\rangle \leq \langle x^*, \bar{x}\rangle \quad \text{for all} \quad (p, x) \in \mathrm{gph}\,\mathcal{F}. \tag{7.21}$$

This allows us, by taking into account that $0 \in \mathrm{int}(\mathrm{dom}\,\mathcal{F})$, to arrive at

$$p^* \in D^*\mathcal{F}(0, \bar{x})\,(0) \Longleftrightarrow \langle p^*, p\rangle \leq 0 \;\; \text{for all} \;\; p \in \mathrm{dom}\mathcal{F} \Longleftrightarrow p^* = 0. \tag{7.22}$$

Observe furthermore that, since $\bar{x}$ is not a strong Slater point for $\mathcal{F}$ at $p = 0$, we have $(0, \bar{x}) \notin \mathrm{int}(\mathrm{gph}\,\mathcal{F})$ and thus conclude by the classical separation theorem that there is a pair $(p^*, x^*) \neq (0, 0)$ for which condition (7.21) holds. Employing (7.22), we have $x^* \neq 0$ and $p^* \in D^*\mathcal{F}(0, \bar{x})\,(-x^*)$.

Take now $p^* \in D^*\mathcal{F}(0, \bar{x})\,(-x^*)$ with $\|x^*\| \leq 1$, and suppose that $x^* \neq 0$; the arguments above ensure the existence of such an element. By Corollary 7.6, there is a net $\{\lambda_\nu\}_{\nu\in\mathcal{N}} \subset \mathbb{R}_+^{(T)}$ for which

$$\gamma_\nu := \sum_{t\in T} \lambda_{t\nu} \to \left\|p^*\right\| = -\langle p^*, e\rangle, \;\; x_\nu^* := \sum_{t\in T} \lambda_{t\nu} a_t^* \xrightarrow{w^*} x^*, \;\; \sum_{t\in T} \lambda_{t\nu} b_t \to \langle x^*, \bar{x}\rangle.$$

Taking $M \geq \|a_t^*\|$ for every $t \in T$, we get the estimate

$$\|x_\nu^*\| \leq M\gamma_\nu \quad \text{whenever} \quad \nu \in \mathcal{N}$$

and also the limiting relationships

$$0 < \|x^*\| \leq \liminf_{\nu \in \mathcal{N}} \|x_\nu^*\| \leq M \liminf_{\nu \in \mathcal{N}} \gamma_\nu = M \|p^*\|,$$

which ensure that $p^* \neq 0$. It follows furthermore that

$$\|p^*\|^{-1} \big(x^*, \langle x^*, \bar{x}\rangle\big) \in \mathrm{cl}^* C(0).$$

Remembering finally that $0 < \|x^*\| \leq 1$, we arrive at the estimates

$$\|p^*\| \leq \Big\| \|p^*\|^{-1} x^* \Big\|^{-1} \leq \max\Big\{ \|u^*\|^{-1} \Big| \big(u^*, \langle u^*, \bar{x}\rangle\big) \in \mathrm{cl}^* C(0) \Big\},$$

which justify the inequality "$\leq$" in (7.18) and thus complete the proof. △

7.1.3 Coderivative Characterization of Lipschitzian Stability

In this subsection, we employ the above coderivative analysis married to appropriate techniques in linear SIPs to establish the *coderivative criterion* of Lipschitzian stability (in the sense of the validity of the Lipschitz-like property) for infinite linear systems (7.1) with precise calculation of the *exact Lipschitzian bound* $\mathrm{lip}\,\mathcal{F}(0, \bar{x})$. Surprisingly, the obtained results look exactly like in the finite-dimensional setting of Theorem 3.3 for general closed-graph multifunctions, while in the case here we can express both the coderivative criterion and exact Lipschitzian bound entirely in terms of the given data of (7.1).

First, we present necessary and sufficient condition for the Lipschitz-like property of $\mathcal{F}$ around the reference point $(0, \bar{x})$ in the form of (3.9).

Theorem 7.9 (Coderivative Criterion for the Lipschitz-Like Property of Linear Infinite Systems). *Let* $\bar{x} \in \mathcal{F}(0)$ *for the infinite inequality system* (7.1). *Then* $\mathcal{F}$ *is Lipschitz-like around* $(0, \bar{x})$ *if and only if*

$$D^*\mathcal{F}(0, \bar{x})(0) = \{0\}. \tag{7.23}$$

Proof. The "only if" part follows from the proof in Step 1 of Theorem 3.3 valid in arbitrarily Banach spaces. To justify now the "if" part of the theorem, suppose on the contrary that $D^*\mathcal{F}(0, \bar{x})(0) = \{0\}$, while the mapping $\mathcal{F}$ is *not* Lipschitz-like around $(0, \bar{x})$. Then, by the equivalence between properties (i) and (iv) in Theorem 7.2, we get the inclusion

$$(0, 0) \in \mathrm{cl}^*\mathrm{co}\big\{(a_t^*, b_t) \in X^* \times \mathbb{R} \big|\ t \in T\big\}$$

meaning that there is a net $\{\lambda_\nu\}_{\nu\in\mathcal{N}} \in \mathbb{R}^{(T)}_+$ with $\sum_{t\in T} \lambda_{t\nu} = 1$, $\nu \in \mathcal{N}$, and

$$w^*\text{-}\lim_{\nu\in\mathcal{N}} \sum_{t\in T} \lambda_{t\nu}(a_t^*, b_t) = (0,0). \tag{7.24}$$

Since the net $\{\sum_{t\in T} \lambda_{t\nu}(-\delta_t)\}_{\nu\in\mathcal{N}}$ is obviously bounded in $l^\infty(T)^*$, the Alaoglu-Bourbaki theorem ensures the existence of its subnet (no relabeling) that w^*-converges to some element $p^* \in l^\infty(T)^*$, i.e.,

$$p^* = w^*\text{-}\lim_{\nu\in\mathcal{N}} \sum_{t\in T} \lambda_{t\nu}(-\delta_t). \tag{7.25}$$

It follows from (7.25) by the Dirac function definition that

$$\langle p^*, -e\rangle = \lim_{\nu\in\mathcal{N}} \sum_{t\in T} \lambda_{t\nu} = 1, \quad \text{where } \ e = (e_t)_{t\in T} \ \text{ with } \ e_t = 1 \ \text{ for all } \ t \in T,$$

which yields $p^* \neq 0$. Furthermore, combining (7.24) and (7.25) tells us that

$$(p^*, 0, 0) = w^*\text{-}\lim_{\nu\in\mathcal{N}} \sum_{t\in T} \lambda_{t\nu}(-\delta_t, a_t^*, b_t) \ \text{ with } \ p^* \neq 0,$$

and therefore, by the explicit coderivative description of Theorem 7.5, we get the inclusion $p^* \in D^*\mathcal{F}(0,\bar{x})(0)\setminus\{0\}$, which contradicts the assumed condition (7.23). This verifies the sufficiency part of the coderivative criterion (7.23) for the Lipschitz-like property and thus completes the proof of the theorem. △

Our next goal is to calculate the *exact Lipschitzian bound* $\text{lip}\,\mathcal{F}(0,\bar{x})$. To proceed, observe the following limiting representation of $\text{lip}\,F(\bar{x},\bar{y})$ via the distance function to a set that holds for any mapping $F\colon X \rightrightarrows Y$:

$$\text{lip}\,F(\bar{z},\bar{y}) = \limsup_{(z,y)\to(\bar{z},\bar{y})} \frac{\text{dist}\big(y; F(z)\big)}{\text{dist}\big(z; F^{-1}(y)\big)} \quad \text{where } \ 0/0 := 0. \tag{7.26}$$

To begin with, form the closed affine half-space

$$H(x^*,\alpha) := \big\{x \in X \big|\ \langle x^*, x\rangle \le \alpha\big\} \ \text{ for } \ (x^*,\alpha) \in X^* \times \mathbb{R}$$

and derive the distance function representation known as the *Ascoli formula*.

Proposition 7.10 (Ascoli Formula). *We have*

$$\text{dist}\big(x; H(x^*,\alpha)\big) = \frac{\big[\langle x^*, x\rangle - \alpha\big]_+}{\|x^*\|}, \tag{7.27}$$

where $[\gamma]_+ := \max\{\gamma, 0\}$ *for* $\gamma \in \mathbb{R}$ *and* $0/0 := 0$.

Proof. In the case of $x \in H(x^*,\alpha)$, representation (7.27) is obvious. Consider now that case of $x \notin H(x^*,\alpha)$, and define the associated *optimization problem*

$$\text{minimize } \|u - x\| \text{ subject to } u \in H(x^*, \alpha), \tag{7.28}$$

where an optimal solution exists; see Exercise 7.75. Let $\bar{u} \in H(x^*, \alpha)$ be any solution to (7.28). Applying the generalized Fermat rule and then the subdifferential sum rule b valid due to the continuity of $u \mapsto \|u - x\|$ yields

$$0 \in \partial\| \cdot - x\|(\bar{u}) + N\big(\bar{u}; H(x^*, \alpha)\big) \tag{7.29}$$

with $\bar{u} \neq x$. Since we have in this case that

$$\partial\| \cdot - x\|(\bar{u}) = \big\{u^* \in X^* \big|\ \|u^*\| = 1,\ \langle u^*, \bar{u} - x\rangle = \|\bar{u} - x\|\big\}$$

and that $N(\bar{u}; H(x^*, \alpha)) = \text{cone}\{x^*\}$ if $\langle x^*, \bar{u}\rangle = \alpha$ with $N(\bar{u}; H(x^*, \alpha)) = \{0\}$ otherwise, it tells us by (7.29) that

$$\langle x^*, \bar{u}\rangle = \alpha \ \text{ and } \ \|x^*\| \cdot \|\bar{u} - x\| = \langle x^*, x - \bar{u}\rangle.$$

This implies in turn the equalities

$$\|\bar{u} - x\| = \frac{\langle x^*, x\rangle - \langle x^*, \bar{u}\rangle}{\|x^*\|} = \frac{\langle x^*, x\rangle - \alpha}{\|x^*\|} = \frac{\big[\langle x^*, x\rangle - \alpha\big]_+}{\|x^*\|}$$

and thus justifies the Ascoli formula (7.27). △

The next two propositions, which are certainly of their own interest, establish extensions of the Ascoli formula first to the case of *convex* inequalities and then to *infinite* systems of linear inequalities instead of the single one as in (7.27). These results play a significant role in what follows for computing the exact Lipschitzian bound lip $\mathcal{F}(0, \bar{x})$. In their proofs, we use elements of the classical *duality theory* of convex analysis in Banach spaces; see, e.g., [757].

Given a proper (may not be convex) function $\varphi\colon X \to \overline{\mathbb{R}}$, recall that its (always convex) *Fenchel conjugate* $\varphi^*\colon X^* \to \overline{\mathbb{R}}$ is defined by

$$\varphi^*\big(x^*\big) := \sup\big\{\big\langle x^*, x\big\rangle - \varphi(x) \,\big|\, x \in X\big\}. \tag{7.30}$$

First, we provide an extension of the Ascoli formula from (single) linear to convex inequalities by using the Fenchel conjugate (7.30).

Proposition 7.11 (Extended Ascoli Formula for Single Convex Inequalities). *Let $g\colon X \to \overline{\mathbb{R}}$ be a (proper) convex function, and let*

$$Q := \big\{y \in X \big|\ g(y) \le 0\big\}. \tag{7.31}$$

Assume the fulfillment of the classical Slater condition: there is $\widehat{x} \in X$ with $g(\widehat{x}) < 0$. Then the distance function to the set Q in (7.31) is calculated by

$$\text{dist}(x; Q) = \max_{(x^*, \alpha) \in \text{epi}\, g^*} \frac{\big[\langle x^*, x\rangle - \alpha\big]_+}{\|x^*\|}. \tag{7.32}$$

Proof. Observe that the nonemptiness of Q in (7.31) yields $\alpha \geq 0$ whenever $(0, \alpha) \in \operatorname{epi} g^*$ and that the possibility of $x^* = 0$ is not an obstacle in (7.32) under the convention $0/0 := 0$. The distance function $\operatorname{dist}(x; Q)$ is nothing else but the optimal *value function* in the parametric *convex optimization* problem.

$$\text{minimize } \|y - x\| \text{ subject to } g(y) \leq 0. \tag{7.33}$$

Since the Slater condition holds for (7.33) by our assumption, we have the strong Lagrange duality in (7.33) by, e.g., [757, Theorem 2.9.3], which yields

$$\begin{aligned}
\operatorname{dist}(x; Q) &= \max_{\lambda \geq 0} \inf_{y \in X} \big\{\|y - x\| + \lambda g(y)\big\} \\
&= \max \left\{ \max_{\lambda > 0} \inf_{y \in X} \big\{\|y - x\| + \lambda g(y)\big\},\ \inf_{y \in X} \|y - x\| \right\} \\
&= \max \left\{ \max_{\lambda > 0} \inf_{y \in X} \big\{ \|y - x\| + \lambda g(y)\big\},\ 0 \right\}.
\end{aligned}$$

Applying now the classical Fenchel duality theorem to the inner infimum problem above for a fixed $\lambda > 0$, we get

$$\inf_{y \in X} \big\{\|y - x\| + \lambda g(y)\big\} = \max_{y^* \in X^*} \big\{ - \|\cdot - x\|^*(-y^*) - (\lambda g)^*(y^*)\big\}. \tag{7.34}$$

Furthermore, it is well known in convex analysis that

$$\|\cdot - x\|^* (-y^*) = \begin{cases} \langle -y^*, x\rangle & \text{if } \|y^*\| \leq 1, \\ \infty & \text{otherwise.} \end{cases}$$

Substituting it into formula (7.34) leads us to

$$\begin{aligned}
\inf_{y \in X} \big\{ \|y - x\| + \lambda g(y) \big\} &= \max_{\|y^*\| \leq 1} \big\{ \langle y^*, x\rangle - (\lambda g)^* (y^*) \big\} \\
&= \max_{\|y^*\| \leq 1,\ (\lambda g)^*(y^*) \leq \eta} \big\{ \langle y^*, x\rangle - \eta\big\} \\
&= \max_{\|y^*\| \leq 1,\ \lambda g^*(y^*/\lambda) \leq \eta} \big\{ \langle y^*, x\rangle - \eta\big\} \\
&= \max_{\|y^*\| \leq 1,\ (1/\lambda)(y^*, \eta) \in \operatorname{epi} g^*} \big\{ \langle y^*, x\rangle - \eta\big\}.
\end{aligned}$$

This ensures, by denoting $x^* := (1/\lambda) y^*$ and $\alpha := (1/\lambda)\eta$, that

$$\inf_{y \in X} \big\{ \|y - x\| + \lambda g(y)\big\} = \max_{(x^*, \alpha) \in \operatorname{epi} g^*,\ \|x^*\| \leq 1/\lambda} \lambda \big\{ \langle x^*, x\rangle - \alpha\big\}.$$

Combining the latter with the formulas above, we arrive at

$$\begin{aligned}\operatorname{dist}(x;Q) &= \max\Big\{\max_{(x^*,\alpha)\in\operatorname{epi}g^*,\ \|x^*\|\le 1/\lambda}\lambda\big\{\langle x^*,x\rangle-\alpha\big\},\ 0\Big\}\\ &= \max_{(x^*,\alpha)\in\operatorname{epi}g^*,\ \|x^*\|\le 1/\lambda}\Big\{\lambda\big[\langle x^*,x\rangle-\alpha\big]_+\Big\}.\end{aligned}\tag{7.35}$$

It is easy to observe the following relationships hold for any $\lambda>0$:

$$\begin{aligned}\max_{(0,\alpha)\in\operatorname{epi}g^*}\lambda\big\{\langle 0,x\rangle-\alpha\big\} &= \max_{g^*(0)\le\alpha}\lambda\big(\langle 0,x\rangle-\alpha\big)=\lambda\big(-g^*(0)\big)\\ &\le\lambda\inf_{x\in X}g(x)\le\lambda g(\widehat{x})<0.\end{aligned}$$

Taking this into account, we deduce from (7.35) the equalities

$$\begin{aligned}\operatorname{dist}(x;Q) &= \max_{(x^*,\alpha)\in\operatorname{epi}g^*,\ \|x^*\|\le 1/\lambda}\Big\{\lambda\big[\langle x^*,x\rangle-\alpha\big]_+\Big\}\\ &= \max_{(x^*,\alpha)\in\operatorname{epi}g^*}\ \max_{\|x^*\|\le 1/\lambda}\Big\{\lambda\big[\langle x^*,x\rangle-\alpha\big]_+\Big\}\\ &= \max_{(x^*,\alpha)\in\operatorname{epi}g^*}\frac{\big[\langle x^*,x\rangle-\alpha\big]_+}{\|x^*\|},\end{aligned}$$

which justify (7.32) and thus complete the proof of the proposition. △

The next proposition provides the required extension of the Ascoli formula (7.27) to the case of the *infinite* inequality systems (7.1) in Banach spaces.

Proposition 7.12 (Extended Ascoli Formula for Infinite Linear Systems). *Assume that the infinite linear system* (7.1) *satisfies the strong Slater condition at* $p=(p_t)_{t\in T}$. *Then for any* $x\in X$ *and* $p\in l^\infty(T)$, *we have*

$$\operatorname{dist}\big(x;\mathcal{F}(p)\big)=\max_{(x^*,\alpha)\in\operatorname{cl}^*C(p)}\frac{\big[\langle x^*,x\rangle-\alpha\big]_+}{\|x^*\|}.\tag{7.36}$$

If in addition X *is reflexive, then* (7.36) *can be simplified by*

$$\operatorname{dist}\big(x;\mathcal{F}(p)\big)=\max_{(x^*,\alpha)\in C(p)}\frac{\big[\langle x^*,x\rangle-\alpha\big]_+}{\|x^*\|}.\tag{7.37}$$

Proof. Observe that the infinite linear system (7.1) can be represented as

$$\mathcal{F}(p=\big\{x\in X\big|\ g(x)\le 0\big\},\tag{7.38}$$

where the convex function $g\colon X\to\overline{\mathbb{R}}$ is given in the supremum form

$$g(x):=\sup_{t\in T}\big(f_t(x)-p_t\big)\ \text{ with }\ f_t(x):=\langle a_t^*,x\rangle-b_t.\tag{7.39}$$

The assumed strong Slater condition for $\mathcal{F}(p)$ ensures the validity of the classical Slater condition for g from Proposition 7.11. To employ the result therein in the framework of (7.38), we need to calculate the Fenchel conjugate to the supremum function in (7.39). It can be done by (see Exercise 7.77)

$$\begin{cases} \operatorname{epi} g^* = \operatorname{epi}\Big\{\sup_{t\in T}(f_t - p_t)\Big\}^* = \operatorname{cl}^*\operatorname{co}\Big(\bigcup_{t\in T}\operatorname{epi}\big(f_t - p_t\big)^*\Big) \\ \qquad\quad = \operatorname{cl}^* C(p) + \mathbb{R}_+(0,1) \;\text{ with }\; 0 \in X^*, \end{cases} \tag{7.40}$$

where the weak* closedness of the set $\operatorname{cl}^* C(p) + \mathbb{R}_+(0,1)$ is a consequence of the classical Dieudonné theorem; see, e.g., [757, Theorem 1.1.8]. Thus we get the distance formula (7.36) from Proposition 7.11 in general Banach spaces.

To justify the simplified distance formula (7.37) in the case of reflexive spaces, suppose on the contrary that it doesn't hold. Then there is a scalar $\beta \in \mathbb{R}$ such that we have the strict inequalities

$$\max_{(x^*,\alpha)\in \operatorname{cl}^* C(p)} \frac{\big[\langle x^*, x\rangle - \alpha\big]_+}{\|x^*\|} > \beta > \sup_{(x^*,\alpha)\in C(p)} \frac{\big[\langle x^*, x\rangle - \alpha\big]_+}{\|x^*\|}. \tag{7.41}$$

This yields the existence of $(\bar{x}^*, \bar{\alpha}) \in \operatorname{cl}^* C(p)$ with $\bar{x}^* \neq 0$ satisfying

$$\frac{\big[\langle \bar{x}^*, x\rangle - \bar{\alpha}\big]_+}{\|\bar{x}^*\|} > \beta.$$

Taking into account that X is reflexive and that $C(p)$ is convex and then employing the Mazur weak closure theorem, we can replace the weak* closure of $C(p)$ by its norm closure in X^*. This allows us to find a sequence $(x_k^*, \alpha_k) \in C(p)$ converging by norm to $(\bar{x}^*, \bar{\alpha})$ as $k \to \infty$. Hence we get

$$\lim_{k\to\infty} \frac{\big[\big\langle x_k^*, x\big\rangle - \alpha_k\big]_+}{\big\|x_k^*\big\|} = \frac{\big[\langle \bar{x}^*, x\rangle - \bar{\alpha}\big]_+}{\|\bar{x}^*\|} > \beta,$$

and therefore there exists $k_0 \in \mathbb{N}$ for which

$$\frac{\Big[\Big\langle x_{k_0}^*, x\Big\rangle - \alpha_{k_0}\Big]_+}{\Big\|x_{k_0}^*\Big\|} > \beta.$$

The latter contradicts (7.41) and thus completes the proof. △

The following example shows that the reflexivity of the decision space X is an essential requirement for the validity of the simplified distance formula (7.37), even in the framework of (nonreflexive) Asplund spaces.

Example 7.13 (Failure of the Simplified Distance Formula in Nonreflexive Asplund Spaces). Consider the classical space c_0 of real number sequences converging to zero and endowed with the supremum norm. This space is known to be Asplund while not reflexive. Let us show that the simplified distance formula (7.37) *fails* in $X = c_0$ (the classical space of sequences converging to zero, with the supremum norm) for a rather plain linear system of *countable* inequalities. Of course, we need to demonstrate that the inequality "$\le$" is generally violated in (7.37), since the opposite inequality holds in any Banach space. Form the infinite (countable) linear inequality system

$$\mathcal{F}(0) := \big\{x \in c_0 \big|\ \langle e_1^* + e_t^*, x\rangle \le -1,\ t \in \mathbb{N}\big\}, \tag{7.42}$$

where $e_t^* \in l_1$ has 1 as its t^{th}-component, while all the remaining components are zeros. System (7.42) can be rewritten as

$$x \in \mathcal{F}(0) \Longleftrightarrow x(1) + x(t) \le -1 \ \text{ for all }\ t \in \mathbb{N}.$$

Observe that for $z = 0$, we have $\operatorname{dist}(0; \mathcal{F}(0)) = 1$, and the distance is realized at, e.g., $u = (-1, 0, 0, \ldots)$. Indeed, passing to the limit in $x(1)+x(t) \le -1$ as $t \to \infty$ and taking into account that $x(t) \to 0$ by the structure of the space of c_0, we get $x(1) \le -1$. Furthermore, it can be checked that

$$\begin{aligned}&(e_1^*, -1) \in \operatorname{cl}^* C(0),\ \ \langle e_1^*, x - u\rangle \le 0 \ \text{ for all }\ x \in \mathcal{F}(0),\\ &\operatorname{dist}\big(z; \mathcal{F}(0)\big) = \|z - u\| = \langle e_1^*, z - u\rangle = \frac{\langle e_1^*, z\rangle - (-1)}{\|e_1^*\|}.\end{aligned}$$

On the other hand, for the pair $(x^*, \alpha) \in X^* \times \mathbb{R}$ given by

$$(x^*, \alpha) := \Big(e_1^* + \sum_{t\in\mathbb{N}} \lambda_t e_t^*, -1\Big) \in C(0) \ \text{ with }\ \lambda \in \mathbb{R}_+^{(\mathbb{N})} \ \text{ and }\ \sum_{t\in\mathbb{N}} \lambda_t = 1,$$

we can directly verify that $\|x^*\| = 2$ and hence

$$\frac{\big[\langle x^*, z\rangle - \alpha\big]_+}{\|x^*\|} = \frac{1}{2},$$

which shows that the equality in (7.37) is violated for the countable system (7.42) in the nonreflexive Asplund space $X = c_0$.

Prior to deriving the main result of this subsection on the precise calculation of the exact Lipschitzian bound for the infinite system (7.1) at the reference point, we need the following technical assertion.

Lemma 7.14 (Closed-Graph Property of Characteristic Sets). *The set-valued mapping $l^\infty(T) \ni p \mapsto \operatorname{cl}^* C(p) \subset X^* \times \mathbb{R}$ generated by the characteristic sets (7.5) is closed-graph in the norm×weak* topology of $\ell^\infty(T) \times (X^* \times \mathbb{R})$, i.e., for any nets $\{p_\nu\}_{\nu\in\mathcal{N}} \subset l^\infty(T)$, $\{x_\nu^*\}_{\nu\in\mathcal{N}} \subset X^*$, and $\{\beta_\nu\}_{\nu\in\mathcal{N}} \subset \mathbb{R}$, satisfying the conditions $p_\nu \to p$, $x_\nu^* \xrightarrow{w^*} x^*$, $\beta_\nu \to \beta$, and $\big(x_\nu^*, \beta_\nu\big) \in \operatorname{cl}^* C(p_\nu)$ for every $\nu \in \mathcal{N}$, we have the inclusion $(x^*, \beta) \in \operatorname{cl}^* C(p)$.*

Proof. Arguing by contradiction, suppose that $(x^*, \beta) \notin \text{cl}^* C(p)$. Then the classical strict separation indexconvex separation theorem allows us to find a nonzero pair $(x, \alpha) \in X \times \mathbb{R}$ and real numbers γ and γ' satisfying

$$\langle x^*, x \rangle + \beta\alpha < \gamma' < \gamma \leq \langle a_t^*, x \rangle + (b_t + p_t)\,\alpha \quad \text{for all} \quad t \in T.$$

Hence there exists a net index $\nu_0 \in \mathcal{N}$ such that

$$\langle x_\nu^*, x \rangle + \beta_\nu\alpha < \gamma' \quad \text{and} \quad \|\alpha\,(p - p_\nu)\| \leq \gamma - \gamma' \quad \text{whenever} \quad \nu \succeq \nu_0.$$

This ensures therefore the validity of the estimates

$$\begin{aligned}\langle a_t^*, x \rangle + \alpha\,(b_t + p_{t\nu}) &= \langle a_t^*, x \rangle + \alpha\,(b_t + p_t) + \alpha\,(p_{t\nu} - p_t)\\ &\geq \gamma - \|\alpha\,(p_\nu - p)\| \geq \gamma' \quad \text{for all} \quad t \in T.\end{aligned}$$

The latter implies that $\gamma' \leq \langle z^*, x \rangle + \eta\alpha$ for all $(z^*, \eta) \in \text{cl}^* C(p_\nu)$ whenever $\nu \succeq \nu_0$. Thus we arrive at the contradiction

$$\langle x_\nu^*, x \rangle + \beta_\nu\alpha < \gamma' \leq \langle x_\nu^*, x \rangle + \beta_\nu\alpha, \qquad \nu \succeq \nu_0,$$

which completes the proof of the lemma. △

Now we are ready to provide a precise calculation of the exact Lipschitzian bound of $\mathcal{F}$ around $(0, \bar{x})$ in the general Banach space setting.

Theorem 7.15 (Calculating the Exact Lipschitzian Bound of Infinite Linear Systems). *Let $\bar{x} \in \mathcal{F}(0)$ for the linear infinite inequality system* (7.1)*. Suppose that $\mathcal{F}$ satisfies the strong Slater condition at $p = 0$ and that the coefficient set $\{a_t^* |\ t \in T\}$ is bounded in X^*. The following assertions hold:*

(i) *If $\bar{x}$ is a strong Slater point for $\mathcal{F}$ at $p = 0$, then* $\text{lip}\,\mathcal{F}(0, \bar{x}) = 0$.

(ii) *If $\bar{x}$ is not a strong Slater point for $\mathcal{F}$ at $p = 0$, then the exact of $\mathcal{F}$ around* $(0, \bar{x})$ *is calculated by*

$$\text{lip}\,\mathcal{F}(0, \bar{x}) = \max\left\{\|x^*\|^{-1} \middle|\ \left(x^*, \langle x^*, \bar{x} \rangle\right) \in \text{cl}^* C(0)\right\} > 0 \tag{7.43}$$

via the w^*-closure of the characteristic set (7.5) *at $p = 0$.*

Proof. To verify (i), recall from the proof of Theorem 7.8(i) that the assumptions made imply that $(0, \bar{x}) \in \text{int}(\text{gph}\,\mathcal{F})$, which in turn yields $\text{lip}\,\mathcal{F}(0, \bar{x}) = 0$ by the definition of the exact Lipschitzian bound.

Next we justify the more difficult assertion (ii) of the theorem while assuming that $\bar{x}$ is *not* a strong Slater point for $\mathcal{F}$ at $p = 0$. Observe that by Proposition 7.7, the set under the maximum operation on the right-hand side in (7.43) is nonempty and w^*-compact in X^*. Thus the maximum over this set is realized and is finite. The inequality "$\geq$" in (7.43) follows from the estimate

$$\text{lip}\,\mathcal{F}(0, \bar{x}) \geq \|D^*\mathcal{F}(0, \bar{x})\|$$

taken from Exercise 3.35(i) and then combined with formula (7.18) for calculating the coderivative norm of the infinite inequality system (7.1) derived in Theorem 7.8. Thus it remains to verify the opposite inequality "$\le$" in (7.43).

To proceed, let $M := \sup_{t\in T} \|a_t^*\| < \infty$, and observe that the inequality "$\le$" in (7.43) is obvious when $L := \mathrm{lip}\,\mathcal{F}(0,\bar{x}) = 0$. Suppose now that $L > 0$, and consider any pair (p,x) sufficiently close to $(0,\bar{x})$ in representation (7.26) of the exact Lipschitzian bound $\mathrm{lip}\,\mathcal{F}(0,\bar{x})$. By $L > 0$, we can confine ourselves to the case of $(p,x) \notin \mathrm{gph}\,\mathcal{F}$. It follows from the structure of $\mathcal{F}$ that

$$0 < \mathrm{dist}\big(p;\mathcal{F}^{-1}(x)\big) = \sup_{t\in T}\big[\langle a_t^*,x\rangle - b_t - p_t\big]_+. \tag{7.44}$$

Moreover, we have the relationships

$$\begin{aligned}\big\langle a_t^*,x\big\rangle - b_t - p_t &= \big\langle a_t^*,x-\bar{x}\big\rangle + \big\langle a_t^*,\bar{x}\big\rangle - b_t - p_t\\ &\le M\,\|x-\bar{x}\| + \|p\| \quad \text{for all } \ t\in T,\end{aligned}$$

which allow us to conclude that

$$\begin{aligned}0 &< \sup_{(x^*,\beta)\in \mathrm{cl}^*C(p)}\big[\big\langle x^*,x\big\rangle - \beta\big]_+ = \sup_{(x^*,\beta)\in \mathrm{cl}^*C(p)}\big\{\big\langle x^*,x\big\rangle - \beta\big\}\\ &\le M\,\|x-\bar{x}\| + \|p\| \quad \text{for all } \ x\in X \ \text{ and } \ p\in P.\end{aligned} \tag{7.45}$$

Consider further the set

$$C_+(p,x) := \big\{\big(x^*,\beta\big)\in \mathrm{cl}^*C(p)\,\big|\,\big\langle x^*,x\big\rangle - \beta > 0\big\},$$

which is obviously nonempty, and denote

$$M_{(p,x)} := \sup\big\{\,\|x^*\|^{-1}\,\big|\,\big(x^*,\beta\big)\in C_+(p,x)\,\big\}.$$

In our setting, we get $0 \in \mathrm{int}(\mathrm{dom}\,\mathcal{F})$ (see Exercise 7.72(i)) and therefore $p \in \mathrm{dom}\,\mathcal{F}$ for all $p \in l^\infty(T)$ sufficiently close to the origin. In this case, the set $C_+(p,x)$ cannot contain any element of the form $(0,\beta)$, since the contrary would yield $\beta < 0$ by the definition of $C_+(p,x)$, while Proposition 7.3 tells us that $\beta \ge 0$. Thus we conclude that $0 < \|x^*\| \le M$ whenever $(x^*,\beta) \in C_+(p,x)$ and, in particular, $M_{(p,x)} \in (0,\infty]$. It follows furthermore that

$$\frac{\sup_{(x^*,\beta)\in \mathrm{cl}^*C(p)} \dfrac{\big[\langle x^*,x\rangle - \beta\big]_+}{\|x^*\|}}{\sup_{(x^*,\beta)\in \mathrm{cl}^*C(p)}\big[\langle x^*,x\rangle - \beta\big]_+} = \frac{\sup_{(x^*,\beta)\in \mathrm{cl}^*C(p)} \dfrac{\langle x^*,x\rangle - \beta}{\|x^*\|}}{\sup_{(x^*,\beta)\in \mathrm{cl}^*C(p)}\big\{\langle x^*,x\rangle - \beta\big\}} \le M_{(p,x)},$$

where the latter inequality ensures the estimate

$$L \le \limsup_{(p,x)\to(0,\bar{x}),\ x\notin\mathcal{F}(p)\neq\emptyset} M_{(p,x)} := K.$$

Considering next a sequence $(p_k, x_k) \to (0, \bar{x})$ with $x_k \notin \mathcal{F}(p_k) \neq \emptyset$ and

$$L \leq \lim_{k\to\infty} M_{(p_k,x_k)} = K,$$

we select a sequence $\{\alpha_k\}_{k=1}^{\infty} \subset \mathbb{R}$ such that

$$\lim_{k\to\infty} \alpha_k = K \text{ and } 0 < \alpha_k < M_{(p_k,x_k)} \text{ as } k \in \mathbb{N}.$$

Take now $\left(x_k^*, \beta_k\right) \in C_+(p_k, x_k)$ with $\alpha_k < \left\|x_k^*\right\|^{-1}$ for all $k \in \mathbb{N}$. Since the sequence $\left\{x_k^*\right\}_{k\in\mathbb{N}} \subset X^*$ is bounded, it contains a subnet $\left\{x_\nu^*\right\}_{\nu\in\mathcal{N}}$ that w^*-converges to some $x^* \in X^*$. Denoting by $\{p_\nu\}$, $\{x_\nu\}$, $\{\beta_\nu\}$, and $\{\alpha_\nu\}$ the corresponding subnets of $\{p_k\}$, $\{x_k\}$, $\{\beta_k\}$, and $\{\alpha_k\}$, we get from (7.45) that

$$0 < \left\langle x_\nu^*, x_\nu\right\rangle - \beta_\nu \leq M\left\|x_\nu - \bar{x}\right\| + \left\|p_\nu\right\|.$$

Hence $\left\langle x_\nu^*, x_\nu\right\rangle - \beta_\nu \to 0$, which implies by the constructions above that $\beta_\nu \to \langle x^*, \bar{x}\rangle$. We deduce from Lemma 7.14 that

$$\left(x^*, \left\langle x^*, \bar{x}\right\rangle\right) \in \mathrm{cl}^* C\,(0)\,,$$

and then Theorem 7.2 ensures that $x^* \neq 0$.

To finalize verifying the inequality "$\leq$" in (7.43), observe that

$$\left\|x^*\right\| \leq \liminf_{\nu\in\mathcal{N}} \left\|x_\nu^*\right\| \leq \lim_{\nu\in\mathcal{N}} \frac{1}{\alpha_\nu} = \frac{1}{K}$$

due to $\left\|x_\nu^*\right\| \leq \alpha_\nu^{-1}$ and $\lim_{\nu\in\mathcal{N}} \alpha_\nu = K$, which gives us

$$L \leq K \leq \frac{1}{\|x^*\|} \leq \max\left\{ \left\|z^*\right\|^{-1} \,\middle|\, \left(z^*, \left\langle z^*, \bar{x}\right\rangle\right) \in \mathrm{cl}^* C\,(0) \right\}.$$

Remembering the notation above, we complete the proof of the theorem. △

Summarizing the obtained results on the calculations of the coderivative norm in Theorem 7.8 and the exact Lipschitzian bound in Theorem 7.15 allows us to arrive at the *unconditional* relationship between these quantities for the infinite linear inequality system $\mathcal{F}$ with an arbitrary Banach decision space X that is expressed by the same formula as the one (3.10) derived in Theorem 3.3 for set-valued mappings between finite-dimensional spaces.

Corollary 7.16 (Relationship Between the Exact Lipschitzian Bound and Coderivative Norm). *Let $\bar{x} \in \mathcal{F}(0)$ for the infinite system (7.1) satisfying the strong Slater condition at $p = 0$, and let the coefficient set $\{a_t^* |\ t \in T\}$ be bounded in X^*. Then we have the equality*

$$\mathrm{lip}\,\mathcal{F}(0, \bar{x}) = \|D^*\mathcal{F}(0, \bar{x})\|. \tag{7.46}$$

Proof. If $\bar{x}$ is a strong Slater point for $\mathcal{F}$ at $p = 0$, then we get equality (7.46) by comparing assertions (i) in Theorem 7.8 and Theorem 7.15, which yield

$$\operatorname{lip} \mathcal{F}(0, \bar{x}) = \|D^*\mathcal{F}(0, \bar{x})\| = 0.$$

If $\bar{x}$ is not a strong Slater point for $\mathcal{F}$ at $p = 0$, then (7.46) follows from comparing assertions (ii) in Theorem 7.8 and Theorem 7.15, which give us the same formula for calculating both $\|D^*\mathcal{F}(0, \bar{x})\|$ and $\operatorname{lip} \mathcal{F}(0, \bar{x})$. △

7.2 Optimization Under Infinite Linear Constraints

In this section, we derive necessary optimality conditions for SIPs with general nonsmooth cost functions over feasible solution sets governed by infinite linear constraint systems of type (7.1). The calculation of the coderivative of the feasible solution map given in Section 7.1 plays a crucial role in deriving necessary optimality conditions of both upper and lower subdifferential types presented below. The results obtained are then applied to solving an optimization problem of a practical interest arising in water resource modeling.

7.2.1 Two-Variable SIPs with Infinite Inequality Constraints

We deal here with the following SIP problem:

$$\text{minimize } \ \varphi(p, x) \ \text{ subject to } \ x \in \mathcal{F}(p), \tag{7.47}$$

where $\varphi\colon P \times X \to \overline{\mathbb{R}} := (-\infty, \infty]$ is an extended-real-valued *cost* function (generally nonsmooth and nonconvex) defined on the product of *Banach* spaces and where $\mathcal{F}\colon P \rightrightarrows X$ is a set-valued mapping of *feasible solutions*

$$\mathcal{F}(p) := \big\{x \in X \big|\ \langle a_t^*, x\rangle \le b_t + \langle c_t^*, p\rangle, \quad t \in T\big\} \tag{7.48}$$

with an *arbitrary* (possibly infinite) *index set* T and with some fixed elements $a_t^* \in X^*$, $c_t^* \in P^*$, and $b_t \in \mathbb{R}$ for all $t \in T$. Note that our considerations in Section 7.1, conducted mainly from the viewpoints of Lipschitzian stability of parametric mappings $\mathcal{F}(p)$, concern the case of (7.48) with $P = l^\infty(T)$ and $c_t^* = \delta_t$ (Dirac measure), but the coderivative calculation given therein can be easily adapted to the case of (7.48).

Observe that the optimization in (7.47) is taken with respect to both variables (p, x), which are interconnected through the infinite inequality system (7.48). This means in fact that we have *two groups of decision variables* represented by x and p. One player specifies p, and the other solves (7.47) in x subject to (7.48) with the specified p as a parameter. The first one, having the same objective, varies his/her parameter p to get the best outcome via the so-called *optimistic approach*. We could treat this as a *two-level design*: optimizing the basic parameter p at the upper level, while at the lower level, the cost function is optimized with respect to x for the given p. The reader is referred to, e.g., [442], and the bibliography therein for various tuning and tolerancing problems of such types arising in engineering design. Another area where two-variable SIPs governed by (7.47) and (7.48) with Banach decision spaces X and P naturally appear concerns optimization of water resources. A practical problem of this type is introduced and studied in Subsection 7.2.4.

We can notice some similarity between the two-variable optimization problem in (7.47) and (7.48), treated above as a two-level optimistic design, and the optimistic model of bilevel programming that was considered in Chapter 6 for finitely many constraints and will be studied in Section 7.5.4 for infinitely many ones. The main difference between these classes is that (7.48) is a *constraint system* described by finitely many or infinitely many inequalities, while the corresponding parameter-dependent set $S(\cdot)$ at the upper level of bilevel programming is given by a *variational system* of optimal solutions to a lower-level problem of parametric optimization.

Keeping the same notation as in Section 7.1, we proceed now with deriving two types of necessary optimality conditions for the SIP given in (7.47) and (7.48).

7.2.2 Upper Subdifferential Optimality Conditions for SIPs

Let us begin with upper subdifferential optimality conditions for problem (7.47) and (7.48) that utilize the upper regular subdifferential (6.2) of the cost function (7.47) along with the precise coderivative calculation for the infinite inequality constraint system in (7.48).

Recall the well-known *Farkas-Minkowski property* for (7.48) that amounts to saying that the conic convex hull

$$\text{cone}\big\{(-c_t^*, a_t^*, b_t) \in P^* \times X^* \times \mathbb{R}\big|\ t \in T\} \tag{7.49}$$

is weak* closed in the dual space $P^* \times X^* \times \mathbb{R}$.

Now we are ready to formulate and prove upper subdifferential necessary optimality conditions for the SIP in (7.47) and (7.48) in general Banach spaces.

Theorem 7.17 (Upper Subdifferential Conditions for SIPs with Linear Inequality Constraints). *Let* $(\bar{p}, \bar{x}) \in \text{gph}\,\mathcal{F} \cap \text{dom}\,\varphi$ *be a local minimizer for the two-variable SIP given by* (7.47) *and* (7.48). *Then every upper regular subgradient* $(p^*, x^*) \in \widehat{\partial}^+\varphi(\bar{p}, \bar{x})$ *satisfies the asymptotic optimality condition*

$$-\big(p^*, x^*, \langle p^*, \bar{p}\rangle + \langle x^*, \bar{x}\rangle\big) \in \text{cl}^*\text{cone}\,\big\{(-c_t^*, a_t^*, b_t)\big|\ t \in T\big\}. \tag{7.50}$$

If furthermore the Farkas-Minkowski property (7.49) *holds for* (7.48), *then* (7.50) *can be equivalently written in the upper subdifferential KKT form: for every* $(p^*, x^*) \in \widehat{\partial}^+\varphi(\bar{p}, \bar{x})$, *there are multipliers* $\lambda = (\lambda_t)_{t\in T} \in \mathbb{R}_+^{(T)}$ *satisfying*

$$(p^*, x^*) + \sum_{t \in T(\bar{p}, \bar{x})} \lambda_t(-c_t^*, a_t^*) = 0, \tag{7.51}$$

where $\mathbb{R}_+^{(T)}$ *is defined in* (7.3) *and where*

$$T(\bar{p}, \bar{x}) := \big\{t \in T \big|\ \langle a_t^*, \bar{x}\rangle - \langle c_t^*, \bar{p}\rangle = b_t\big\}. \tag{7.52}$$

Proof. Pick any $(p^*, x^*) \in \widehat{\partial}^+\varphi(\bar{p}, \bar{x})$, and, employing the first part of Theorem 1.27 which holds in arbitrary Banach spaces (see Exercise 1.64), construct a function $s\colon P \times X \to \mathbb{R}$ such that

$$s(\bar{p}, \bar{x}) = \varphi(\bar{p}, \bar{x}), \quad \varphi(p, x) \le s(p, x) \ \text{ for all } \ (p, x) \in P \times X, \tag{7.53}$$

and $s(\cdot)$ is Fréchet differentiable at $(\bar{p}, \bar{x})$ with $\nabla s(\bar{p}, \bar{x}) = (p^*, x^*)$. Taking into account that $(\bar{p}, \bar{x})$ is a local minimizer in (7.47), (7.48) and that

$$s(\bar{p}, \bar{x}) = \varphi(\bar{p}, \bar{x}) \le \varphi(p, x) \le s(p, x) \ \text{ for all } \ (p, x) \in \operatorname{gph} \mathcal{F} \ \text{ near } \ (\bar{p}, \bar{x})$$

by (7.53), we deduce that $(\bar{p}, \bar{x})$ is a local minimizer for the auxiliary problem

$$\text{minimize } \ s(p, x) \ \text{ subject to } \ (p, x) \in \operatorname{gph} \mathcal{F} \tag{7.54}$$

with the objective $s(\cdot)$ that is Fréchet differentiable at $(\bar{p}, \bar{x})$. Rewriting (7.54) in the infinite-penalty unconstrained form

$$\text{minimize } \ s(p, x) + \delta\big((p, x); \operatorname{gph} \mathcal{F}\big)$$

via the indicator function of $\operatorname{gph} \mathcal{F}$, observe directly from definition (1.33) of the regular subdifferential at a local minimizer that

$$(0, 0) \in \widehat{\partial}\big[s + \delta(\cdot; \operatorname{gph} \mathcal{F})\big](\bar{p}, \bar{x}). \tag{7.55}$$

Since $s(\cdot)$ is Fréchet differentiable at $(\bar{p}, \bar{x})$, we easily get from (7.55) that

$$(0, 0) \in \nabla s(\bar{p}, \bar{x}) + N\big((\bar{p}, \bar{x}); \operatorname{gph} \mathcal{F}\big),$$

which implies by $\nabla s(\bar{p}, \bar{x}) = (p^*, x^*)$ and the coderivative definition (1.15) that $-p^* \in D^*\mathcal{F}(\bar{p}, \bar{x})(x^*)$. It follows from the proof of Theorem 7.5 that the latter coderivative condition can be constructively described in terms of the initial problem data as follows:

$$\big(-p^*, -x^*, -(\langle p^*, \bar{p}\rangle + \langle x^*, \bar{x}\rangle)\big) \in \operatorname{cl}^*\operatorname{cone}\big\{(-c_t^*, a_t^*, b_t)\big|\ t \in T\big\}. \tag{7.56}$$

Thus (7.56) justifies the asymptotic condition (7.50) for the given upper subgradient $(p^*, x^*) \in \widehat{\partial}^+\varphi(\bar{p}, \bar{x})$. If the Farkas-Minkowski property (7.49) is satisfied, then the operation cl^* in (7.50) can be omitted, and we arrive at the KKT condition (7.51) while completing the proof of the theorem. △

The essence of upper subdifferential conditions in the general framework of minimization has been discussed above in Remark 6.2, which equally applies to the SIP setting of Theorem 7.17. The following consequence of the obtained results is used in Subsection 7.2.4 when both spaces X and P are Banach.

Corollary 7.18 (Necessary Conditions for SIPs with Fréchet Differentiable Costs). *In the setting of Theorem 7.17, suppose that the cost function φ is Fréchet differentiable at the local optimal solution $(\bar{p}, \bar{x})$ with the derivative $(p^*, x^*) = \nabla\varphi(\bar{p}, \bar{x})$. Then (7.50) holds and further reduces to (7.51) if in addition system (7.48) enjoys the Farkas-Minkowski property.*

Proof. It follows directly from Theorem 7.17 since in this case we have $\widehat{\partial}^+\varphi(\bar{p}, \bar{x}) = \{\nabla\varphi(\bar{p}, \bar{x})\}$ for the regular upper subdifferential of φ. △

Observe that in the general settings of Theorem 7.17 and Corollary 7.18, the necessary optimality condition (7.50) is obtained in the *normal form* meaning that we have a nonzero ($\lambda_0 = 1$) multiplier associated with the cost function without any constraint qualification. However, this condition is expressed in the *asymptotic form* involving the weak* closure of the set on the right-hand side of (7.50). This feature partly relates to considering arbitrary index sets in the SIP constraint (7.48) but may also be exhibited in problems with compact index sets as shown in Subsection 7.2.4.

The latter phenomenon doesn't appear under the validity of Farkas-Minkowski property (7.49), which ensures the more conventional KKT form (7.51). Let us present another consequence of Theorem 7.17, where the Farkas-Minkowski property holds and gives us KKT (7.51).

To proceed, we need the following adaptation of the *strong Slater condition* (SSC) from Definition 7.1 to the case of the constraint system (7.48): SSC holds for (7.48) if there is a pair $(\widehat{p}, \widehat{x}) \in P \times X$ such that

$$\sup_{t \in T} \left[\langle a_t^*, \widehat{x} \rangle - \langle c_t^*, \widehat{p} \rangle - b_t \right] < 0. \tag{7.57}$$

The reader can easily check the validity of the equivalent descriptions of SSC for (7.48) similar to those given in Theorem 7.2.

Corollary 7.19 (Upper Subdifferential Conditions in KKT Form). *Suppose that T is a compact Hausdorff space, that both X and P are finite-dimensional, that the mapping $t \mapsto (a_t^*, c_t^*, b_t)$ is continuous on T, and that SSC (7.57) holds. Then for any $(p^*, x^*) \in \widehat{\partial}^+\varphi(\bar{p}, \bar{x})$, there are multipliers $\lambda = (\lambda_t)_{t \in T} \in \mathbb{R}_+^{(T)}$ such that the KKT condition (7.51) is satisfied.*

Proof. To check the fulfillment of the Farkas-Minkowski property under the assumptions imposed in the corollary, we observe first that the boundedness and closedness of the set $\{(c_t^*, a_t^*, b_t) |\ t \in T\}$ (and hence of its convex hull by the classical Carathéodory theorem) follow from the continuity of $t \mapsto (c_t^*, a_t^*, b_t)$ and compactness of T. Using this boundedness and the equivalence (ii)⇔(iii) in the counterpart of Theorem 7.2 for (7.48) gives us the condition

$$(0,0,0) \notin \operatorname{co}\big\{(-c_t^*, a_t^*, b_t)\big|\ t \in T\big\}. \tag{7.58}$$

As well known in convex analysis (see, e.g., [667, Corollary 9.6.1]), the validity of (7.58) in this setting yields the closedness of the convex conic hull of $\{-c_t^*, a_t^*, b_t^*)|\ t \in T\}$, and thus the Farkas-Minkowski property holds. △

7.2.3 Lower Subdifferential Optimality Conditions for SIPs

Now we turn to lower subdifferential optimality conditions for the SIP under consideration, which use the basic subgradients (1.24) of the cost function φ in (7.47). Our standing assumption in this subsection is that both spaces X and P are *Asplund*. Recall also that the lower semicontinuity of φ, which is the standing assumption in this book, is essential here, while it is not needed for the upper subdifferential results of Subsection 7.2.2.

The lower subdifferential conditions for the SIP in (7.47) and (7.48) derived below differ from their upper subdifferential counterparts in assumptions as well as in conclusions even for the case of finite-dimensional decision spaces. Observe that the following theorem utilizes both basic (1.24) and singular (1.25) subgradients of the cost function.

Theorem 7.20 (Lower Subdifferential Conditions for SIPs with Linear Inequality Constraints). *Let* $(\bar p, \bar x) \in \operatorname{gph} \mathcal{F} \cap \operatorname{dom} \varphi$ *be a local minimizer for the SIP under consideration. Suppose also that:*

(a) *either* φ *is locally Lipschitzian around* $(\bar p, \bar x)$*;*

(b) *or* $\operatorname{int}(\operatorname{gph} \mathcal{F}) \neq \emptyset$ (*which is true, in particular, when SSC* (7.57) *holds and the set* $\{(a_t^*, c_t^*)|\ t \in T\}$ *is bounded in* $X^* \times P^*$) *and the system*

$$\begin{array}{c} (p^*, x^*) \in \partial^\infty \varphi(\bar p, \bar x), \\ -\big(p^*, x^*, \langle (p^*, x^*), (\bar p, \bar x)\rangle\big) \in \operatorname{cl}^* \operatorname{cone}\big\{(-c_t^*, a_t^*, b_t)\big|\ t \in T\big\} \end{array} \tag{7.59}$$

admits only the trivial solution $(p^*, x^*) = (0,0)$.

Then there is a basic subgradient pair $(p^*, x^*) \in \partial \varphi(\bar p, \bar x)$ *satisfying the asymptotic optimality condition* (7.50). *If in addition the Farkas-Minkowski property* (7.49) *holds for* (7.48)*, then there are subgradients* $(p^*, x^*) \in \partial\varphi(\bar p, \bar x)$ *and multipliers* $\lambda = (\lambda_t)_{t \in T} \in \mathbb{R}_+^{(T)}$ *satisfying the KKT condition*

$$(p^*, x^*) + \sum_{t \in T(\bar p, \bar x)} \lambda_t(-c_t^*, a_t^*) = 0 \tag{7.60}$$

with the active index set $T(\bar p, \bar x)$ *defined in* (7.52).

Proof. The SIP in (7.47) and (7.48) can be equivalently written as

$$\text{minimize}\ \ \varphi(p, x) + \delta\big((p, x); \operatorname{gph} \mathcal{F}\big). \tag{7.61}$$

Applying the generalized Fermat rule to $(\bar{p}, \bar{x})$ in (7.61) gives us

$$(0, 0) \in \partial\big[\varphi + \delta(\cdot; \operatorname{gph} \mathcal{F})\big](\bar{p}, \bar{x}) \tag{7.62}$$

in terms of the basic subdifferential of the summation function in (7.62). By using an Asplund space version of the subdifferential sum rule from Exercise 2.54(i) with taking into account that the product of two Asplund spaces is Asplund and that the SNC property holds for solid convex sets by Exercise 2.29(ii), we deduce from (7.62) the validity of the inclusion

$$(0, 0) \in \partial\varphi(\bar{p}, \bar{x}) + N\big((\bar{p}, \bar{x}); \operatorname{gph} \mathcal{F}\big) \tag{7.63}$$

provided that either φ is locally Lipschitzian around $(\bar{p}, \bar{x})$ as assumed in (a) or the interior of $\operatorname{gph} \mathcal{F}$ is nonempty and the qualification condition

$$\partial^{\infty}\varphi(\bar{p}, \bar{x}) \cap \big[-N\big((\bar{p}, \bar{x}); \operatorname{gph} \mathcal{F}\big)\big] = \{(0, 0)\} \tag{7.64}$$

is satisfied as assumed in (b). It follows from the proof of Theorem 7.2 that the strong Slater condition (7.57) and the boundedness of $\{(a_t^*, c_t^*)|\ t \in T\}$ surely imply that the interior of $\operatorname{gph} \mathcal{F}$ is nonempty. Using now the coderivative description obtained in Theorem 7.5 while modifying it for the case of $\mathcal{F}$ from (7.48) shows that the qualification condition (7.64) can be equivalently written as the triviality of solutions to system (7.59) imposed above. In the same way, we reduce (7.63) to the validity of (7.50) for some $(p^*, x^*) \in \partial\varphi(\bar{p}, \bar{x})$. If furthermore the Farkas-Minkowski property (7.49) is satisfied for (7.48), then the operation cl^* in (7.50) can be omitted. Thus we arrive at the KKT condition (7.60) and complete the proof of the theorem. $\triangle$

Similarly to Subsection 7.2.2, we can derive from Theorem 7.20 the lower subdifferential counterpart of Corollary 7.19. Observe that the corresponding consequence of Theorem 7.20 involving an appropriate differentiability of the cost function in (7.47) holds under more restrictive assumptions in comparison with Corollary 7.18: besides the Asplund property of X and P, we have to assume the strict differentiability of φ at $(\bar{p}, \bar{x})$.

7.2.4 Applications to Water Resource Optimization

This subsection provides applications of the obtained general results for SIPs to a water resource optimization problem of a practical interest. We formulate the water recourse model and reduce it to a two-variable SIP over a compact index set with Banach decision spaces. The usage of the necessary optimality conditions for such problems established above allows us to determine optimal decision strategies and suggest efficient ways of their realizations.

The *water resource problem* under consideration is inspired by a continuous-time network flow model formulated in [15]. Consider a system of n reservoirs $R_1, \ldots, R_n$ from which a time-varying water demand is required during a fixed continuous-time period $T = [\underline{t}, \overline{t}]$. Let c_i be the capacity of the reservoir R_i, and let water flow into R_i at rate $r_i(t)$ for each $i = 1, \ldots, n$ and $t \in T$. Denote by $D(t)$ the rate of water demand at t, and suppose that all these nonnegative functions $r_1, \ldots, r_n$ and D are piecewise continuous on the compact interval T and are known in advance. If there is enough water to fill all the reservoir capacity, then the rest can be sold to a neighboring dry area provided that the demand is satisfied. Conversely, if the inflows are short and the reservoirs have free capability for holding additional water, then some water can be bought from outside to meet the inner demand in the region; see Fig. 7.1.

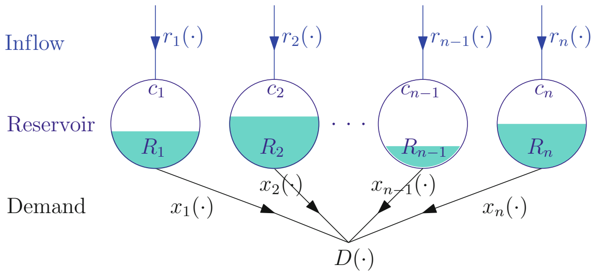


Fig. 7.1 Reservoirs.

Denote by $x_i(t)$ the rate at which water is fed from the reservoir R_i at time $t \in T$. It is natural to assume in our basic model that $x_i \in \mathcal{C}(T)$ for all $i = 1, \ldots, n$. The *feeder constraints* can be expressed by

$$0 \leq x_i(t) \leq \eta_i, \quad i = 1, \ldots, n, \tag{7.65}$$

with fixed bounds $\eta_i \geq 0$. The *selling rate* of water from the reservoir R_i at time t is given by $dp_i(t)$, which means that $p_i(t)$ is the quantity of water sold until instant t and depending on t continuously on the time interval T. Without loss of generality, suppose that $p_i\left(\underline{t}\right) = 0$ for all $i = 1, \ldots, n$. Observe that we are actually buying water at time $t \in T$ if the selling rate $dp_i(t)$ is negative. Denoting further by $s_i \geq 0$ the amount of water initially stored in R_i, we describe the *storage constraints* by

$$\begin{aligned} 0 &\leq \int_{\underline{t}}^{t} \left[r_i(\tau) - x_i(\tau)\right] d\tau - \int_{\underline{t}}^{t} dp_i(\tau) + s_i \\ &= \int_{\underline{t}}^{t} \left[r_i(\tau) - x_i(\tau)\right] d\tau - p_i(t) + s_i \\ &\leq c_i \quad \text{for all} \quad t \in T \quad \text{and} \quad i = 1, \ldots, n \end{aligned} \tag{7.66}$$

and arrive at the following problem of *water resource optimization*:

$$\begin{cases} \text{minimize} \ \ \varphi(p,x) \ \ \text{subject to} \ \ (7.65), (7.66), \\ \text{and} \ \sum_{i=1}^{n} x_i(t) \ge D(t) \ \ \text{for all} \ \ t \in T, \end{cases} \tag{7.67}$$

where the cost function $\varphi(p,x)$ is determined by the cost of water, environmental requirements in the region, and the technology of reservoir processes in the water resource problem. It is clear that we should impose the relationship

$$D(t) \le \sum_{i=1}^{n} \eta_i, \quad t \in T,$$

in order to ensure the consistency of the constraints in (7.67).

Let us show that problem (7.67) can be reduced to the SIP form in (7.47), (7.48) with two groups of variables $(p,x) \in \mathcal{C}(T)^n \times \mathcal{C}(T)^n$. To proceed, define the following t-parametric families of functions on T:

$$\delta_t(\tau) := \begin{cases} 0 & \text{if} \ \ \underline{t} \le \tau < t, \\ 1 & \text{otherwise;} \end{cases} \qquad \alpha_t(\tau) := \begin{cases} \tau & \text{if} \ \ \underline{t} \le \tau < t, \\ t & \text{otherwise.} \end{cases}$$

Both families $\{\delta_t \mid t \in T\}$ and $\{\alpha_t \mid t \in T\}$ are subsets of the dual space $\mathcal{C}(T)^*$. In fact, the Riesz representation theorem ensures that each function $\gamma\colon T \to \mathbb{R}$ of bounded variation on T determines a linear functional on $\mathcal{C}(T)$ by

$$z \mapsto \langle \gamma, z \rangle := \int_{\underline{t}}^{\overline{t}} z(\tau)\, d\gamma(\tau), \quad z \in \mathcal{C}(T),$$

via the Stieltjes integral. It is easy to check that

$$\int_{\underline{t}}^{t} x_i(\tau)\, d\tau = \langle \alpha_t, x_i \rangle, \quad d\alpha_t(\tau) = \chi_{[\underline{t},t]}(\tau)\, d\tau \ \ \text{for} \ \ t \in T,$$

where $\chi_{[\underline{t},t]}$ is the standard characteristic function of the interval $[\underline{t}, t]$. Moreover, for each element $z \in \mathcal{C}(T)$, we have

$$\langle \delta_t, z \rangle = z\,(t)\,, \quad t \in T,$$

and thus δ_t can be identified in this context with the Dirac measure at t, which justifies the δ-notation above. Consider further the functions

$$\beta_i(t) := \int_{\underline{t}}^{t} r_i(\tau)\, d\tau \ \ \text{for} \ \ i = 1, \ldots, n, \quad t \in T,$$

and notice that the constraints in (7.66) can be rewritten as

$$\begin{cases} \langle \delta_t, p_i \rangle + \langle \alpha_t, x_i \rangle \le \beta_i(t) + s_i, \\ -\langle \delta_t, p_i \rangle - \langle \alpha_t, x_i \rangle \le c_i - s_i - \beta_i(t), \end{cases} \tag{7.68}$$

while the one in (7.67) admits the form

$$\sum_{i=1}^{n} \langle \delta_t, x_i \rangle \ge D(t), \quad t \in T. \tag{7.69}$$

Observing finally that the constraints in (7.65) can be equivalently given by

$$0 \le \langle \delta_t, x_i \rangle \le \eta_i, \quad i = 1, \ldots, n, \quad t \in T, \tag{7.70}$$

we arrive at the following reduction result.

Proposition 7.21 (Water Resource Problem as SIP in Banach Spaces). *The problem of water resource optimization* (7.67) *is equivalent to the two-variable SIP of type* (7.47) and (7.48) *in the space* $\mathcal{C}(T) \times \mathcal{C}(T)$*:*

$$\textit{minimize } \varphi(p, x) \textit{ subject to } (7.68), (7.69), \textit{ and } (7.70) \tag{7.71}$$

with the data δ_t, α_t, β_t, c_i, s_i, η_i, *and* D *defined above.*

Now we examine the possibility to apply the obtained necessary optimality conditions for SIPs to the case of the water resource model (7.71). Since the space $\mathcal{C}(T)$ for both variables x and p in our model is not Asplund, we proceed with applying the upper subdifferential optimality conditions of Theorem 7.17 and consider for definiteness the case where the cost function φ is Fréchet differentiable at the reference point, i.e., apply the optimality conditions of Corollary 7.18. For simplicity of notation, suppose in what follows that $n = 1$ in (7.71), and write $(p, x, \beta, c, s, \eta)$ instead of $(p_1, x_1, \beta_1, c_1, s_1, \eta_1)$.

Using the initial data of problem (7.71), define the following convex conic hull in the dual space $\mathcal{C}(T)^* \times \mathcal{C}(T)^* \times \mathbb{R}$ by

$$K(T) := \text{cone}\left\{ \begin{matrix} \Big[\big(\delta_t, \alpha_t, \beta(t) + s\big), \big(-\delta_t, -\alpha_t, c - s - \beta(t)\big), \\ \big(0, -\delta_t, -D(t)\big), \big(0, \delta_t, \eta\big) \text{ over all } t \in T\Big] \end{matrix} \right\}, \tag{7.72}$$

which is a specification of (7.49) for the water recourse problem (7.71). Given a solution pair $(\bar{p}, \bar{x})$, consider the sets of *active indices* corresponding to all the inequality constraints in (7.71) formed as

$$\begin{cases} T_1(\bar{p}, \bar{x}) := \big\{t \in T \,\big|\, \langle \delta_t, \bar{p} \rangle + \langle \alpha_t, \bar{x} \rangle = \beta(t) + s\big\}, \\ T_2(\bar{p}, \bar{x}) := \big\{t \in T \,\big|\, -\langle \delta_t, \bar{p} \rangle - \langle \alpha_t, x \rangle = c - s - \beta(t)\big\}, \\ T_3(\bar{p}, \bar{x}) := \big\{t \in T \,\big|\, -\langle \delta_t, x \rangle = -D(t)\big\}, \\ T_4(\bar{p}, \bar{x}) := \big\{t \in T \,\big|\, \langle \delta_t, \bar{x} \rangle = \eta\big\}. \end{cases} \tag{7.73}$$

The next result provides necessary conditions for local minimizers in the water recourse optimization problem (7.71).

Proposition 7.22 (Necessary Optimality Conditions for Water Resource Optimization). *Let $(\bar{p}, \bar{x})$ be a local minimizer in problem (7.71). Assume that the cost function $\varphi\colon \mathcal{C}(T) \times \mathcal{C}(T) \to \overline{\mathbb{R}}$ is Fréchet differentiable at $(\bar{p}, \bar{x})$, and consider the cone $K(T)$ defined in (7.72). Then we have the inclusion*

$$-\big(\nabla_p \varphi(\bar{p}, \bar{x}), \nabla_x \varphi(\bar{p}, \bar{x}), \langle \nabla_p \varphi(\bar{p}, \bar{x}), \bar{p}\rangle + \langle \nabla_x \varphi(\bar{p}, \bar{x}), \bar{x}\rangle\big) \in \mathrm{cl}^* K(T).$$

If furthermore the cone $K(T)$ is weak closed, then there exist generalized multipliers $\lambda = (\lambda_t)_{t\in T}$, $\mu = (\mu_t)_{t\in T}$, $\gamma = (\gamma_t)_{t\in T}$, and $\rho = (\rho_t)_{t\in T} \in \mathbb{R}_+^{(T)}$ satisfying the following KKT relationship:*

$$\begin{cases} -\big(\nabla_p \varphi(\bar{p}, \bar{x}), \nabla_x \varphi(\bar{p}, \bar{x})\big) = \displaystyle\sum_{t\in T_1(\bar{p},\bar{x})} \lambda_t\,(\delta_t, \alpha_t) \\ + \displaystyle\sum_{t\in T_2(\bar{p},\bar{x})} \mu_t\,(-\delta_t, -\alpha_t) + \sum_{t\in T_3(\bar{p},\bar{x})} \gamma_t\,(0, -\delta_t) + \sum_{t\in T_4(\bar{p},\bar{x})} \rho_t\,(0, \delta_t)\,, \end{cases} \tag{7.74}$$

where the sets of active indices $T_i(\bar{p}, \bar{x})$, $i = 1, \ldots, 4$, are defined in (7.73).

Proof. This follows from the necessary optimality conditions in Corollary 7.18 applied to problem (7.71) taking into account the specification of the characteristic cone (7.49) for problem (7.71) obtained in (7.72) and then expressed via the active index sets from (7.73) corresponding to the infinite inequality constraints in (7.68)–(7.70). △

Observe that the optimality conditions obtained in Proposition 7.22 provide a valuable insight to our understanding of *optimal strategies* for the water resource problem. Indeed, it follows from the structures of constraints in (7.71) and their active index sets that the time inclusion $t \in T_1(\bar{p}, \bar{x})$ means that at this moment t the reservoir is empty, while the one of $t \in T_2(\bar{p}, \bar{x})$ means that at this time the quantity of water inside the reservoir given by $\langle \delta_t, p\rangle + \langle \alpha_t, x\rangle - s - \beta(t)$ attains its maximum level c, i.e., the reservoir is full. Similarly the inclusions $t \in T_i(\bar{p}, \bar{x})$ for $i = 3, 4$ signify, respectively, that the water is flowing at its minimum rate or at its maximum rate to satisfy the demand. The KKT relationship (7.74), valid under the Farkas-Minkowski condition, reflects therefore that the "dual action" (p^*, x^*) is a linear combination of these "bang-bang" strategies with the corresponding weights $(\lambda, \mu, \gamma, \rho)$. The general *asymptotic* optimality condition of the proposition indicates from this viewpoint that, in the absence of the Farkas-Minkowski property, the optimal impulse can be approximated by such combinations.

Finally in this section, we fully characterize the setting of Proposition 7.22 in which the Farkas-Minkowski property is satisfied for problem (7.71).

Proposition 7.23 (Farkas-Minkowski Property in Water Resource Optimization). *Let $\widetilde{T}$ be a nonempty subset of the time interval $T = [\underline{t}, \bar{t}]$ in (7.71). Then the cone $K(\widetilde{T})$ from (7.72) is weak* closed in $\mathcal{C}(T)^* \times \mathcal{C}(T)^* \times \mathbb{R}$ if and only if the set $\widetilde{T}$ consists of a finite number of indices.*

Proof. The "if" part easily follows from the definitions. Let us justify the "only if" part arguing by contradiction and taking into account that the space $\mathcal{C}(T)$ is separable. Suppose that the set $\widetilde{T}$ is infinite and pick for simplicity a strictly monotone (increasing or decreasing) sequence $\{t_k\}_{k\in\mathbb{N}}$ in $\widetilde{T}$, which therefore converges to some point of T. It is not hard to check that the sequence in $\mathcal{C}(T)^* \times \mathcal{C}(T)^* \times \mathbb{R}$ given by

$$\left\{\sum_{j=1}^{k} \frac{1}{j^2}\big(\delta_{t_j}, \alpha_{t_j}, \beta(t_j) + s\big)\right\}_{k\in\mathbb{N}} \tag{7.75}$$

weak* converges to the triple (δ, α, b) defined by

$$\langle(\delta, \alpha, b), (p, x, q)\rangle := \langle\delta, p\rangle + \langle\alpha, x\rangle + bq \tag{7.76}$$

via the componentwise relationships

$$\langle\delta, p\rangle := \sum_{j=1}^{\infty} \frac{1}{j^2} p\left(t_j\right), \ \ \langle\alpha, x\rangle := \sum_{j=1}^{\infty} \frac{1}{j^2} \int_{\underline{t}}^{t_j} x(t)\,dt, \ \ b := \sum_{j=1}^{\infty} \frac{1}{j^2}\big(\beta(t_j) + s\big).$$

Indeed, the weak* convergence of the above sequence follows directly from the boundedness of the set $\left\{\big(\delta_{t_j}, \alpha_{t_j}, \beta(t_j) + s\big)\right\}_{k\in\mathbb{N}}$ in $\mathcal{C}(T)^* \times \mathcal{C}(T)^* \times \mathbb{R}$ and the convergence of the series $\sum_{j=1}^{\infty} \frac{1}{j^2}$.

Let us now show that $(\delta, \alpha, b) \notin K(\widetilde{T})$, and thus the cone $K(\widetilde{T})$ is not weak* closed. To verify it, observe that the inclusion $(\delta, \alpha, b) \in K(\widetilde{T})$ yields

$$\delta = \sum_{t\in\widetilde{T}} \lambda_t \delta_t \ \text{ for some } \ \lambda \in \mathbb{R}_+^{(\widetilde{T})},$$

which gives us a function $\delta \in \mathcal{C}(T)^*$ that is discontinuous only on a finite subset of T. It is easy to check at the same time that this component δ of the triple above is the weak* limit of the functions $\sum_{j=1}^{k} \frac{1}{j^2}\delta_{t_j}$ as $k \to \infty$, and hence it is discontinuous on the infinite set $\{t_k\}_{k\in\mathbb{N}}$. The obtained contradiction completes the proof of the proposition. △

One of the remarkable consequences of Proposition 7.23 is that the Farkas-Minkowski property *doesn't* hold for the water resource problem (7.71) on the compact continuous-time interval $T = [\underline{t}, \bar{t}]$. On the other hand, this result justifies yet another interpretation of the optimality conditions of Proposition 7.22 corresponding to the efficient realization of control strategies for reservoirs. Since in practice the measuring and control processes for the water resource model under consideration are implemented only at discrete instants of time, we can consider a *discretization* $\widetilde{T}$ of the time interval T and then apply the KKT conditions of Proposition 7.22 on $\widetilde{T}$.

7.3 Infinite Linear Systems Under Block Perturbations

In this section, we consider a class of infinite inequality constraint systems under *block perturbations*. Besides being of an undoubted interest in semilinear programming for its own sake, systems of this type eventually cover infinite *convex* inequality systems by using Fenchel duality. For brevity, we consider only the issues related to coderivative analysis of infinite linear block-perturbed and convex systems and its applications to characterizing Lipschitzian stability, i.e., we aim to develop convex counterparts of the results given in Section 7.1. It is not hard to observe that the coderivatives results obtained in this way can be equally applied to deriving both upper and lower subdifferential optimality conditions for SIPs with infinite constraints under consideration similarly to those obtained in Section 7.2 for the linear ones.

Our approach is as follows. We first consider infinite linear systems with block perturbations and extend to this case the results of Section 7.1. Then the results obtained are applied to infinite convex systems by using their *linearization* via Fenchel conjugates. As a by-product of our developments, we remove the boundedness assumption previously imposed on the coefficient of linear and convex systems in the case of reflexive decision spaces.

7.3.1 Description of Infinite Linear Block-Perturbed Systems

Given an arbitrary set $T \neq \emptyset$, consider its *partition*

$$\mathcal{J} := \left\{T_j \mid j \in J\right\} \text{ with } T_j \neq \emptyset \text{ for all } j \in J$$

indexed by a fixed set $J \neq \emptyset$ so that we have

$$T = \bigcup_{j \in J} T_j \text{ with } T_i \cap T_j = \emptyset \text{ if } i \neq j,$$

where the sets T_j, $j \in J$, in the partition are referred to as *blocks*.

Given further a decision Banach space and coefficients $(a_t^*, b_t) \in X^* \times \mathbb{R}$, $t \in T$, consider the *block-perturbed* system

$$\sigma_{\mathcal{J}}(p) := \left\{\left\langle a_t^*, x\right\rangle \leq b_t + p_j,\ t \in T_j,\ j \in J\right\} \tag{7.77}$$

with the perturbation parameter $p = \left(p_j\right)_{j \in J}$ ranging in the Banach space $l^\infty(J)$. The zero function $\bar{p} = 0$ is regarded as the *nominal parameter*, which corresponds to the *nominal system*

$$\sigma(0) := \left\{ \langle a_t^*, x \rangle \le b_t, \ t \in T \right\} \tag{7.78}$$

independently on the partition $\mathcal{J}$. The two extreme partitions

$$\mathcal{J}_{\min} := \{T\} \ \text{ and } \ \mathcal{J}_{\max} := \left\{ \{t\} \,\middle|\, t \in T \right\} \tag{7.79}$$

are called the *minimum partition* and the *maximum partition*, respectively.

Our major attention is focused in what follows on coderivative analysis of the *feasible solution map* $\mathcal{F}_{\mathcal{J}} : l^\infty(J) \rightrightarrows X$ generated by (7.77) as

$$\mathcal{F}_{\mathcal{J}}(p) := \left\{ x \in X \,\middle|\, x \ \text{ is a solution to } \ \sigma_{\mathcal{J}}(p) \right\} \tag{7.80}$$

and its applications to a complete characterization of Lipschitzian stability for (7.80) via the given data of the nominal system (7.78). Then we proceed with further applications to infinite convex inequality systems.

7.3.2 Stability of Block-Perturbed Systems via Coderivatives

First, we present the following coderivative calculation for $\mathcal{F}_{\mathcal{J}}$ at the reference point, where δ_j stands for the Dirac measure at $j \in J$ given by

$$\langle \delta_j, p \rangle := p_j \ \text{ for } \ p = \left(p_j\right)_{j \in J} \in l^\infty(J)\,.$$

Proposition 7.24 (Coderivative Calculation for Block-Perturbed Linear Systems). *Let* $\bar{x} \in \mathcal{F}_{\mathcal{J}}(0)$ *for the mapping* $\mathcal{F}_{\mathcal{J}} : l^\infty(J) \rightrightarrows X$ *from* (7.80)*. Then we have* $p^* \in D^*\mathcal{F}_{\mathcal{J}}(0, \bar{x})(x^*)$ *if and only if*

$$\left(p^*, -x^*, -\langle x^*, \bar{x} \rangle\right) \in \text{cl}^*\text{cone}\left\{ \left(-\delta_j, a_t^*, b_t\right) \,\middle|\, j \in J, \ t \in T_j \right\}.$$

Proof. It can be done by following the lines in the proof of Theorem 7.5 and the preceding propositions of Subsection 7.1.2. △

Similarly to (7.5), define the characteristic set for (7.77) by

$$C_{\mathcal{J}}(p) := \text{co}\left\{ (a_t^*, b_t + p_j) \,\middle|\, t \in T_j, \ j \in J \right\} \subset X^* \times \mathbb{R} \tag{7.81}$$

at $p \in l^\infty(J)$ and consider its specification at $p = 0$, which actually doesn't depend on $\mathcal{J}$ but just on the nominal system (7.78):

$$C(0) = \text{co}\left\{ \left(a_t^*, b_t\right) \,\middle|\, t \in T \right\}.$$

The strong Slater condition (SSC) for the nominal system $\sigma(0)$ and the corresponding strong Slater point $\widehat{x}$ are specifications of Definition 7.1 for $p = 0$.

We have the following equivalent relationships, which extend the equivalencies in Theorem 7.2 to the case of linear block-perturbed systems with taking into account some other results and proofs developed in Section 7.1.

Proposition 7.25 (Characterizations of the Lipschitz-Like Property for Linear Systems Under Block Perturbations). *Given $\bar{x} \in \mathcal{F}_{\mathcal{J}}(0)$ for the feasible solution map* (7.80)*, the following are equivalent:*

(i) $\mathcal{F}_{\mathcal{J}}$ *is Lipschitz-like around* $(0, \bar{x})$.
(ii) $D^*\mathcal{F}_{\mathcal{J}}(0, \bar{x})(0) = \{0\}$.
(iii) *SSC holds for* $\sigma(0)$.
(iv) $0 \in \operatorname{int}(\operatorname{dom} \mathcal{F}_{\mathcal{J}})$.
(v) $\mathcal{F}_{\mathcal{J}}$ *is Lipschitz-like around* $(0, x)$ *for all* $x \in \mathcal{F}_{\mathcal{J}}(0)$.
(vi) $(0, 0) \notin \operatorname{cl}^* C(0)$.

Proof. Implication (i)⇒(ii) is verified, due to $D^*_M\mathcal{F}_{\mathcal{J}}(0, \bar{x}) = D^*_N\mathcal{F}_{\mathcal{J}}(0, \bar{x})$ by the graph convexity of $\mathcal{F}_{\mathcal{J}}(0, \bar{x})$, in Step 1 of Theorem 3.3 the proof of which holds without change in any change in arbitrary Banach space; see Exercise 3.35. The verification of the converse application (ii)⇒(i) follows the lines in the proof of Theorem 7.9 with the usage of Proposition 7.24. Since the conditions involved in (iii) and (vi) don't depend on partitions, the equivalence between them reduces to (iii)⇔(iv) for $p = 0$ in Theorem 7.2. Following the proof of (ii)⇔(iii) in Theorem 7.2 allows us to establish the equivalence between (iii) and (iv) for the maximum partition $\mathcal{J} = \mathcal{J}_{\max}$ in (7.79), which obviously implies that (iii)⇒(iv) for an arbitrary partition $\mathcal{J}$. The converse implication (iv)⇒(iii) holds by considering a constant perturbation $p \equiv \varepsilon$ with $\varepsilon > 0$ being sufficiently small to ensure that $p \in \operatorname{int}(\operatorname{dom} \mathcal{F}_{\mathcal{J}})$ by taking into account that constant perturbations (corresponding to the minimum partition $\mathcal{J} = \mathcal{J}_{\min}$ in (7.79)) are surely a particular case of block perturbations. The equivalent relationships in (i)⇔(iv) and (iv)⇔(v) follow from the classical Robinson-Ursescu theorem and the equivalence between the Lipschitz-like property of a mapping and the metric regularity/covering properties of the inverse; see Theorem 3.2, Corollary 3.6, and the corresponding commentaries in Section 3.5. This completes the proof of the proposition. △

Now we proceed with evaluating the exact Lipschitzian bound of the mapping (7.80) under block perturbations. Prior to establishing the main result in this direction, we present several propositions of their independent interest.

Proposition 7.26 (Relationships Between Exact Lipschitzian Bounds of Block-Perturbed Systems). *Let $\bar{x} \in \mathcal{F}_{\mathcal{J}}(0)$ for the feasible solution map from* (7.80)*. Then we have in the notation of* (7.79) *that*

$$\operatorname{lip} \mathcal{F}_{\min}(0, \bar{x}) \le \operatorname{lip} \mathcal{F}_{\mathcal{J}}(0, \bar{x}) \le \operatorname{lip} \mathcal{F}_{\max}(0, \bar{x}) .$$

Proof. We rely on the Lipschitzian bound representation given in (7.26). Consider the nontrivial case where SSC is satisfied at the nominal system $\sigma(0)$; otherwise all the exact Lipschitzian bounds above are equal to ∞ according to the equivalence (i)⇔(iii) in Proposition 7.25. Note that the mappings $\mathcal{F}_{\min}$, $\mathcal{F}_{\mathcal{J}}$, and $\mathcal{F}_{\max}$ act in the spaces $\mathbb{R}$, $l^\infty(J)$, and $l^\infty(T)$, respectively. For each $\rho \in \overline{\mathbb{R}}$, let p_ρ be the constant function $p_\rho \equiv \rho$ on J, and for each $p \in l^\infty(J)$, denote by p_T the piecewise constant function on T defined as p_j on the block T_j, $j \in J$. Let us further verify the two inequalities:

$$\text{dist}\left(\rho; \mathcal{F}_{\min}^{-1}(x)\right) \geq \text{dist}\left(p_\rho; \mathcal{F}_{\mathcal{J}}^{-1}(x)\right), \ \text{dist}\left(p; \mathcal{F}_{\mathcal{J}}^{-1}(x)\right) \geq \text{dist}\left(p_T; \mathcal{F}_{\max}^{-1}(x)\right)$$

valid for any $x \in X$. Indeed, we obviously have that $\mathcal{F}_{\mathcal{J}}^{-1}(x) = \emptyset$ yields $\mathcal{F}_{\min}^{-1}(x) = \emptyset$ and similarly for the second inequality above.

Consider now the nontrivial case where both of these sets are nonempty. Thus we get for some sequence $\{\rho_r\}_{r \in \mathbb{N}} \subset \mathcal{F}_{\min}^{-1}(x)$ that

$$\text{dist}\left(\rho; \mathcal{F}_{\min}^{-1}(x)\right) = \lim_{r \in \mathbb{N}} |\rho - \rho_r| = \lim_{r \in \mathbb{N}} \left\| p_\rho - p_{\rho_r} \right\| \geq \text{dist}\left(p_\rho; \mathcal{F}_{\mathcal{J}}^{-1}(x)\right)$$

by taking into account that $\rho_r \in \mathcal{F}_{\min}^{-1}(x)$ if and only if $p_{\rho_r} \in \mathcal{F}_{\mathcal{J}}^{-1}(x)$.

Finally, we appeal to representation (7.26) of the exact Lipschitzian bound combined with the directly verifiable equalities

$$\mathcal{F}_{\min}(\rho) = \mathcal{F}_{\mathcal{J}}\left(p_\rho\right) \ \text{ and } \ \mathcal{F}_{\mathcal{J}}(p) = \mathcal{F}_{\max}(p_T),$$

which thus allow us to complete the proof of the proposition. △

The next proposition establishes relationships between the coderivative norms of (7.80) corresponding to different partitions.

Proposition 7.27 (Coderivative Norms for Block-Perturbed Systems). *Consider the feasible solution mappings* (7.80) *corresponding to an arbitrary partition* $\mathcal{J}$ *and to the minimum one* (7.79)*. Then for any* $\bar{x} \in \mathcal{F}_{\mathcal{J}}(0)$*, we have*

$$\left\| D^*\mathcal{F}_{\min}(0, \bar{x}) \right\| \leq \left\| D^*\mathcal{F}_{\mathcal{J}}(0, \bar{x}) \right\|. \tag{7.82}$$

Proof. Observe that $\mathcal{F}_{\mathcal{J}}(0) = \mathcal{F}_{\min}(0)$ since both sets therein reduce to the nominal one; hence $\bar{x} \in \mathcal{F}_{\min}(0)$. According to the coderivative norm definition, pick arbitrarily $x^* \in X^*$ with $\|x^*\| \leq 1$, and consider the nontrivial case where there exists $\mu \in \mathbb{R}\backslash\{0\}$ with $\mu \in D^*\mathcal{F}_{\min}(0, \bar{x})(x^*)$. The coderivative calculation in Proposition 7.24 entails the existence of a net $\{\lambda_\nu\}_{\nu \in \mathcal{N}}$ with $\lambda_\nu = (\lambda_{t\nu})_{t \in T} \in \mathbb{R}_+^{(T)}$ as $\nu \in \mathcal{N}$ satisfying the condition

$$\left(\mu, -x^*, -\left\langle x^*, \bar{x}\right\rangle\right) = w^*\text{-}\lim_{\nu \in \mathcal{N}} \sum_{t \in T} \lambda_{t\nu}\left(-1, a_t^*, b_t\right). \tag{7.83}$$

Looking at the first coordinates in (7.83) and setting $\gamma_\nu := \sum_{t \in T} \lambda_{t\nu}$, we get $-\mu = \lim_{\nu \in \mathcal{N}} \gamma_\nu > 0$, and hence $\gamma_\nu > 0$ for ν sufficiently advanced in the directed set $\mathcal{N}$, say for all ν without loss of generality. This gives us

$$\left(\mu^{-1}x^*, \left\langle \mu^{-1}x^*, \bar{x}\right\rangle\right) = w^*\text{-}\lim_{\nu \in \mathcal{N}} \sum_{t \in T} \gamma_\nu^{-1}\lambda_{t\nu}\left(a_t^*, b_t\right) \in \text{cl}^*C(0). \tag{7.84}$$

For each $\nu \in \mathcal{N}$, consider next $\eta_\nu = \left(\eta_{j\nu}\right)_{j \in J} \in \mathbb{R}_+^{(J)}$ with $\eta_{j\nu} := \sum_{t \in T_j} \gamma_\nu^{-1}\lambda_{t\nu}$, which yields $\sum_{j \in J} \eta_{j\nu} = 1$. Since the net $\{\sum_{j \in J} \eta_{j\nu}\left(-\delta_j\right)\}_{\nu \in \mathcal{N}}$ is contained in the ball $\mathbb{B}_{l^\infty(J)^*}$, the Alaoglu-Bourbaki theorem tells us that a certain subnet (indexed without relabeling by $\nu \in \mathcal{N}$) weak* converges to some $p^* \in l^\infty(J)^*$ with $\|p^*\| \leq 1$. Denoting by $e \in l^\infty(J)$ the function whose coordinates are identically 1, we get the equality

$$\langle p^*, -e \rangle = \lim_{\nu \in \mathcal{N}} \sum_{j \in J} \eta_{j\nu} = 1, \quad \text{and so} \quad \|p^*\| = 1.$$

Appealing now to (7.84) shows for the subnet under consideration that

$$\left(p^*, \mu^{-1}x^*, \left\langle \mu^{-1}x^*, \bar{x} \right\rangle\right) = w^*\text{-}\lim_{\nu \in \mathcal{N}} \sum_{j \in J} \sum_{t \in T_j} \gamma_\nu^{-1} \lambda_{t\nu} \left(-\delta_j, a_t^*, b_t\right).$$

Employing then the coderivative description from Proposition 7.24 yields

$$p^* \in D^*\mathcal{F}_{\mathcal{J}}(0, \bar{x})\left(-\mu^{-1}x^*\right).$$

Since $-\mu > 0$, the positive homogeneity of the coderivative implies that

$$-\mu p^* \in D^*\mathcal{F}_{\mathcal{J}}(0, \bar{x})\left(x^*\right),$$

which ensures in turn by the coderivative norm definition that

$$\left\|D^*\mathcal{F}_{\mathcal{J}}(0, \bar{x})\right\| \geq \left\|-\mu p^*\right\| = -\mu = |\mu|.$$

Since the number $\mu \in D^*\mathcal{F}_{\min}(0, \bar{x})(x^*)$ was chosen arbitrarily, we arrive at (7.82) and thus complete the proof of the proposition. △

To proceed further, we make for notational convenience the convention that $\sup \emptyset := 0$, which allows us to get the equality

$$\sup\left\{\left\|u^*\right\|^{-1} \mid \left(u^*, \langle u^*, \bar{x} \rangle\right) \in \text{cl}^*C(0)\right\} = 0$$

for a strong Slater point $\bar{x}$ of $\sigma(0)$. Indeed, it is easy to check that for such a point $\bar{x}$, there is no element $u^* \in X^*$ satisfying $\left(u^*, \langle u^*, \bar{x} \rangle\right) \in \text{cl}^*C(0)$.

Note that the converse statement doesn't hold in general. To illustrate it, consider the system $\sigma(0) := \{tx \leq 1/t \ \text{ as } \ t = 1, 2, \ldots\}$ in $\mathbb{R}$. On the one hand, observe that $\bar{x} = 0$ is not a strong Slater point of this system. On the other hand, we have $\{u^* \in \mathbb{R} |\ (u^*, \langle u^*, \bar{x} \rangle) \in \text{cl}^*C(0)\} = \emptyset$.

Recall also that the failure of SSC for $\sigma(0)$ tells us by Proposition 7.25 that $(0, 0) \in \text{cl}^*C(0)$, which ensures under the convention $1/0 := \infty$ that for any feasible point $\bar{x}$ of $\sigma(0)$, we have the relationship

$$\sup\left\{\left\|u^*\right\|^{-1} \mid \left(u^*, \langle u^*, \bar{x} \rangle\right) \in \text{cl}^*C(0)\right\} = \infty.$$

These observations are useful in deriving the following lower estimate of the coderivative norm for the minimum partition, which is an important step to obtain the main result of this section.

Proposition 7.28 (Lower Estimate of the Coderivative Norm for the Minimum Partition). *Consider the mapping $\mathcal{F}_{\min}: \mathbb{R} \rightrightarrows X$ defined by the minimum partition $\mathcal{J}_{\min}$ in (7.79), and pick any $\bar{x} \in \mathcal{F}_{\min}(0)$. Then we have*

$$\sup\left\{\left\|u^*\right\|^{-1} \mid \left(u^*, \left\langle u^*, \bar{x}\right\rangle\right) \in \mathrm{cl}^*C(0)\right\} \leq \left\|D^*\mathcal{F}_{\min}(0, \bar{x})\right\|. \tag{7.85}$$

Proof. Let us check first that $\|D^*\mathcal{F}_{\min}(0, \bar{x})\| = \infty$ provided that SSC for $\sigma(0)$. Indeed, in this case, Proposition 7.25 tells us that $(0, 0) \in \mathrm{cl}^*C(0)$, which yields in turn the existence of a net $\{\lambda_\nu\}_{\nu\in\mathcal{N}}$ with $\lambda_\nu = (\lambda_{t\nu})_{t\in T} \in \mathbb{R}_+^{(T)}$ and $\sum_{t\in T}\lambda_{t\nu} = 1$ as $\nu \in \mathcal{N}$ satisfying the condition

$$\left(0, 0\right) = w^*\text{-}\lim_{\nu\in\mathcal{N}} \sum_{t\in T} \lambda_{t\nu}\left(a_t^*, b_t\right).$$

The latter obviously implies that $\left(-1, 0, 0\right) = w^*\text{-}\lim_{\nu\in\mathcal{N}} \sum_{t\in T}\lambda_{t\nu}\left(-1, a_t^*, b_t\right)$, i.e., by Proposition 7.24 we get the inclusion

$$-1 \in D^*\mathcal{F}_{\min}(0, \bar{x})(0).$$

Since $D^*\mathcal{F}_{\min}(0, \bar{x})$ is positively homogeneous, the coderivative norm definition ensures the validity of the claimed condition $\|D^*\mathcal{F}_{\min}(0, \bar{x})\| = \infty$.

Next we consider the nontrivial case where SSC holds for $\sigma(0)$ and the set of elements $u^* \in X^*$ with $(u^*, \langle u^*, \bar{x}\rangle) \in \mathrm{cl}^*C(0)$ is nonempty. Take such an element u^*, and observe that the fulfillment of SSC for $\sigma(0)$ yields $u^* \neq 0$ according to Proposition 7.25. The choice of u^* allows us to find a net $\{\lambda_\nu\}_{\nu\in\mathcal{N}}$ with $\lambda_\nu = (\lambda_{t\nu})_{t\in T} \in \mathbb{R}_+^{(T)}$ and $\sum_{t\in T}\lambda_{t\nu} = 1$ as $\nu \in \mathcal{N}$ satisfying

$$\left(u^*, \left\langle u^*, \bar{x}\right\rangle\right) = w^*\text{-}\lim_{\nu\in\mathcal{N}} \sum_{t\in T} \lambda_{t\nu}\left(a_t^*, b_t\right),$$

which can be equivalently rewritten in the form

$$\left(-1, u^*, \left\langle u^*, \bar{x}\right\rangle\right) = w^*\text{-}\lim_{\nu\in\mathcal{N}} \sum_{t\in T} \lambda_{t\nu}\left(-1, a_t^*, b_t\right).$$

This implies that $-1 \in D^*\mathcal{F}_{\min}(0, \bar{x})(-u^*)$, and hence

$$-\left\|u^*\right\|^{-1} \in D^*\mathcal{F}_{\min}(0, \bar{x})\left(-\left\|u^*\right\|^{-1}u^*\right),$$

which ensures by the definition of the coderivative norm that

$$\left\|D^*\mathcal{F}_{\min}(0, \bar{x})\right\| \geq \left\|u^*\right\|^{-1}.$$

Since the element u^* was chosen arbitrarily from those satisfying the inclusion $(u^*, \langle u^*, \bar{x}\rangle) \in \mathrm{cl}^*C(0)$, we arrive at the claimed lower estimate (7.85) of the coderivative norm and thus complete the proof of the proposition. $\triangle$

Now we are ready to establish the main result of this subsection.

Theorem 7.29 (Evaluation of Coderivative Norms for Block-Perturbed Systems). *For any $\bar{x} \in \mathcal{F}_{\mathcal{J}}(0)$, we have the relationships*

$$\sup\left\{\|u^*\|^{-1} \,\Big|\, \big(u^*, \langle u^*, \bar{x}\rangle\big) \in \operatorname{cl}^* C(0)\right\} \le \|D^*\mathcal{F}_{\min}(0,\bar{x})\| \le \big\|D^*\mathcal{F}_{\mathcal{J}}(0,\bar{x})\big\|$$
$$\le \operatorname{lip}\mathcal{F}_{\mathcal{J}}(0,\bar{x}) \le \operatorname{lip}\mathcal{F}_{\max}(0,\bar{x}).$$

Furthermore, if either the coefficient set $\{a_t^* \mid t \in T\}$ is bounded in X^* or the space X is reflexive, then all the above inequalities hold as equalities.

Proof. Recall as above that the lower estimate

$$\|D^*\mathcal{F}_{\mathcal{J}}(0,\bar{x})\| \le \operatorname{lip}\mathcal{F}_{\mathcal{J}}(0,\bar{x}) \tag{7.86}$$

follows from the proof of Step 1 Theorem 3.3 in arbitrary Banach spaces. Applying now (in this order) Propositions 7.28 and 7.27, formula (7.86), and Proposition 7.26 verifies the chain of inequalities claimed in the theorem.

To verify the equalities therein under the additional assumptions made, consider first the case where the coefficient set $\{a_t^* \mid t \in T\}$ is bounded in X^*. Then using Theorem 7.15 adapted to the current notation gives us

$$\operatorname{lip}\mathcal{F}_{\max}(0,\bar{x}) \le \sup\left\{\|u^*\|^{-1} \mid \big(u^*, \langle u^*, \bar{x}\rangle\big) \in \operatorname{cl}^* C(0)\right\} \tag{7.87}$$

in the nontrivial case where SSC holds for the nominal system $\sigma(0)$.

It remains to consider the case where the space X is reflexive and to justify the upper estimate (7.87) provided the validity of SSC for $\sigma(0)$. Employing in this case the Mazur weak closure theorem allows us to replace the weak* closure $\operatorname{cl}^* C(0)$ of the convex set $C(0)$ by its norm closure $\operatorname{cl} C(0)$. Suppose that (7.87) fails, and choose $\beta > 0$ such that

$$\operatorname{lip}\mathcal{F}_{\max}(0,\bar{x}) > \beta > \sup\left\{\|u^*\|^{-1} \mid \big(u^*, \langle u^*, \bar{x}\rangle\big) \in \operatorname{cl} C(0)\right\}. \tag{7.88}$$

Using the distance representation (7.26) of the exact Lipschitzian bound and the first inequality in (7.88) gives us sequences $p_r = (p_{tr})_{t\in T} \to 0$ and $x_r \to \bar{x}$ along which we have the relationship

$$\operatorname{dist}\big(x_r; \mathcal{F}_{\max}(p_r)\big) > \beta \operatorname{dist}\big(p_r; \mathcal{F}_{\max}^{-1}(x_r)\big) \text{ for all } r \in \mathbb{N}, \tag{7.89}$$

which readily implies that the quantity

$$\begin{aligned}\operatorname{dist}\big(p_r; \mathcal{F}_{\max}^{-1}(x_r)\big) &= \sup_{t\in T}\big[\langle a_t^*, x_r\rangle - b_t - p_{tr}\big]_+ \\ &= \sup_{(x^*,\alpha)\in C_{\max}(p_r)} \big[\langle x^*, x_r\rangle - \alpha\big]_+\end{aligned} \tag{7.90}$$

is finite. It follows from Proposition 7.25 due to the assumed SSC that $\mathcal{F}_{\max}(p_r) \neq \emptyset$ for $r \in \mathbb{N}$ sufficiently large, say, for all $r \in \mathbb{N}$ without loss of generality. Furthermore, under this condition, we have

$$\lim_{r\to\infty} \operatorname{dist}\big(x_r; \mathcal{F}_{\max}(p_r)\big) = 0; \tag{7.91}$$

see Exercise 7.86 for more discussions. Assume without loss of generality the validity of SSC for the system $\sigma_{\max}(p_r)$ and then deduce from the extended Ascoli formula (7.37) for infinite linear systems in Proposition 7.12, which holds in reflexive spaces, the representation

$$\operatorname{dist}\big(x_r; \mathcal{F}_{\max}(p_r)\big) = \sup_{(x^*,\alpha)\in C_{\max}(p_r)} \frac{\big[\langle x^*, x_r\rangle - \alpha\big]_+}{\|x^*\|}, \quad r \in \mathbb{N}.$$

This allows us to find $(x_r^*, \alpha_r) \in C_{\max}(p_r)$ as $r \in \mathbb{N}$ satisfying

$$0 < \operatorname{dist}\big(x_r, \mathcal{F}_{\max}(p_r)\big) - \frac{\langle x_r^*, x_r\rangle - \alpha_r}{\|x_r^*\|} < \frac{1}{r}. \tag{7.92}$$

Furthermore, by (7.89) and (7.90), we can choose (x_r^*, α_r) in (7.92) so that

$$\beta \operatorname{dist}\big(p_r; \mathcal{F}_{\max}^{-1}(x_r)\big) < \frac{\langle x_r^*, x_r\rangle - \alpha_r}{\|x_r^*\|} \le \frac{\operatorname{dist}\big(p_r; \mathcal{F}_{\max}^{-1}(x_r)\big)}{\|x_r^*\|}. \tag{7.93}$$

Since $\operatorname{dist}(p_r; \mathcal{F}_{\max}^{-1}(x_r)) > 0$ (otherwise both sides of (7.89) would be equal to zero, which is not possible), it follows from (7.93) that $\|x_r^*\| < \beta^{-1}$ for all $r \in \mathbb{N}$. Thus, by the weak* sequential compactness of the unit balls in duals to reflexive spaces, we select a subsequence $\{x_{r_k}^*\}_{k\in\mathbb{N}}$ that weak* converges to some $x^* \in X^*$ with $\|x^*\| \le \beta^{-1}$. Then (7.91) and (7.92) yield

$$\lim_{k\in\mathbb{N}} \frac{\langle x_{r_k}^*, x_{r_k}\rangle - \alpha_{r_k}}{\|x_{r_k}^*\|} = 0, \quad \text{and so} \quad \lim_{k\in\mathbb{N}} \big(\langle x_{r_k}^*, x_{r_k}\rangle - \alpha_{r_k}\big) = 0.$$

The latter implies by the normal convergence of $\{x_{r_k}\}_{k\in\mathbb{N}}$ to $\bar{x}$ that

$$\lim_{k\in\mathbb{N}} \alpha_{r_k} = \lim_{k\in\mathbb{N}} \langle x_{r_k}^*, x_{r_k}\rangle = \langle x^*, \bar{x}\rangle.$$

Then we deduce from $(x_{r_k}^*, \alpha_{r_k}) \in C_{\max}(p_{r_k})$ the existence of multiplies $\lambda_{r_k} = (\lambda_{tr_k})_{t\in T}$ such that $\lambda_{tr_k} \ge 0$, only finitely many of them are not zero, and

$$\sum_{t\in T} \lambda_{tr_k} = 1, \quad \text{and} \quad (x_{r_k}^*, \alpha_{r_k}) = \sum_{t\in T} \lambda_{tr_k}\big(a_t^*, b_t + p_{tr_k}\big), \quad k \in \mathbb{N}.$$

Combining the above equations gives us the relationships

$$\left(x^*, \langle x^*, \bar{x}\rangle\right) = w^* \text{-}\lim_{k\in\mathbb{N}}(x^*_{r_k}, \alpha_{r_k}) = w^* \text{-}\lim_{k\in\mathbb{N}} \sum_{t\in T} \lambda_{tr_k} \left(a_t^*, b_t + p_{tr_k}\right)$$
$$= w^* \text{-}\lim_{k\in\mathbb{N}} \sum_{t\in T} \lambda_{tr_k} \left(a_t^*, b_t\right) \in \operatorname{cl} C\,(0)\,,$$

where the last equality comes from $\lim_{k\to\infty} \|p_{r_k}\| = 0$. Observe finally that $x^* \neq 0$ due to the validity of SSC for $\sigma\,(0)$ by Proposition 7.25. Hence

$$\sup\left\{\|u^*\|^{-1} \mid \left(u^*, \langle u^*, \bar{x}\rangle\right) \in \operatorname{cl} C\,(0)\right\} \geq \|x^*\|^{-1} \geq \beta,$$

which contradicts (7.88) and thus completes the proof of the theorem. △

7.3.3 Applications to Infinite Convex Inequality Systems

Here we consider *parameterized convex inequality systems* given by

$$\sigma(p) := \left\{\varphi_j\,(x) \leq p_j,\;\; j \in J\right\}, \tag{7.94}$$

where J is an arbitrary index set and where the functions $\varphi_j \colon X \to \overline{\mathbb{R}}$, $j \in J$, are l.s.c. (our standing assumption) and *convex* on the Banach space X. As above, the functional parameter p belongs to $l^\infty(J)$ and the zero function $\bar{p} = 0$ is the nominal parameter. Our goal is to characterize Lipschitzian stability of the convex system (7.94) around $\bar{p} = 0$ by applying the obtained results for block-perturbed linear systems. We can do it with the help of the *Fenchel conjugate* (7.30) defined for each function φ_j by

$$\varphi_j^*(u^*) := \sup\left\{\langle u^*, x\rangle - \varphi_j(x)\big|\; x \in X\right\} = \sup\left\{\langle u^*, x\rangle - \varphi_j(x)\big|\; x \in \operatorname{dom}\varphi_j\right\}.$$

Indeed, the classical Fenchel duality theorem tells us that relationship

$$\varphi_j^{**} = \varphi_j \;\;\text{on}\;\; X \;\;\text{with}\;\; \varphi_j^{**} := \left(\varphi_j^*\right)^*$$

holds under the assumptions made. Using this, we get for each $j \in J$ that the convex inequality $\varphi_j\,(x) \leq p_j$ turns out to be equivalent to the *linear system*

$$\left\{\langle u^*, x\rangle - \varphi_j^*\left(u^*\right) \leq p_j,\;\; u^* \in \operatorname{dom}\varphi_j^*\right\} \tag{7.95}$$

in the sense that they have the same solution sets. Denote

$$T := \left\{\left(j, u^*\right) \in J \times X^* \big|\; u^* \in \operatorname{dom}\varphi_j^*\right\}$$

and observe that T can be partitioned as

$$T = \bigcup_{j \in J} T_j \quad \text{with} \quad T_j := \{j\} \times \operatorname{dom} \varphi_j^*. \tag{7.96}$$

In this way the right-hand side perturbations on the nominal convex system $\sigma(0)$ correspond to block perturbations of the linearized nominal system $\sigma_{\mathcal{J}}(0)$ with the partition $\mathcal{J} := \{T_j \mid j \in J\}$. It is important to realize to this end that the feasible solution map $\mathcal{F} : l^\infty(J) \rightrightarrows X$ to (7.94) given by

$$\mathcal{F}(p) := \{x \in X \mid x \text{ is a solution to } \sigma(p)\} \tag{7.97}$$

and the one for the block-perturbed linearized system $\mathcal{F}_{\mathcal{J}}$ with the partition $\mathcal{J} := \{T_j \mid j \in J\}$ are *exactly the same mapping*. This allows us to implement the results of Subsection 7.3.1 to characterizing Lipschitzian stability of infinite convex systems. It is not hard to check that the convex counterpart of the characteristic set $C_{\mathcal{J}}(p)$ from (7.81) is

$$\begin{aligned} C(p) &:= \operatorname{co}\Big\{\Big(u^*, \varphi_j^*(u^*) + p_j\Big) \,\Big|\, j \in J,\ u^* \in \operatorname{dom}\varphi_j^*\Big\} \\ &= \operatorname{co}\Big(\bigcup\nolimits_{j \in J} \operatorname{gph}(\varphi_j - p_j)^*\Big) \subset X^* \times \mathbb{R}. \end{aligned} \tag{7.98}$$

Observe that for the convex system $\sigma(0)$ under consideration, the corresponding SSC reads as $\sup_{t \in T} \varphi_t(\widehat{x}) < 0$ for some $\widehat{x} \in X$ and that $\widehat{x}$ is a strong Slater point for $\sigma(0)$ if and only if

$$\sup_{(j,u^*) \in T} \big\{ \langle u^*, \widehat{x} \rangle - \varphi_j^*(u^*) \big\} < 0.$$

The next result provides calculating the coderivative of the solution map (7.97) to the original infinite convex system (7.94) in terms of its initial data.

Proposition 7.30 (Calculating Coderivatives for Infinite Convex Systems). *Take* $\bar{x} \in \mathcal{F}(0)$ *for the solution map* (7.97) *to the convex system* (7.94). *Then we have* $p^* \in D^*\mathcal{F}(0, \bar{x})(x^*)$ *if and only if*

$$\big(p^*, -x^*, -\langle x^*, \bar{x} \rangle\big) \in \operatorname{cl}^* \operatorname{cone}\Big(\bigcup_{j \in J} \big(\{-\delta_j\} \times \operatorname{gph} \varphi_j^*\big)\Big). \tag{7.99}$$

Proof. It follows directly from its linear counterpart in Proposition 7.24. △

Now we are ready to present the major result of this subsection proving an evaluation of the exact Lipschitzian bound for the feasible solution map (7.97) for infinite convex inequality systems.

Theorem 7.31 (Evaluation of the Coderivative Norm for Infinite Convex Systems). *For any* $\bar{x} \in \mathcal{F}(0)$ *from* (7.97), *we have the relationships*

$$\sup\left\{ \|u^*\|^{-1} \,\Big|\, \big(u^*, \langle u^*, \bar{x}\rangle\big) \in \mathrm{cl}^*\mathrm{co}\Big(\bigcup_{j\in J} \mathrm{gph}\,\varphi_j^*\Big)\right\} \le \big\|D^*\mathcal{F}(0,\bar{x})\big\| \le \mathrm{lip}\mathcal{F}(0,\bar{x})\,.$$

If furthermore either the set $\bigcup_{j\in J} \mathrm{dom}\,\varphi_j^*$ is bounded in X^* or the space X is reflexive, then the above inequalities hold as equalities.

Proof. It follows from Theorem 7.29 applied to the linear system (7.95) with block perturbations by employing the above linearization procedure and the coderivative calculation given in Proposition 7.30. △

The next example shows that the boundedness assumption, which looks quite natural in the linear setting, may fail for very simple convex systems.

Example 7.32 (Failure of the Boundedness Assumption for Infinite Convex Inequality Systems). Consider the following single convex inequality involving one-dimensional decision and parameter variables:

$$x^2 \le p \quad \text{with} \quad x, p \in \mathbb{R}.$$

The linearized system associated with it reads as follows:

$$\left\{ux \le \frac{u^2}{4} + p,\ u \in \mathbb{R}\right\},$$

and thus the boundedness assumption of Theorem 7.31 fails.

7.4 Metric Regularity of Infinite Convex Systems

In this section, we develop another approach to well-posedness of infinite convex constraint systems concentrating mainly on their metric regularity. The study of well-posedness in Chapter 3 reveals that, although metric regularity of general multifunctions is equivalent to the Lipschitz-like property of their inverses, the former is unnatural (fails as a rule), while the latter holds under unrestrictive qualification conditions for broad classes of set-valued mappings known as parametric *parametric variational systems (PVS)*; see Section 3.3. The situation is *parametric constraint systems (PCS)* different for parametric *constraint* systems, where both metric regularity and Lipschitzian properties can be studied in parallel and are satisfied under similar (symmetric) constraint qualifications; cf. Section 3.3 and [522, Section 4.3]. The infinite constraint systems considered in Sections 7.1 and 7.3 belong to the latter category, and so their metric regularity and Lipschitzian stability can be studied and characterized in a parallel way.

In fact, full characterizations of metric regularity for the infinite linear and convex inequality systems considered in Sections 7.1 and 7.3 can be derived from the *equalities* for their exact Lipschitzian bounds, which are reciprocal to the exact bounds of metric regularity. However, the aforementioned calculation of the exact Lipschitzian bound in Theorem 7.31 (which extends the previous ones for linear systems) is justified under the imposed *boundedness* assumption, which is rather restrictive (as shown in Example 7.32) while cannot be removed in the given proof unless the decision space is reflexive.

The new approach to characterizing metric regularity of infinite convex systems developed below is completely different from the one employed in the previous sections of this chapter. It first concerns the study of metric regularity of general multifunctions with closed and convex graphs for which we establish formulas for the *precise calculation* of the *exact regularity bound* in arbitrary Banach spaces without imposing any qualification conditions while with involving ε-*coderivatives*. Our approach to these issues is based on reducing metric regularity of such mappings to the unconstrained minimization of *DC* (difference of convex) functions. In this way we obtain regularity criteria for general convex-graph multifunctions and then apply them to metric regularity of infinite convex systems. It allows us not only to cover the case of infinite convex inequalities in arbitrary Banach spaces without imposing the aforementioned boundedness assumption but also to include additional linear equality and convex geometric constraints into consideration.

7.4.1 DC Optimization Approach to Metric Regularity

Recall in accordance with (3.2) in Definition 3.1, a set-valued mapping $F\colon X \rightrightarrows Y$ between metric spaces is *metrically regular* around $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ with modulus $\mu > 0$ if there are neighborhoods U of $\bar{x}$ and V of $\bar{y}$ such that

$$\operatorname{dist}\big(x; F^{-1}(y)\big) \le \mu \operatorname{dist}\big(y; F(x)\big) \text{ for any } x \in U \text{ and } y \in Y.$$

The *exact regularity bound* $\operatorname{reg} F(\bar{x}, \bar{y})$ of F around $(\bar{x}, \bar{y})$ is the infimum of all such moduli μ. It is easy to observe directly from the definition that the metric regularity (3.2) is amount to saying that $(\bar{x}, \bar{y})$ is a *local minimizer* of the following unconstrained optimization problem:

$$\text{minimize } \mu \operatorname{dist}\big(y; F(x)\big) - \operatorname{dist}\big(x; F^{-1}(y)\big) \tag{7.100}$$

over $(x, y) \in X \times Y$. Throughout this and the next subsections, we consider, unless otherwise stated, multifunctions F between *arbitrary Banach* spaces with *closed* and *convex* graphs. Observe that (7.100) is a *DC minimization problem*. Problems of this type are briefly studied in Section 6.1 and in much more details in Section 7.5 while from different prospectives.

To proceed, we need to recall some notions and facts from convex analysis and DC optimization. Given a convex function $\varphi\colon X \to \overline{\mathbb{R}}$ and $\varepsilon \ge 0$, the ε-*subdifferential* of φ at $\bar{x} \in \operatorname{dom} \varphi$ is defined by

$$\partial_\varepsilon \varphi(\bar{x}) := \big\{x^* \in X^* \big|\ \langle x^*, x - \bar{x}\rangle \le \varphi(x) - \varphi(\bar{x}) + \varepsilon,\ x \in X\big\}, \tag{7.101}$$

which reduces to the subdifferential of convex analysis for $\varepsilon = 0$; this construction is also known as the *approximate subdifferential* of φ at $\bar{x}$ if $\varepsilon > 0$. We put $\partial_\varepsilon \varphi(\bar{x}) := \emptyset$

if $\bar{x} \notin \operatorname{dom}\varphi$. Note that (7.101) for $\varepsilon > 0$ is different from the ε-enlargement $\widehat{\partial}_\varepsilon\varphi(\bar{x})$ of the regular subdifferential from (1.34) in the case of convex functions under consideration; see Proposition 1.25. The following *ε-subdifferential sum rule* is well known in convex analysis:

$$\partial_\varepsilon(\varphi_1+\varphi_2)(\bar{x}) = \bigcup_{\substack{\varepsilon_1+\varepsilon_2=\varepsilon\\ \varepsilon_1,\varepsilon_2\ge 0}} \Big[\partial_{\varepsilon_1}\varphi_1(\bar{x}) + \partial_{\varepsilon_2}\varphi_2(\bar{x})\Big] \tag{7.102}$$

provided that one of the functions φ_i is continuous at $\bar{x} \in \operatorname{dom}\varphi_1 \cap \operatorname{dom}\varphi_2$; see Exercise 7.93 for more discussions.

Given a convex set $\Omega \subset X$, we have the collection of (*convex*) *ε-normals*

$$N_\varepsilon(\bar{x};\Omega) := \partial_\varepsilon\delta(\bar{x};\Omega) = \big\{x^* \in X^*\big|\ \langle x^*, x-\bar{x}\rangle \le \varepsilon \ \text{ for all } \ x \in \Omega\big\}, \quad \varepsilon \ge 0,$$

which can be equivalently represented in the form

$$N_\varepsilon(\bar{x};\Omega) = \big\{x^* \in X^*\big|\ \sigma_\Omega(x^*) \le \langle x^*, \bar{x}\rangle + \varepsilon\big\}, \tag{7.103}$$

where σ_Ω stands for the *support function* of Ω defined by

$$\sigma_\Omega(x^*) := \sup\big\{\langle x^*, x\rangle\big|\ x \in \Omega\big\}, \quad x^* \in X^*.$$

Again note that convex ε-normals in (7.103) are different as $\varepsilon > 0$ from regular ε-normals in $\widehat{N}_\varepsilon(\bar{x};\Omega)$ defined in (1.6) for general (including convex) sets.

The *ε-coderivative* of a set-valued mapping $F\colon X \rightrightarrows Y$ at $(\bar{x},\bar{y}) \in \operatorname{gph} F$ is defined by the usual scheme via ε-normals to the graph

$$D^*_\varepsilon F(\bar{x},\bar{y})(y^*) := \big\{x^* \in X^*\big|\ (x^*,-y^*) \in N_\varepsilon\big((\bar{x},\bar{y});\operatorname{gph} F\big)\big\} \tag{7.104}$$

for $\varepsilon \ge 0$ with $D^*_0 F(\bar{x},\bar{y}) = D^*F(\bar{x},\bar{y})$. The *ε-coderivative norm* is given by

$$\|D^*_\varepsilon F(\bar{x},\bar{y})\| := \sup\big\{\|x^*\|\ \big|\ x^* \in D^*_\varepsilon F(\bar{x},\bar{y})(y^*),\ y^* \in \mathbb{B}_{Y^*}\big\}. \tag{7.105}$$

If F is metrically regular around $(\bar{x},\bar{y})$, we get from Theorem 3.3(ii), by observing that this part holds in any Banach space, that $D^*F^{-1}(\bar{y},\bar{x})(0) = \{0\}$, and thus arrive at the norm representation via the unit sphere S_{X^*}:

$$\|D^*F^{-1}(\bar{y},\bar{x})\| = \sup\big\{\|y^*\|\ \big|\ y^* \in D^*F^{-1}(\bar{y},\bar{x})(x^*),\ x^* \in S_{X^*}\big\}. \tag{7.106}$$

The following two results from DC programming in Banach spaces involving ε-subgradients of convex functions (7.101) are important in the proof of the main theorem in the next subsection.

Lemma 7.33 (Necessary and Sufficient Conditions for Global DC Minimizers). *Let $\varphi_1,\varphi_2 : X \to \overline{\mathbb{R}}$ be convex functions. Then $\bar{x}$ is a global minimizer of the unconstrained DC program given by*

$$\textit{minimize}\ \ \varphi_1(x) - \varphi_2(x)\ \ \textit{over}\ \ x \in X \tag{7.107}$$

if and only if $\partial_\varepsilon \varphi_2(\bar{x}) \subset \partial_\varepsilon \varphi_1(\bar{x})$ *for all* $\varepsilon \geq 0$.

Note that the *necessity* of the obtained subdifferential inclusion with $\varepsilon = 0$ for *local* minimizers of (7.107) is established in Proposition 6.3 as a consequence of *upper* subdifferential conditions in unconstrained optimization; see more discussions and references in Exercise 7.94(i,ii). The next result provides a *sufficient* condition of this type for *local* minimizers of (7.107); see Exercise 7.94(iii,iv) for the proof and discussions.

Lemma 7.34 (Sufficient Conditions for Local DC Minimizers). *Let* $\varphi_1, \varphi_2 : X \to \overline{\mathbb{R}}$ *be convex functions, and let* φ_2 *be continuous at the point* $\bar{x} \in \mathrm{dom}\,\varphi_1 \cap [\mathrm{int}(\mathrm{dom}\,\varphi_2)]$. *Then* $\bar{x}$ *is a local minimizer of* (7.107) *if there is* $\varepsilon_0 > 0$ *such that* $\partial_\varepsilon \varphi_2(\bar{x}) \subset \partial_\varepsilon \varphi_1(\bar{x})$ *for all* $\varepsilon \in [0, \varepsilon_0]$.

7.4.2 Metric Regularity of Convex-Graph Multifunctions

Now we are ready to establish the main result on calculating the exact regularity bound of closed- and convex-graph multifunctions via their ε-coderivatives at the reference points. The next theorem presents two limiting formulas for calculating this bound in general Banach spaces.

Theorem 7.35 (ε-Coderivative Formulas for the Exact Regularity Bound). *Given a point* $(\bar{x}, \bar{y}) \in \mathrm{gph}\, F$, *assume that* $\bar{y} \in \mathrm{int}(\mathrm{rge}\, F)$. *Then we have*

$$\mathrm{reg}\, F(\bar{x}, \bar{y}) = \lim_{\varepsilon \downarrow 0} \|D^*_\varepsilon F^{-1}(\bar{y}, \bar{x})\|, \tag{7.108}$$

$$\mathrm{reg}\, F(\bar{x}, \bar{y}) = \lim_{\varepsilon \downarrow 0} \Big[\sup \Big\{ \frac{1}{\|x^*\|} \Big| \ x^* \in D^*_\varepsilon F(\bar{x}, \bar{y})(y^*),\ y^* \in S_{Y^*} \Big\} \Big]. \tag{7.109}$$

Proof. Since $\bar{y} \in \mathrm{int}\,(\mathrm{rge}\, F)$, it follows from the Robinson-Ursescu theorem in Banach spaces (see Corollary 3.6 and Exercise 3.49) that F is metrically regular around $(\bar{x}, \bar{y})$, i.e., there are $\eta, \mu > 0$ such that

$$\mathrm{dist}\big(x; F^{-1}(y)\big) \leq \mu\, \mathrm{dist}\big(y; F(x)\big) \ \text{ for all } (x, y) \in B_\eta(\bar{x}, \bar{y}). \tag{7.110}$$

Consider now the convex functions φ_1, φ_2 on $X \times Y$ defined by

$$\varphi_1(x, y) := \mathrm{dist}\big(y; F(x)\big) \ \text{ and } \ \varphi_2(x, y) := \mathrm{dist}\big(x; F^{-1}(y)\big) \tag{7.111}$$

and deduce from the covering property of F equivalent to metric regularity that there is $r > 0$ such that $B_{2r}(\bar{y}) \subset F(\bar{x} + \mathbb{B}_X)$. Combining this with the construction of φ_2 in (7.111) provides the estimate

$$\varphi_2(x, y) \leq \|x - \bar{x}\| + 1 \ \text{ whenever } \ y \in B_{2r}(\bar{y}),$$

which tells us that φ_2 is upper bounded around $(\bar{x}, \bar{y})$, and thus it is locally Lipschitzian around this point due to the well-known result of convex analysis; see, e.g., [757, Corollary 2.2.13]. Implementing our approach to metric regularity, we conclude that $(\bar{x}, \bar{y})$ is a local minimizer of the DC program:

$$\text{minimize } \ \mu\varphi_1(x, y) - \varphi_2(x, y) \ \text{ subject to } \ (x, y) \in X \times Y, \tag{7.112}$$

and consequently it is a global minimizer of the DC function

$$\big(\mu\varphi_1 + \delta(\cdot; B_\eta(\bar{x}, \bar{y}))\big)(x, y) - \varphi_2(x, y) \ \text{ over } \ (x, y) \in X \times Y. \tag{7.113}$$

Applying Lemma 7.33 to the DC program (7.113) gives us the inclusion

$$\partial_\varepsilon\varphi_2(\bar{x}, \bar{y}) \subset \partial_\varepsilon\big(K\varphi_1 + \delta(\cdot; B_\eta(\bar{x}, \bar{y}))\big)(\bar{x}, \bar{y}) \ \text{ for all } \ \varepsilon \ge 0.$$

Since the function $\delta((\cdot, \cdot); B_\eta(\bar{x}, \bar{y}))$ is continuous at $(\bar{x}, \bar{y})$, it follows from the ε-subdifferential sum rule (7.102) that the latter inclusion reduces to

$$\begin{aligned}\partial_\varepsilon\varphi_2(\bar{x}, \bar{y}) &\subset \bigcup_{\substack{\varepsilon_1+\varepsilon_2=\varepsilon\\ \varepsilon_1,\varepsilon_2\ge 0}} \Big[\partial_{\varepsilon_1}(K\varphi_1)(\bar{x}, \bar{y}) + \partial_{\varepsilon_2}\delta\big(\cdot; B_\eta(\bar{x}, \bar{y})\big)(\bar{x}, \bar{y})\Big]\\ &= \bigcup_{\substack{\varepsilon_1+\varepsilon_2=\varepsilon\\ \varepsilon_1,\varepsilon_2\ge 0}} \Big[\partial_{\varepsilon_1}(K\varphi_1)(\bar{x}, \bar{y}) + \frac{\varepsilon_2}{\eta}\mathbb{B}_{X^*\times Y^*}\Big]\end{aligned} \tag{7.114}$$

due to the fact that $\partial_\varepsilon\delta(\cdot; B_r(x))(x) = \frac{\varepsilon}{r}\mathbb{B}_{X^*}$ for all $\varepsilon \ge 0$ and $r > 0$.

Let us next calculate the ε-subdifferentials of the functions $K\varphi_1$ and φ_2 from (7.111) at $(\bar{x}, \bar{y})$ by using their Fenchel conjugates (7.30) and the obvious ε-subdifferential representation for any convex function $\varphi\colon X \to \overline{\mathbb{R}}$:

$$\partial_\varepsilon\varphi(\bar{x}) = \big\{x^* \in X^* \big|\ \varphi^*(x^*) \le \langle x^*, \bar{x}\rangle - \varphi(\bar{x}) + \varepsilon\big\}, \quad \varepsilon \ge 0.$$

In this way we get that $(x^*, y^*) \in \partial_{\varepsilon_1}(\mu\varphi_1)(\bar{x}, \bar{y})$ if and only if

$$(\mu\varphi_1)^*(x^*, y^*) \le \langle x^*, \bar{x}\rangle + \langle y^*, \bar{y}\rangle + \varepsilon_1, \tag{7.115}$$

which ensures in turn by elementary transformations that

$$\begin{aligned}(\mu\varphi_1)^*(x^*, y^*) &= \sup_{x,y}\Big(\langle x^*, x\rangle + \langle y^*, y\rangle - \mu\,\text{dist}\big(y; F(x)\big)\Big)\\ &= \sup_{x,y}\Big(\langle x^*, x\rangle + \langle y^*, y\rangle - \inf_u\big(\mu\|y - u\| + \delta(u; F(x))\big)\Big)\\ &= \sup_{u,x,y}\Big(\langle x^*, x\rangle + \langle y^*, y-u\rangle + \langle y^*, u\rangle - \mu\|y - u\| - \delta\big(u; F(x)\big)\Big)\\ &= \sup_{u,x,y}\Big(\langle x^*, x\rangle + \langle y^*, u\rangle - \delta\big(u; F(x)\big) + \langle y^*, y\rangle - \mu\|y\|\Big)\\ &= \sigma_{\text{gph}F}(x^*, y^*) + \delta\big(y^*; \mu\mathbb{B}_{Y^*}\big).\end{aligned}$$

By using (7.103) and (7.115), the latter implies that

$$\partial_{\varepsilon_1}(\mu\varphi_1)(\bar{x},\bar{y}) = N_{\varepsilon_1}\big((\bar{x},\bar{y});\operatorname{gph} F\big)\cap\big(X^*\times\mu\mathbb{B}_{Y^*}\big). \tag{7.116}$$

Similarly, by taking into account the form of φ_2 in (7.111), we arrive at

$$\partial_{\varepsilon}\varphi_2(\bar{x},\bar{y}) = N_{\varepsilon}\big((\bar{x},\bar{y});\operatorname{gph} F\big)\cap\big(\mathbb{B}_{X^*}\times Y^*\big). \tag{7.117}$$

Thus the inclusion in (7.114) reduces to the following:

$$\begin{aligned} N_{\varepsilon}\big((\bar{x},\bar{y});\operatorname{gph} F\big)\cap\big(\mathbb{B}^*\times Y^*\big) \subset & \bigcup_{\substack{\varepsilon_1+\varepsilon_2=\varepsilon\\ \varepsilon_1,\varepsilon_2\ge 0}} N_{\varepsilon_1}\big((\bar{x},\bar{y});\operatorname{gph} F\big)\cap\big(X^*\times\mu\mathbb{B}^*\big)\\ & +\frac{\varepsilon_2}{\eta}\mathbb{B}_{X^*\times Y^*}. \end{aligned} \tag{7.118}$$

To justify the equality in (7.108), let us fix $\varepsilon > 0$ and pick any $(x^*,y^*)\in\mathbb{B}_{X^*}\times Y^*$ satisfying $y^*\in D^*_{\varepsilon}F^{-1}(\bar{y},\bar{x})(x^*)$, which means that $(-x^*,y^*)\in N_{\varepsilon}((\bar{x},\bar{y});\operatorname{gph} F)$. It follows from (7.118) that there exist a number $\varepsilon_1\in[0,\varepsilon]$ and ε_1-normals $(u^*,v^*)\in N_{\varepsilon_1}((\bar{x},\bar{y});\operatorname{gph} F)$ satisfying the estimates $\|v^*\|\le\mu$ and $\|y^*-v^*\|\le(\varepsilon-\varepsilon_1)\eta^{-1}$. Hence, we get the inequalities

$$\|y^*\|\le\|v^*\|+(\varepsilon-\varepsilon_1)\eta^{-1}\le\mu+\varepsilon\eta^{-1}.$$

Observe from (7.105) that the function $\varepsilon\mapsto\|D^*_{\varepsilon}F^{-1}(\bar{y},\bar{x})\|$ is nondecreasing, which implies therefore the relationships

$$\lim_{\varepsilon\downarrow 0}\|D^*_{\varepsilon}F^{-1}(\bar{y},\bar{x})\| = \inf_{\varepsilon>0}\|D^*_{\varepsilon}F^{-1}(\bar{y},\bar{x})\|\le\inf_{\varepsilon>0}\Big(\mu+\varepsilon\eta^{-1}\Big).$$

Letting $\mu\downarrow\operatorname{reg} F(\bar{x};\bar{y})$ above gives us the estimate

$$\lim_{\varepsilon\downarrow 0}\|D^*_{\varepsilon}F^{-1}(\bar{y},\bar{x})\|\le\operatorname{reg} F(\bar{x},\bar{y}). \tag{7.119}$$

It follows from (7.119) that the equality in (7.108) is trivial if $\operatorname{reg} F(\bar{x},\bar{y}) = 0$. Considering further the case of $\operatorname{reg} F(\bar{x},\bar{y}) > 0$, we deduce from the definition of the exact regularity bound that $(\bar{x},\bar{y})$ is *not* a local minimizer of the DC problem (7.112) when $0<\mu<\operatorname{reg} F(\bar{x},\bar{y})$. Then Lemma 7.34 allows us to find sequences $\varepsilon_k\downarrow 0$ and $(x_k^*,y_k^*)\in\partial_{\varepsilon_k}\varphi_2(\bar{x},\bar{y})$ such that $(x_k^*,y_k^*)\notin\partial_{\varepsilon_k}(\mu\varphi_1)(\bar{x},\bar{y})$ as $k\in\mathbb{N}$. Combining this with (7.116) and (7.117) implies that

$$\|x_k^*\|\le 1\ \text{ and }\ \|y_k^*\|>\mu\ \text{ for all }\ k\in\mathbb{N}. \tag{7.120}$$

Since $B_{2r}(\bar{y})\subset F(\bar{x}+\mathbb{B}_X)$ as mentioned, (7.117) and (7.120) yield

$$\begin{aligned} \varepsilon_k \ge & \sup_{(x,y)\in\operatorname{gph}F}\Big(\langle x_k^*,x-\bar{x}\rangle+\langle y_k^*,y-\bar{y}\rangle\Big)\\ & \ge \sup\nolimits_{y\in B_{2r}(\bar{y})}\big(\langle y_k^*,y-\bar{y}\rangle\big)-\|x_k^*\|\ge 2r\|y_k^*\|-\|x_k^*\|\ge 2r\mu-\|x_k^*\|. \end{aligned} \tag{7.121}$$

By $\varepsilon_k \downarrow 0$ as $k \to \infty$, we have $\|x_k^*\| \geq 2r\mu - \varepsilon_k \geq r\mu$ for sufficiently large k. Suppose without loss of generality that $\|x_k^*\| \geq r\mu$ for all $k \in \mathbb{N}$, and define

$$\widetilde{y}_k^* := y_k^* \|x_k^*\|^{-1}, \;\; \widetilde{x}_k^* := -x_k^* \|x_k^*\|^{-1}, \;\; \text{and} \;\; \widetilde{\varepsilon}_k := \varepsilon_k \|x_k^*\|^{-1}.$$

Then $\|\widetilde{x}_k^*\| = 1$, $\widetilde{\varepsilon}_k \downarrow 0$, and $\widetilde{y}_k^* \in D^*_{\widetilde{\varepsilon}_k} F^{-1}(\bar{y}, \bar{x})(\widetilde{x}_k^*)$. We get from (7.120) that

$$\sup\left\{\|y^*\| \;\middle|\; y^* \in D^*_{\widetilde{\varepsilon}_k} F^{-1}(\bar{y}, \bar{x})(y^*), \;\; x^* \in S_{X^*}\right\} \geq \|\widetilde{y}_k^*\| = \|y_k^*\| \cdot \|x_k^*\|^{-1} > \mu.$$

Letting $k \to \infty$ and $\mu \uparrow \operatorname{reg} F(\bar{x}, \bar{y})$ tells us that

$$\limsup_{\varepsilon \downarrow 0} \left\{\|y^*\| \Big|\;\; y^* \in D^*_{\varepsilon} F^{-1}(\bar{y}, \bar{x})(x^*), \;\; x^* \in S_{X^*}\right\} \geq \operatorname{reg} F(\bar{x}; \bar{y}),$$

which yields the equality in (7.108) by using (7.119).

It remains to prove formula (7.109). By the arguments similar to those following (7.121), we arrive at the relationships

$$D^*_{\varepsilon} F(\bar{x}, \bar{y})(y^*) \cap r\mathbb{B}_{X^*} = \emptyset \;\; \text{for all} \;\; 0 < \varepsilon < r \;\; \text{and} \;\; y^* \in S_{Y^*}. \tag{7.122}$$

Pick any $(x^*, y^*) \in X^* \times S_{Y^*}$ such that $x^* \in D^*_{\varepsilon} F(\bar{x}, \bar{y})(y^*)$ for some $0 < \varepsilon < r$. Define further $\widehat{x}^* := -x^*\|x^*\|^{-1}$, $\widehat{y}^* := -y^*\|x^*\|^{-1}$, and $\widehat{\varepsilon} := \varepsilon\|x^*\|^{-1}$. This ensures that $\widehat{x}^* \in S_{X^*}$, $\|\widehat{y}^*\| = \|x^*\|^{-1}$, and $\widehat{y}^* \in D^*_{\widehat{\varepsilon}} F^{-1}(\bar{y}, \bar{x})(\widehat{x}^*)$. Observe from (7.122) that $\widehat{\varepsilon} \leq \varepsilon r^{-1}$, and thus we have

$$\|x^*\|^{-1} = \|\widehat{y}^*\| \leq \|D^*_{\widehat{\varepsilon}} F^{-1}(\bar{y}, \bar{x})\| \leq \|D^*_{\varepsilon r^{-1}} F^{-1}(\bar{y}, \bar{x})\|.$$

This together with (7.108) yields the inequality "$\geq$" in (7.109) by letting $\varepsilon \downarrow 0$.

To justify the converse inequality in (7.109), note first that it obviously holds when $\operatorname{reg} F(\bar{x}, \bar{y}) = 0$. If $\operatorname{reg} F(\bar{x}, \bar{y}) > 0$, we get from the equality in (7.108) and the norm definition in (7.105) that there exists a sufficiently small number $0 < s < \operatorname{reg} F(\bar{x}, \bar{y})$ ensuring the validity of the condition

$$D^*_{\varepsilon} F^{-1}(\bar{y}, \bar{x})(x^*) \cap s\mathbb{B}_{Y^*} = \emptyset \;\; \text{for all} \;\; 0 < \varepsilon < s \;\; \text{and} \;\; x^* \in S_{X^*}.$$

The arguments similar to those after (7.122) give us the estimate

$$\|D^*_{\varepsilon} F^{-1}(\bar{y}, \bar{x})\| \leq \sup\left\{\frac{1}{\|x^*\|} \middle| x^* \in D^*_{\varepsilon s^{-1}} F(\bar{x}, \bar{y})(y^*), \, y^* \in S_{Y^*}\right\}. \tag{7.123}$$

Indeed, pick $(y^*, x^*) \in Y^* \times S_{X^*}$ with $y^* \in D^*_{\varepsilon} F^{-1}(\bar{y}, \bar{x})(x^*)$ and get $\|y^*\| > s$. Then for $\widetilde{x}^* := x^*\|y^*\|^{-1}$ and $\widetilde{y}^* := y^*\|y^*\|^{-1}$, we have $\|\widetilde{x}^*\|^{-1} = \|y^*\|$, $\widetilde{y}^* \in S_{Y^*}$, and $\widetilde{x}^* \in D^*_{\frac{\varepsilon}{\|y^*\|}} F(\bar{x}, \bar{y})(\widetilde{y}^*) \subset D^*_{\varepsilon s^{-1}} F(\bar{x}, \bar{y})(\widetilde{y}^*)$, which yields (7.123). Combining finally (7.108) with (7.123) justifies the inequality "$\leq$" in (7.109) and thus completes the proof of the theorem. △

The following consequence of Theorem 7.35 and the classical Brøndsted-Rockafellar density theorem of convex analysis (see, e.g., [638, Theorem 3.17]) establish a precise formula for the exact regularity bound of a closed convex multifunction F between Banach spaces by using the coderivative of F^{-1} instead of its ε-counterparts while involving points around the reference one.

Corollary 7.36 (Calculating the Exact Regularity Bound via Coderivatives at Nearby Points). *In the setting of Theorem 7.35, we have*

$$\operatorname{reg} F(\bar{x},\bar{y}) = \lim_{\varepsilon\downarrow 0}\Big[\sup\Big\{\|D^*F^{-1}(y,x)\| \ \ \textit{over} \ \ (x,y)\in \operatorname{gph} F\cap B_\varepsilon(\bar{x},\bar{y})\Big\}\Big]. \tag{7.124}$$

Proof. To verify the inequality "$\ge$" in (7.124), observe from (7.110) that for any $\mu > \operatorname{reg} F(\bar{x},\bar{y})$ and any sufficiently small $\varepsilon > 0$, we get

$$\operatorname{dist}\big(x; F^{-1}(y)\big) \le \mu \operatorname{dist}\big(y; F(x)\big) \ \text{ for all } \ (x,y)\in B_\varepsilon(\widetilde{x},\widetilde{y})$$

whenever $(\widetilde{x},\widetilde{y})\in B_\varepsilon(\bar{x},\bar{y})$. It follows from (7.108) that

$$\mu \ge \lim_{\eta\downarrow 0}\|D^*_\eta F^{-1}(\widetilde{y},\widetilde{x})\| \ge \|D^*F^{-1}(\widetilde{y},\widetilde{x})\| \ \text{ for all } \ (\widetilde{x},\widetilde{y})\in \operatorname{gph} F\cap B_\varepsilon(\bar{x},\bar{y}).$$

This clearly implies the estimate

$$\mu \ge \lim_{\varepsilon\downarrow 0}\Big[\sup\Big\{\|D^*F^{-1}(y,x)\| \ \Big|\ (x,y)\in \operatorname{gph} F\cap B_\varepsilon(\bar{x},\bar{y})\Big\}\Big].$$

Letting there $\mu \downarrow \operatorname{reg} F(\bar{x},\bar{y})$, we arrive at the inequality "$\ge$" in (7.124).

To prove the converse inequality in (7.124), take an arbitrary $\varepsilon > 0$, and observe from Theorem 7.35 that $\operatorname{reg} F(\bar{x},\bar{y}) \le \|D^*_{\varepsilon^2}F^{-1}(\bar{y},\bar{x})\|$. This allows us to find $(x^*,y^*)\in X^*\times Y^*$ satisfying the condition $y^*\in D^*_{\varepsilon^2}F^{-1}(\bar{y},\bar{x})(x^*)$, i.e., $(-x^*,y^*)\in N_{\varepsilon^2}((\bar{x},\bar{y}); \operatorname{gph} F)$. We have furthermore that

$$\|x^*\| \le 1 \ \text{ and } \ \|y^*\| + \varepsilon \ge \operatorname{reg} F(\bar{x},\bar{y}). \tag{7.125}$$

By the Brøndsted-Rockafellar theorem, there are $(x_\varepsilon,y_\varepsilon)\in \operatorname{gph} F\cap B_\varepsilon(\bar{x},\bar{y})$ and $(-x^*_\varepsilon,y^*_\varepsilon)\in N((x_\varepsilon,y_\varepsilon); \operatorname{gph} F)$ satisfying $\|x^*_\varepsilon - x^*\|\le\varepsilon$ and $\|y^*_\varepsilon - y^*\|\le\varepsilon$. Thus we get $\|x^*_\varepsilon\| \le \|x^*\|+\varepsilon \le 1+\varepsilon$ and $\|y^*\| \le \|y^*_\varepsilon\|+\varepsilon$, and thus

$$\|y^*\| \le (1+\varepsilon)\|D^*F^{-1}(y_\varepsilon,x_\varepsilon)\| + \varepsilon.$$

Combining this with (7.125) yields the estimate

$$\operatorname{reg} F(\bar{x},\bar{y}) \le (1+\varepsilon)\sup\Big\{\|D^*F^{-1}(y,x)\| \ \Big|\ (x,y)\in \operatorname{gph} F\cap B_\varepsilon(\bar{x},\bar{y})\Big\} + 2\varepsilon,$$

which ensures the inequality "$\le$" in (7.124) while letting $\varepsilon\downarrow 0$. △

The next consequence of Theorem 7.35 concerns calculating the exact covering bound of closed- and convex-graph multifunctions. This is indeed a major result of this section, which accumulates the previous developments.

Corollary 7.37 (Calculating the Exact Covering Bound for Convex-Graph Multifunctions). *Given a point* $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ *with* $\bar{y} \in \operatorname{int}(\operatorname{rge} F)$*, the exact covering bound of* F *at* $(\bar{x}, \bar{y})$ *is calculated by*

$$\operatorname{cov} F(\bar{x}, \bar{y}) = \lim_{\varepsilon \downarrow 0} \Big[\inf_{x^* \in X^*} \inf_{y^* \in S_{Y^*}} \Big(\|x^*\| + \frac{\sigma_{\operatorname{gph} F - (\bar{x}, \bar{y})}(x^*, y^*)}{\varepsilon} \Big) \Big].$$

Proof. Define $\Omega := \operatorname{gph} F - (\bar{x}, \bar{y})$. Since the number $\operatorname{cov} F(\bar{x}, \bar{y})$ is the reciprocal of $\operatorname{reg} F(\bar{x}, \bar{y})$, it suffices to show that

$$\operatorname{reg} F(\bar{x}, \bar{y}) = \lim_{\varepsilon \downarrow 0} \Big[\sup_{x^* \in X^*} \sup_{y^* \in S_{Y^*}} \Big(\|x^*\| + \frac{\sigma_\Omega(x^*, y^*)}{\varepsilon} \Big)^{-1} \Big] =: \alpha. \tag{7.126}$$

By (7.109), we find sequences $\varepsilon_k \downarrow 0$ and $(x_k^*, y_k^*) \in X^* \times S_{Y^*}$ such that $x_k^* \in D^*_{\varepsilon_k} F(\bar{x}, \bar{y})(y_k^*)$, which amounts to $\sigma_\Omega(x_k^*, -y_k^*) \le \varepsilon_k$ due to (7.103), and that $\|x_k^*\|^{-1} \to \operatorname{reg} F(\bar{x}, \bar{y})$ as $k \to \infty$. Hence

$$\begin{aligned} \sup_{x^* \in X^*} \sup_{y^* \in S_{Y^*}} \Big(\|x^*\| + \frac{\sigma_\Omega(x^*, y^*)}{\sqrt{\varepsilon_k}} \Big)^{-1} &\ge \Big(\|x_k^*\| + \frac{\sigma_\Omega(x_k^*, -y_k^*)}{\sqrt{\varepsilon_k}} \Big)^{-1} \\ &\ge \Big(\|x_k^*\| + \sqrt{\varepsilon_k} \Big)^{-1}, \end{aligned}$$

which yields the inequality "$\le$" in (7.126) by passing to the limit as $k \to \infty$.

Conversely, if the right-hand side of (7.126) is 0, the equality in (7.126) is obvious. Otherwise, we find sequences $\widetilde{\varepsilon}_k \downarrow 0$ and $(\widetilde{x}_k^*, \widetilde{y}_k^*) \in X^* \times S_{Y^*}$ with

$$\beta < \Big(\|\widetilde{x}_k^*\| + \frac{\sigma_\Omega(\widetilde{x}_k^*, \widetilde{y}_k^*)}{\widetilde{\varepsilon}_k} \Big)^{-1} \to \alpha \ \text{ as } \ k \to \infty \tag{7.127}$$

for some $\beta > 0$. It follows that $\sigma_\Omega(\widetilde{x}_k^*, \widetilde{y}_k^*) \le \widetilde{\varepsilon}_k \beta^{-1}$ for all $k \in \mathbb{N}$, which gives us $\widetilde{x}_k^* \in D^*_{\widehat{\varepsilon}_k} F(\bar{x}, \bar{y})(-\widetilde{y}_k^*)$ with $\widehat{\varepsilon}_k := \widetilde{\varepsilon}_k \beta^{-1} \to 0$ by (7.103). Hence, we have

$$\begin{aligned} &\Big(\|\widetilde{x}_k^*\| + \frac{\sigma_\Omega(\widetilde{x}_k^*, \widetilde{y}_k^*)}{\widetilde{\varepsilon}_k} \Big)^{-1} \le \|\widetilde{x}_k^*\|^{-1} \\ &\le \sup \Big\{ \|x^*\|^{-1} \Big| \ x^* \in D^*_{\widehat{\varepsilon}_k} F(\bar{x}, \bar{y})(y^*), \ \ y^* \in S_{Y^*} \Big\}. \end{aligned} \tag{7.128}$$

Substituting the regularity formula (7.109) into (7.128) and using (7.127), we arrive at $\alpha \le \operatorname{reg} F(\bar{x}, \bar{y})$ and thus complete the proof of the corollary. △

Finally in this subsection, let us introduce an additional condition, which helps us to remove $\varepsilon > 0$ in the exact bound formula (7.108) and get the precise equality (7.130) for calculating the exact regularity bound of closed- and convex-graph multifunctions between arbitrary Banach spaces as in case (3.8) of set-valued mapping between finite-dimensional spaces. Note that assumption (8.84) holds in the SIP setting of Subsection 7.4.3 and also when $\dim Y < \infty$, while X is an arbitrary Banach space.

Theorem 7.38 (Calculating the Exact Regularity Bound via the Basic Coderivative Norm). *In the setting of Theorem 3.8, assume in addition that*

$$\Lambda(S_{Y^*}) \subset S_{Y^*}, \tag{7.129}$$

where the set $\Lambda(S_{Y^*})$ is defined sequentially by

$$\Lambda(S_{Y^*}) := \Big\{ y^* \in Y^* \Big| \; \exists\, \varepsilon_k \downarrow 0,\;\; y_k^* \in S_{Y^*} \text{ such that } \; D^*_{\varepsilon_k} F(\bar{x}, \bar{y})(y_k^*) \neq \emptyset \\ \text{and } \; y^* \; \text{ is a weak}^* \text{ cluster point of } \; y_k^* \Big\}.$$

Then the exact regularity bound is calculated by

$$\operatorname{reg} F(\bar{x}, \bar{y}) = \|D^* F^{-1}(\bar{y}, \bar{x})\|. \tag{7.130}$$

If furthermore $\operatorname{reg} F(\bar{x}, \bar{y}) > 0$*, we get the improved formula*

$$\operatorname{reg} F(\bar{x}, \bar{y}) = \sup \Big\{ \|x^*\|^{-1} \Big|\; x^* \in D^* F(\bar{x}, \bar{y})(y^*),\;\; y^* \in S_{Y^*} \Big\}. \tag{7.131}$$

Proof. Note that the equality in (7.130) is trivial when $\operatorname{reg} F(\bar{x}, \bar{y}) = 0$. Otherwise, it follows from (7.109) that there are sequences $\varepsilon_k \downarrow 0$ and $x_k^* \in D^*_{\varepsilon_k} F(\bar{x}, \bar{y})(y_k^*)$ such that $\|x_k^*\| > 0$, $\|y_k^*\| = 1$, and

$$\operatorname{reg} F(\bar{x}, \bar{y}) = \lim_{k \to \infty} \|x_k^*\|^{-1}. \tag{7.132}$$

Since the sequence $\{x_k^*\}$ is bounded by (7.132), we get from (7.129) and Alaoglu-Bourbaki theorem that there is a subnet $(x_\alpha^*, y_\alpha^*, \varepsilon_\alpha)$ of $(x_k^*, y_k^*, \varepsilon_k)$ weak* converging to some $(\bar{x}^*, \bar{y}^*, 0) \in X^* \times S_{Y^*} \times \mathbb{R}$. Note further that

$$\langle \bar{x}^*, x - \bar{x} \rangle - \langle \bar{y}^*, y - \bar{y} \rangle = \lim_{\alpha} \langle x_\alpha^*, x - \bar{x} \rangle - \langle y_\alpha^*, y - \bar{y} \rangle \le \limsup_{\alpha} \varepsilon_\alpha = 0$$

for all $(x, y) \in \operatorname{gph} F$, which yields $\bar{x}^* \in D^* F(\bar{x}, \bar{y})(\bar{y}^*)$. Moreover, the classical uniform boundedness principle tells us that $\|\bar{x}^*\| \le \liminf_{\alpha} \|x_\alpha^*\|$. This together with (7.132) ensures the validity of the inequalities

$$\operatorname{reg} F(\bar{x}, \bar{y}) \le \frac{1}{\|\bar{x}^*\|} \le \sup \left\{ \frac{1}{\|x^*\|} \Big| x^* \in D^* F(\bar{x}, \bar{y})(y^*), \|y^*\| = 1 \right\}. \tag{7.133}$$

Combining the latter with (7.109) yields (7.131). Furthermore, observe that $\widehat{x}^* := \bar{x}^* \|\bar{x}^*\|^{-1} \in S_{X^*}$ and $\widehat{y}^* := \bar{y}^* \|\bar{x}^*\|^{-1} \in D^* F^{-1}(\bar{y}, \bar{x})(\widehat{x}^*)$. Hence we get from (7.133) and (7.106) the relationships

$$\operatorname{reg} F(\bar{x}, \bar{y}) \le \|\widehat{y}^*\| = \|\bar{x}^*\|^{-1} \le \|D^* F^{-1}(\bar{y}, \bar{x})\|,$$

which together with (7.108) yield (7.130) and thus complete the proof. △

It is obvious that assumption (7.129) automatically holds when Y is finite-dimensional. More subtle, it also holds under the validity of the condition

$$\mathrm{cl}^*\Big\{y^*\in S_{Y^*}\Big|\ \sigma_\Omega(x^*,y^*)<\infty,\ x^*\in X^*\Big\}\subset S_{Y^*}, \tag{7.134}$$

with $\Omega:=\mathrm{gph}\,F-(\bar{x},\bar{y})$ due to the proper/strict inclusion

$$\begin{aligned}\Lambda(S_{Y^*})&\subset \mathrm{cl}^*\Big[\bigcup_{\varepsilon\ge 0}\Big\{y^*\in S_{Y^*}\Big|\ D^*_\varepsilon F(\bar{x},\bar{y})(y^*)\neq\emptyset\Big\}\Big]\\ &=\mathrm{cl}^*\Big\{y^*\in S_{Y^*}\Big|\ \sigma_\Omega(x^*,y^*)<\infty,\ x^*\in X^*\Big\}.\end{aligned} \tag{7.135}$$

7.4.3 Applications to Infinite Convex Constraint Systems

Here we develop applications of the results obtained in Subsection 7.4.2 to the special class of set-valued mappings $F: X \rightrightarrows Y := Z\times l^\infty(T)$ given by

$$F(x):=\begin{cases}\Big\{(z,p)\in Y\Big|\ Ax=z,\ f_t(x)\le p_t,\ t\in T\Big\} & \text{if}\ \ x\in C,\\ \emptyset & \text{otherwise,}\end{cases} \tag{7.136}$$

which describes, in particular, sets of feasible solutions in parameterized SIPs with infinitely many inequality as well as equality and geometric constraints.

The data of (7.136) are as follows: $A: X\to Z$ is a bounded linear operator between two Banach spaces; the functions $f_t: X\to\overline{\mathbb{R}}$ are l.s.c. and convex for all t from the arbitrary index set T; and C is a closed and convex subset of X with nonempty interior. These assumptions clearly imply that F in (7.136) is closed- and convex-graph multifunction, and so we can implement the results on metric regularity at $(x,(z,p))\in\mathrm{gph}\,F$ obtained above to the infinite constraint system (7.136) provided the validity of the underlying condition

$$(z,p)\in\mathrm{int}(\mathrm{rge}\,F). \tag{7.137}$$

Note that this condition clearly implies that $z\in\mathrm{int}(AX)$, which ensures that A is an open mapping, and hence it must be surjective.

Throughout this section, we denote $f(x):=\sup_{t\in T} f_t(x)$ and suppose that the space $Z\times l^\infty(T)$ is equipped with the maximum product norm

$$\|(z,p)\|=\max\Big\{\|z\|,\|p\|\Big\}\ \text{for all}\ \ z\in Z,\ \ p\in l^\infty(T).$$

As mentioned above, F is metrically regular around $(x,(z,p))\in\mathrm{gph}\,F$ if and only if condition (7.137) holds. This motivates us to introduce a qualification condition via the initial data of (7.136), which ensures the validity of (7.137) and extend the usual strong SSC typically employed for infinite linear and convex inequality systems to the more general constraint case of (7.136).

Definition 7.39 (Bounded Strong Slater Condition). *We say that the infinite system* (7.136) *satisfies the* BOUNDED STRONG SLATER CONDITION (BOUNDED SSC) *at* $(z, p) \in Z \times l^\infty(T)$ *if there is* $\widehat{x} \in \operatorname{int} C$ *such that the function* f *is bounded from above around* $\widehat{x}$*, that* $A\widehat{x} = z$*, and that*

$$\sup_{t \in T}[f_t(\widehat{x}) - p_t] < 0. \tag{7.138}$$

Note that the Slater-type notion introduced in Definition 7.39 is generally different for infinite linear and convex systems from the strong Slater condition studied and applied in Sections 7.1 and 7.3. In the particular case of $C = X$, $Z = \{0\}$, and $f_t(x) = \langle a_t^*, x\rangle - b_t$ with $(a_t^*, b_t) \in X^* \times \mathbb{R}$ considered in Section 7.1, our bounded SSC is clearly weaker than the usual SSC provided that the coefficient set $\{a_t^* |\ t \in T\}$ is bounded in X^*, which is the underlying assumption therein. The following example demonstrates that it may be *strictly weaker* even in the one-dimensional case of $X = \mathbb{R}$.

Example 7.40 (Bounded from Above Linear Constraint Functions with Unbounded Coefficients). Let $X = \mathbb{R}$, $Z = \{0\}$, $T = (0, 1)$, and $f_t(x) = -\frac{1}{t}x + t$ in (7.136). Note that

$$f_t(x) = -\frac{1}{t}x + t = -\frac{1}{t}x - t + 2t \le -2\sqrt{x} + 2t \quad \text{for all} \quad x > 0,\ t \in T.$$

Taking $\widehat{x} = 4$ and $\bar{x} = 1$, we observe that $f_t(\widehat{x}) < -2$, $f_t(\bar{x}) \le 0$, and the supremum function f is bounded from above around $\widehat{x}$. However, the coefficient set $\big\{ -\frac{1}{t} \big|\ t \in T\big\}$ is obviously unbounded.

The next proposition shows that the bounded SSC introduced is a sufficient condition for the validity of (7.137) while being in fact "almost necessary" for this, up to the upper boundedness of the supremum function f.

Proposition 7.41 (Bounded Strong Slater Condition and Metric Regularity). *Let* $(z, p) \in \operatorname{rge} F$ *for the infinite system* (7.136)*. Then the bounded SSC for* F *at* (z, p) *implies the validity of* (7.137)*. Conversely, if* (7.137) *holds, then there is* $\widehat{x} \in \operatorname{int} C$ *such that* $A\widehat{x} = z$ *and that* (7.138) *is satisfied.*

Proof. To verify the first part, suppose that the bounded SSC holds for F at (z, p). Then there are $\widehat{x} \in \imath C$ and $\varepsilon > 0$ such that the supremum function f is upper bounded around $\widehat{x}$ with $A(\widehat{x}) = z$ and $f^p(\widehat{x}) < -\varepsilon$, where

$$f^p(\cdot) := \sup_{t \in T}\big\{f_t(\cdot) - p_t\big\} \quad \text{for} \quad p \in l^\infty(T).$$

Note that the function $f^p(\cdot)$ is obviously a proper, l.s.c., convex, and upper bounded around $\widehat{x}$. We know from convex analysis that in this case it is continuous at $\widehat{x}$. Since A is surjective and $\widehat{x} \in \operatorname{int} C$, the classical open mapping theorem allows us to find $0 < s \le \frac{\varepsilon}{2}$ such that $B_s(z) \subset A(B_r(\widehat{x}) \cap C)$ for $r > 0$. Picking any

$(z', p') \in B_s(z, p)$, there exists $x \in B_r(\widehat{x}) \cap C$ with $Ax = z'$ and so that for each $t \in T$, we have

$$\begin{aligned} f^{p'}(x) &\le f^p(x) + s \le f^p(x) - f^p(\widehat{x}) + s + f^p(\widehat{x}) \\ &\le f^p(x) - f^p(\widehat{x}) + s - \varepsilon \le f^p(x) - f^p(\widehat{x}) - \varepsilon/2 \le 0 \end{aligned}$$

when r is sufficiently small. This yields $(z', p') \in \operatorname{rge} F$, which implies in turn that the inclusion $B_s(z, p) \subset \operatorname{rge} F$ holds.

To justify the necessity part, observe that $(z, (p_t - \varepsilon)_{t\in T}) \in \operatorname{rge} F$ for some $\varepsilon > 0$ if $(z, p) \in \operatorname{int}(\operatorname{rge} F)$. Hence there is $\widehat{x} \in X$ such that $A\widehat{x} = z$ and $f_t(\widehat{x}) - p_t \le -\varepsilon$ as $t \in T$, which thus completes the proof. △

Now we proceed with calculating the exact regularity bound for the constraint system (7.136) at $(\bar{x}, (\bar{z}, 0)) \in \operatorname{gph} F$ based on the results of Subsection 7.4.2. It follows from Theorem 7.35 that $\operatorname{reg} F(\bar{x}, (\bar{z}, 0))$ can be calculated via the norms of ε-coderivatives. The next result, which is certainly of its own interest, accomplishes an important step in this direction.

Theorem 7.42 (Explicit Form of ε-Coderivatives for Infinite Convex Systems). *Let F be the infinite constraint system (7.136), and let $(\bar{x}, (\bar{z}, 0)) \in \operatorname{gph} F$. Then for each $\varepsilon \ge 0$, we have the ε-coderivative representation*

$$D_\varepsilon^* F\big(\bar{x}, (\bar{z}, 0)\big)\big(S_{(Z\times l^\infty(T))^*}\big) = \big\{x^* \big| (x^*, \langle x^*, \bar{x}\rangle + \varepsilon) \in M\big\}, \tag{7.139}$$

where $x^* \in X^*$ and M is defined, with $C_0 := C \cap \operatorname{dom} f$, by

$$M := \bigcup_{z^*\in \mathbb{B}_{Z^*}} \operatorname{cl}^*\Big[(1 - \|z^*\|) \operatorname{co}\Big(\bigcup_{t\in T} \operatorname{epi} f_t^*\Big) + \operatorname{epi} \delta^*(\cdot; C_0)\Big] + (A^*z^*, \langle z^*, \bar{z}\rangle).$$

Proof. To verify the inclusion "$\subset$" in (7.139), pick $(z^*, p^*) \in S_{(Z\times l^\infty(T))^*}$ and $x^* \in D_\varepsilon^* F(\bar{x}, (\bar{z}, 0))(z^*, p^*)$. Then we have $\|z^*\| + \|p^*\| = 1$ and

$$\langle x^*, x - \bar{x}\rangle - \langle z^*, z - \bar{z}\rangle - \langle p^*, p\rangle \le \varepsilon \ \text{ for all } \ (x, z, p) \in \operatorname{gph} F,$$

which can be equivalently represented by

$$\begin{aligned} &\langle x^* - A^*z^*, x - \bar{x}\rangle - \langle p^*, p\rangle \le \varepsilon \\ &\text{if } (x, p) \in C_0 \times l^\infty(T),\ f_t(x) - \langle \delta_t, p\rangle \le 0,\ t \in T, \end{aligned} \tag{7.140}$$

via the Dirac measure $\delta_t \in (l^\infty(T))^*$ at t. It follows from the extended Farkas lemma in Proposition 7.3 that (7.140) reads as

$$\begin{aligned} &\big(p^*, x^* - A^*z^*, \langle x^* - A^*z^*, \bar{x}\rangle + \varepsilon\big) \\ &\in \operatorname{cl}^*\Big[\operatorname{cone}\Big\{\bigcup_{t\in T} \{\delta_t\} \times \operatorname{epi} f_t^*\Big\} + \{0\} \times \operatorname{epi} \delta^*(\cdot; C_0)\Big]. \end{aligned} \tag{7.141}$$

Hence there exist nets $\{\lambda_\nu\}_{\nu\in\mathcal{N}} \subset \mathbb{R}_+^{(T)}$, $\{(v_\nu^*, s_\nu)\}_{\nu\in\mathcal{N}} \subset \operatorname{epi} \delta^*(\cdot; C_0)$, and $\{(u_{t\nu}^*, r_{t\nu})\}_{\nu\in\mathcal{N}} \subset \operatorname{epi} f_t^*$ for each $t \in T$ such that

$$\big(p^*, x^* - A^*z^*, \langle x^* - A^*z^*, \bar{x}\rangle + \varepsilon\big) = w^*\text{-}\lim_{\nu\in\mathcal{N}}\Big[\sum_{t\in T}\lambda_{t\nu}(\delta_t, u^*_{t\nu}, r_{t\nu}) + (0, v^*_\nu, s_\nu)\Big].$$

Observe from the latter equality that $p^* = w^*\text{-}\lim_{\nu\in\mathcal{N}}\sum_{t\in T}\lambda_{t\nu}\delta_t$. Thus we have

$$\begin{aligned}&\limsup_{\nu\in\mathcal{N}}\sum_{t\in T}\lambda_{t\nu} \ge \sup_{\|p\|\le 1}\lim_{\nu\in\mathcal{N}}\sum_{t\in T}\lambda_{t\nu}p_t\\ &= \sup\nolimits_{\|p\|\le 1}\langle p^*, p\rangle = \|p^*\| \ge \langle p^*, e\rangle = \lim_{\nu\in\mathcal{N}}\sum_{t\in T}\lambda_{t\nu}\end{aligned} \tag{7.142}$$

with $e \in l^\infty(T)$ satisfying $e_t = 1$ for all $t \in T$. This yields

$$1 - \|z^*\| = \|p^*\| = \lim_{\nu\in\mathcal{N}}\sum_{t\in T}\lambda_{t\nu}. \tag{7.143}$$

If $\|z^*\| = 1$, we get from the above the relationships

$$\begin{aligned}&\langle x^* - A^*z^*, x - \bar{x}\rangle - \varepsilon = \langle x^* - A^*z^*, x\rangle - (\langle x^* - A^*z^*, \bar{x}\rangle + \varepsilon)\\ &= \lim_{\nu\in\mathcal{N}}\Big[\sum_{t\in T}\lambda_{t\nu}\langle u^*_{t\nu}, x\rangle + \langle v^*_\nu, x\rangle\Big] - \lim_{\nu\in\mathcal{N}}\Big[\sum_{t\in T}\lambda_{t\nu}r_{t\nu} - s_\nu\Big]\\ &\le \limsup_{\nu\in\mathcal{N}}\Big[\sum_{t\in T}\lambda_{t\nu}\big(\langle u^*_{t\nu}, x\rangle - f_t(x) - r_{t\nu} + f(x)\big) + \langle v^*_\nu, x\rangle - s_\nu\Big]\\ &\le \limsup_{\nu\in\mathcal{N}}\Big[\sum_{t\in T}\lambda_{t\nu}\big(f^*_t(u^*_{t_\nu}) - r_{t\nu} + f(x)\big) + \delta^*(\cdot; C_0)(v^*_\nu) - s_\nu\Big]\\ &\le \limsup_{\nu\in\mathcal{N}}\sum_{t\in T}\lambda_{t\nu}f(x) = 0 \ \text{ for any } \ x \in C_0.\end{aligned}$$

It follows from (7.30) that $(x^* - A^*z^*, \langle x^* - A^*z^*, \bar{x}\rangle + \varepsilon) \in \operatorname{epi}\delta^*(\cdot; C_0)$; so

$$\begin{aligned}(x^*, \langle x^*, \bar{x}\rangle + \varepsilon) &\in \operatorname{epi}\delta^*(\cdot; C_0) + (A^*z^*, \langle A^*z^*, \bar{x}\rangle)\\ &= \operatorname{epi}\delta^*(\cdot; C_0) + (A^*z^*, \langle z^*, \bar{z}\rangle) \subset M.\end{aligned}$$

If $\|z^*\| < 1$, it doesn't restrict the generality due to (7.143) to suppose that $\sum_{t\in T}\lambda_{t\nu} > 0$ for all $\nu \in \mathcal{N}$ and to define $\widetilde{\lambda}_{t\nu} := \dfrac{\lambda_{t\nu}}{\sum_{t'\in T}\lambda_{t'\nu}}$ for each $t \in T$ and $\nu \in \mathcal{N}$. It tells us by the "w^*-lim" expression after formula (7.141) that

$$\begin{aligned}&(x^*, \langle x^*, \bar{x}\rangle + \varepsilon) = w^*\text{-}\lim_{\nu\in\mathcal{N}}\Big[\sum_{t\in T}\lambda_{t\nu}(u^*_{t\nu}, r_{t\nu}) + (v^*_\nu, s_\nu)\Big] + (A^*z^*, \langle z^*, \bar{z}\rangle)\\ &= (1 - \|z^*\|)w^*\text{-}\lim_{\nu\in\mathcal{N}}\Big[\sum_{t\in T}\widetilde{\lambda}_{t\nu}(u^*_{t\nu}, r_{t\nu}) + (v^*_\nu, s_\nu)\Big] + (A^*z^*, \langle z^*, \bar{z}\rangle) \subset M.\end{aligned}$$

Thus we get $(x^*, \langle x^*, \bar{x}\rangle + \varepsilon) \in M$ and justify the inclusion "$\subset$" in (7.139).

To verify the converse inclusion in (7.139), pick any element $x^* \in X^*$ satisfying $(x^*, \langle x^*, \bar{x}\rangle + \varepsilon) \in M$. Hence we find a unit functional $z^* \in \mathbb{B}_{Z^*}$ as well as nets $\{\lambda_\nu\}_{\nu\in\mathcal{N}} \subset \mathbb{R}_+^{(T)}$, $\{(v_\nu^*, s_\nu)\}_{\nu\in\mathcal{N}} \subset \operatorname{epi}\delta^*(\cdot; C_0)$, and $\{(u_{t\nu}^*, r_{t\nu})\}_{\nu\in\mathcal{N}} \subset \operatorname{epi} f_t^*$, $t \in T$, such that $\sum_{t\in T}\lambda_{t\nu} = 1$ and

$$(x^*, \langle x^*, \bar{x}\rangle + \varepsilon) = (1 - \|z^*\|)w^*\text{-}\lim_{\nu\in\mathcal{N}}\Big[\sum_{t\in T}\lambda_{t\nu}(u_{t\nu}^*, r_{t\nu}) + (v_\nu^*, s_\nu)\Big]$$
$$+(A^*z^*, \langle z^*, \bar{z}\rangle).$$

Defining $p_\nu^* := (1 - \|z^*\|)\sum_{t\in T}\lambda_{t\nu}\delta_t$, deduce that $\|p_\nu^*\| = 1 - \|z^*\|$ while arguing similarly to the proof of (7.142). It follows from the classical Alaoglu-Bourbaki theorem that there exists a subnet of p_ν^* (without relabeling), which weak* converges to some $p^* \in \mathbb{B}_{(l^\infty(T))^*}$. By using again the arguments as in the proof of (7.142), we get $\|p^*\| = 1 - \|z^*\|$ and then obtain (7.141). Due to the equivalence between (7.140) and (7.141), this justifies the inclusion "$\supset$" in (7.139) and thus completes the proof of the theorem. $\triangle$

In the *coderivative* case of Theorem 7.42 (i.e., if $\varepsilon = 0$), we can equivalently modify the representation in (7.139) and provide its further specification.

Proposition 7.43 (Explicit Forms of the Coderivative for Infinite Convex Systems). *Let* $(\bar{x}, (\bar{z}, 0)) \in \operatorname{gph} F$ *for the constraint system* (7.136). *Then we have the coderivative representation*

$$D^*F\big(\bar{x}, (\bar{z}, 0)\big)\big(S_{(Z\times l^\infty(T))^*}\big) = \big\{x^* \in X^* \big|\ (x^*, \langle x^*, \bar{x}\rangle) \in L\big\} \tag{7.144}$$

with $L := \bigcup_{z^*\in\mathbb{B}_{Z^*}} \operatorname{cl}^*\Big[(1-\|z^*\|)\operatorname{co}\Big(\bigcup_{t\in T}\operatorname{gph} f_t^*\Big) + \operatorname{gph}\delta^*(\cdot; C_0)\Big] + (A^*z^*, \langle z^*, \bar{z}\rangle).$
Furthermore, the term $\operatorname{gph}\delta^*(\cdot; C_0)$ *above can be dropped if* $\bar{x} \in \operatorname{int} C_0$.

Proof. To verify the inclusion "$\subset$" in (7.144), for any $x^* \in D^*F(\bar{x}, (\bar{z}, 0))(z^*, p^*)$ with $\|z^*\| + \|p^*\| = 1$, we deduce from the proof of Theorem 7.42 the validity of inclusion (7.140) with $\varepsilon = 0$. This allows us to find nets $\{\lambda_\nu\}_{\nu\in\mathcal{N}} \subset \mathbb{R}_+^{(T)}$, $\{\rho_\nu\}_{\nu\in\mathcal{N}} \subset \mathbb{R}_+$, $\{(v_\nu^*, s_\nu)\}_{\nu\in\mathcal{N}} \subset \operatorname{gph}\delta^*(\cdot; C_0)$, and $\{(u_{t\nu}^*, r_{t\nu})\}_{\nu\in\mathcal{N}} \subset \operatorname{gph} f_t^*$ for each $t \in T$ providing the limiting representation

$$\begin{aligned}\big(p^*, x^* - A^*z^*, \langle x^* - A^*z^*, \bar{x}\rangle\big) = w^*\text{-}\lim_{\nu\in\mathcal{N}}\sum_{t\in T}\lambda_{t\nu}(\delta_t, u_{t\nu}^*, r_{t\nu})\\ +(0, v_\nu^*, s_\nu) + (0, 0, \rho_\nu).\end{aligned} \tag{7.145}$$

Similarly to the proof of Theorem 7.42, suppose without loss of generality that $\sum_{t\in T}\lambda_{t\nu} = 1 - \|z^*\|$ for all $\nu \in \mathcal{N}$ and then get

$$r_{t\nu} = f_t^*(u_{t\nu}^*) \geq \langle u_{t\nu}^*, \bar{x}\rangle - f_t(\bar{x}) \geq \langle u_{t\nu}^*, \bar{x}\rangle \ \text{ and } \ s_\nu = \delta^*(\cdot; C_0)(v_\nu^*) \geq \langle v_\nu^*, \bar{x}\rangle.$$

This implies together with (7.145) the relationships

$$\langle x^* - A^*z^*, \bar{x}\rangle = \lim_{\nu\in\mathcal{N}}\Big[\sum_{t\in T}\lambda_{t\nu}r_{t\nu}+s_\nu+\rho_\nu\Big] \ge \limsup_{\nu\in\mathcal{N}}\Big[\sum_{t\in T}\lambda_{t\nu}\langle u^*_{t\nu},\bar{x}\rangle+\langle v^*_\nu,\bar{x}\rangle$$
$$+\rho_\nu\Big] \ge \langle x^* - A^*z^*, \bar{x}\rangle + \limsup_{\nu\in\mathcal{N}}\rho_\nu,$$

which ensure that $\limsup_{\nu\in\mathcal{N}}\rho_\nu = 0$. Then it follows from (7.145) that

$$(x^*, \langle x^*, \bar{x}\rangle) = w^*\text{-}\lim_{\nu\in\mathcal{N}}\sum_{t\in T}\lambda_{t\nu}(u^*_{t\nu}, r_{t\nu}) + (v^*_\nu, s_\nu) + (A^*z^*, \langle z^*, \bar{z}\rangle) \in L,$$

and thus we arrive at the inclusion "$\subset$" in (7.144). The verification of the opposite inclusion in (7.144) follows the lines in the proof of Theorem 7.42.

Finally, let $\bar{x} \in \operatorname{int} C_0$ and pick $x^* \in D^*F(\bar{x},(\bar{z},0))(z^*,p^*)$ with $(z^*,p^*) \in S_{(Z\times l^\infty(T))^*}$. Using the notation from the proof of (7.144) above, we have

$$0 = \langle x^* - A^*z^*, \bar{x}\rangle - \langle x^* - A^*z^*, \bar{x}\rangle = \lim_{\nu\in\mathcal{N}}\Big[\sum_{t\in T}\lambda_{t\nu}(\langle u^*_{t\nu},\bar{x}\rangle - r_{t\nu})$$
$$\langle v^*_\nu,\bar{x}\rangle - s_{t\nu}\Big] \le -\limsup_{\nu\in\mathcal{N}}\sup_{x\in C_0}\big[\langle v^*_\nu, x\rangle - \langle v^*_\nu,\bar{x}\rangle\big] \le -\limsup_{\nu\in\mathcal{N}}\eta\|v^*_\nu\|,$$

where $\eta > 0$ is such that $B_\eta(\bar{x}) \subset C_0$. This implies that $\limsup_{\nu\in\mathcal{N}}\|v^*_\nu\| = 0$, and so we can remove $\operatorname{gph}\delta^*(\cdot; C_0)$ in the representation of L in (7.144). △

The next major result provides a precise calculation of the exact regularity bound of the infinite constraint system (7.136) entirely via its initial data.

Theorem 7.44 (Exact Regularity Bound of Infinite Constraint Systems). *Given* $(\bar{x},(\bar{z},0)) \in \operatorname{gph} F$ *for the infinite system in* (7.136)*, assume that the bounded SSC from Definition* 7.39 *holds at* $(\bar{z},0)$. *Then the exact regularity bound of F at* $(\bar{x},(\bar{z},0))$ *is calculated by*

$$\operatorname{reg} F(\bar{x},(\bar{z},0)) = \lim_{\varepsilon\downarrow 0}\Big[\sup\big\{\|x^*\|^{-1}\,\big|\,(x^*, \langle x^*,\bar{x}\rangle + \varepsilon) \in M\big\}\Big], \tag{7.146}$$

where M is defined in Theorem 7.42. *If in addition* $0 < \dim Z < \infty$, *then*

$$\begin{aligned}\operatorname{reg} F(\bar{x},(\bar{z},0)) &= \|D^*F^{-1}((\bar{z},0),\bar{x})\|\\ &= \sup\Big\{\|x^*\|^{-1}\,\Big|\,(x^*,\langle x^*,\bar{x}\rangle) \in L\Big\},\end{aligned} \tag{7.147}$$

where the set L is defined in Proposition 7.43.

Proof. It follows from Proposition 7.41 that $(\bar{z},0) \in \operatorname{int}(\operatorname{rge} F)$, i.e., the mapping F is metrically regular around $(\bar{x},(\bar{z},0))$. Substituting the ε-coderivative expression from Theorem 7.42 into the exact bound formula (7.109) of Theorem 7.35, we arrive at the limiting representation (7.146).

Let us now justify the equalities in (7.147) under the finite dimensionality of Z. By Theorem 7.38 and Proposition 7.43, we need to check that (7.129) holds and that $\operatorname{reg} F(\bar{x},(\bar{z},0)) > 0$. To proceed, take any $\varepsilon > 0$ and $(z^*,p^*) \in S_{(Z\times l^\infty(T))^*}$

satisfying $D_\varepsilon F(\bar{x}, (\bar{z}, 0))(z^*, p^*) \neq \emptyset$. By the same arguments as in the proofs of (7.141) and (7.143), we get the inclusion

$$p^* \in \big(1 - \|z^*\|\big)\,\mathrm{cl}^*\mathrm{co}\,\big\{\delta_t \big|\ t \in T\big\}.$$

It shows that the set $\mathrm{cl}^*\{(z^*, p^*) \in S_{(Z\times l^\infty(T))^*}|\ D^*_\varepsilon F(\bar{x}, (\bar{z}, 0))(z^*, p^*) \neq \emptyset\}$ is contained in the following one:

$$\mathrm{cl}^* \bigcup_{z^* \in \mathbb{B}_{Z^*}} \Big[\big\{z^*\big\} \times \big(1 - \|z^*\|\big)\,\mathrm{cl}^*\mathrm{co}\,\big\{\delta_t \big|\ t \in T\big\}\Big]. \tag{7.148}$$

Further, we deduce from the proof of (7.143) that $\mathrm{cl}^*\mathrm{co}\,\{\delta_t|\ t \in T\} \subset S_{(l^\infty(T))^*}$. Since $\dim Z < \infty$, the latter implies that the set in (7.148) is a subset of $S_{(Z\times l^\infty(T))^*}$, which ensures the validity of (7.129).

It remains to verify that $\mathrm{reg}\, F(\bar{x}, (\bar{z}, 0)) > 0$. We can easily see that

$$D^*F^{-1}\big((\bar{z}, 0), \bar{x}\big)(x^*) \supset \big\{(z^*, 0) \in Z^* \times (l^\infty(T))^* \big|\ A^*z^* = x^*\big\}.$$

Since the operator A is surjective, we clearly have $\|(A^*)^{-1}\| > 0$. This allows us to conclude that $\|D^*F^{-1}((\bar{z}, 0), \bar{x})\| > 0$, which yields $\mathrm{reg}\, F(\bar{x}, (\bar{z}, 0)) > 0$ by Theorem 7.35 and thus completes the proof. △

It immediately follows from Theorem 7.38 that the exact bound formula

$$\mathrm{reg}\, F\big(\bar{x}, (\bar{z}, 0)\big) = \big\|D^*F^{-1}\big((\bar{z}, 0), \bar{x}\big)\big\| \tag{7.149}$$

holds also in the case of $\dim Z = 0$. Recall that the Lipschitzian counterpart of (7.149) is proved for infinite linear inequality systems in Corollary 7.16 and for infinite convex inequality systems in Theorem 7.31 (with no equality and geometric constraints) under the boundedness assumptions therein. As discussed above, these assumptions are essentially stronger than the bounded SSC imposed in Theorem 7.44; see Example 7.40.

A natural question arising from Theorem 7.44 is whether the exact regularity bound expression (7.149) holds for infinite-dimensional spaces Z. The following *counterexample* is constructed for the case of the classical Asplund space $Z = c_0$, which has been already used above (i.e., the space of sequences of real numbers converging to zero and endowed with the supremum norm).

Example 7.45 (Failure of the Exact Bound Formula for Countable Inequality and Equality Constraints in Asplund Spaces.) Let $X = Z = c_0$ and $T = \mathbb{N}$. Define a linear operator $A\colon X \to Z$ by $Ax := (x_2, x_3, \ldots)$ for all $x = (x_1, x_2, \ldots) \in X$. It is easy to see that A is bounded and surjective. We form a set-valued mapping $F\colon c_0 \rightrightarrows c_0 \times l^\infty$ of type (7.136) by

$$F(x) := \big\{(z, p) \in Z \times l^\infty \big|\ Ax = z, x_1 + x_n + 1 \le p_n, n \in \mathbb{N}\big\} \tag{7.150}$$

for any $x \in X$. Take $\bar{x} := \big(-\frac{1}{n}\big)_{n\in\mathbb{N}}$, $\bar{z} := A\bar{x}$, and $\widehat{x} := (-2, -\frac{1}{2}, -\frac{1}{3}, \ldots) \in X$. Observe that the bounded strong Slater condition of Definition 7.39 is satisfied at $\widehat{x}$ for (7.150) and that $\bar{x} \in F^{-1}(\bar{z}, 0)$. Defining further

$$x^k := \Big(-1, -\frac{1}{2}, \ldots, -\frac{1}{k-1}, \frac{1}{k}, -\frac{1}{k+1}, -\frac{1}{k+2}, \ldots\Big),$$
$$z^k := \Big(-\frac{1}{2}, \ldots, -\frac{1}{k-1}, \frac{2}{k}, -\frac{1}{k+1}, -\frac{1}{k+2}, \ldots\Big)$$

shows that $x^k \to \bar{x}$ and $z^k \to \bar{z}$ in c_0. Moreover, we have the equalities

$$\operatorname{dist}\big((z^k, 0); F(x^k)\big) = \max\Big\{\sup_n (x_1^k + x_n^k + 1)_+, \sup_n |\, x_{n+1}^k - z_n^k|\Big\} = \frac{1}{k}$$

with $\alpha_+ = \max\{0, \alpha\}$ as usual. It is easy to calculate the inverse mapping value $F^{-1}(z^k, 0) = \{(a, z_1^k, z_2^k, \ldots) \in c_0 |\; a \le -\frac{2}{k} - 1\}$, which gives us

$$\operatorname{dist}\big(x^k; F^{-1}(z^k, 0)\big) = \max\Big\{\Big(x_1^k + \frac{2}{k} + 1\Big)_+, \sup_n |x_{n+1}^k - z_n^k|\Big\} = \frac{2}{k}.$$

It follows from the distance expressions above that $\operatorname{reg} F(\bar{x}, (\bar{z}, 0)) \ge 2$. Thus the exact bound formula (7.149) fails if we show that

$$\|x^*\| \ge 1 \;\text{ for all }\; x^* \in D^*F\big(\bar{x}, (\bar{z}, 0)\big)\big(S_{(Z\times l^\infty)^*}\big). \tag{7.151}$$

To verify (7.151), employ the explicit coderivative form from Proposition 7.43 that gives some us $z^* \in \mathbb{B}_{Z^*}$ with

$$(x^*, \langle x^*, \bar{x}\rangle) \in \operatorname{cl}^*\Big[(1 - \|z^*\|)\operatorname{co}\big\{(\delta_1 + \delta_n, -1)\big|\; n \in \mathbb{N}\big\}\Big] + \big(A^*z^*, \langle z^*, \bar{z}\rangle\big),$$

where $\delta_n \in c_0^*$ and $\langle \delta_n, x\rangle = x_n$ for all $x \in c_0$ and $n \in \mathbb{N}$. Hence there is a net $(\lambda_\nu)_{\nu\in\mathcal{N}} \subset \mathbb{R}^{(\mathbb{N})}$ such that $\sum_{n\in\mathbb{N}} \lambda_{n\nu} = 1 - \|z^*\|$ for all $\nu \in \mathcal{N}$ and that

$$(x^*, \langle x^*, \bar{x}\rangle) = w^*\text{-}\lim_{\nu\in\mathcal{N}} \sum_{n\in\mathbb{N}} \lambda_{n\nu}(\delta_1 + \delta_n, -1) + (A^*z^*, \langle z^*, \bar{z}\rangle),$$

which readily implies the limiting relationships

$$0 = \lim_{\nu\in\mathcal{N}} \sum_{n\in\mathbb{N}} \lambda_{n\nu}(-\langle \delta_1 + \delta_n, \bar{x}\rangle - 1) = \lim_{\nu\in\mathcal{N}} \sum_{n\in\mathbb{N}} \frac{\lambda_{n\nu}}{n}. \tag{7.152}$$

Since $c_0^* = l_1$, we write z^* in the form $(z_1^*, z_2^*, \ldots) \in l_1$ and observe that $A^*z^* = (0, z_1^*, z_2^*, \ldots) \in l_1$. Thus for any $\varepsilon > 0$, there is $k \in \mathbb{N}$ sufficiently large and such that $\sum_{n=k+1}^{\infty} |z_n^*| \le \varepsilon$, which ensures that $\|A^*z^* - \widehat{z}_k^*\| \le \varepsilon$ with $\widehat{z}_k^* := (0, z_1^*, \ldots, z_k^*, 0, 0, \ldots) \in l_1$. Define further $\widehat{x}_k^*$ by

$$\widehat{x}_k^* := w^*\text{-}\lim_{\nu\in\mathcal{N}}\sum_{n\in\mathbb{N}}\lambda_{n\nu}(\delta_1+\delta_n)+\widehat{z}_k^*,$$

take $e^k := (1, \operatorname{sign}(z_1^*), \ldots, \operatorname{sign}(z_k^*), 0, \ldots) \in c_0$, and get $\|e^k\| = 1$ with

$$\begin{aligned}\|\widehat{x}_k^*\| \ge \langle \widehat{x}_k^*, e^k\rangle &= \lim_{\nu\in\mathcal{N}}\sum_{n\in\mathbb{N}}\lambda_{n\nu}(e_1^k+e_n^k)+\sum_{n=1}^{k} z_n^* e_{n+1}^k\\ &\ge \lim_{\nu\in\mathcal{N}}\sum_{n\in\mathbb{N}}\lambda_{n\nu}+\sum_{n=1}^{k}|z_n^*|-\limsup_{\nu\in\mathcal{N}}\sum_{n=1}^{k+1}\lambda_{n\nu}.\end{aligned}\tag{7.153}$$

It follows from the equations in (7.152) that

$$0 \le \limsup_{\nu\in\mathcal{N}}\sum_{n=1}^{k+1}\lambda_{n\nu} \le (k+1)\limsup_{\nu\in\mathcal{N}}\sum_{n\in\mathbb{N}}\frac{\lambda_{n\nu}}{n}=0.$$

Combining this with (7.153) gives us the estimates

$$\|\widehat{x}_k^*\| \ge 1-\|z^*\|+\sum_{n=1}^{k}|z_n^*| \ge 1-\|z^*\|+\|z^*\|-\varepsilon = 1-\varepsilon.$$

It is clear furthermore that $\|x^*-\widehat{x}_k^*\| = \|A^*z^*-\widehat{z}_k^*\| \le \varepsilon$. Thus we arrive at

$$\|x^*\| \ge \|\widehat{x}_k^*\|-\|x^*-\widehat{x}_k^*\| \ge 1-\varepsilon-\varepsilon = 1-2\varepsilon \quad\text{for all}\quad \varepsilon>0,$$

yielding $\|x^*\| \ge 1$ and (7.151). This confirms the failure of (7.149).

The next example shows that the formula (7.149) for calculating the exact regularity bound fails when $\dim Z = \infty$ even for constraint systems (7.136) with a *single* convex inequality, while both spaces X and Z are Asplund.

Example 7.46 (Failure of the Exact Bound Formula for Single Inequality and Infinite-Dimensional Equality Constraints). Let $X = Z = c_0$ and $T = \{1\}$. Define the linear operator $A : X \to Z$ as in Example 7.45, and consider $F : X \rightrightarrows Z\times\mathbb{R}$ given by

$$F(x) := \big\{(z,p)\in Z\times\mathbb{R}\big|\ Ax = z,\ \ f(x)\le p\big\}\ \text{ for any } x\in X,$$

where $f(x) := \sup\{x_1+x_n+1|\ n\in\mathbb{N}\}$ with $\operatorname{dom} f = X$. Then we have

$$\operatorname{dist}\big((z^k,0);F(x^k)\big)=k^{-1}\ \text{ and }\ \operatorname{dist}\big(x^k;F^{-1}(z^k,0)\big)=2k^{-1}$$

in the notation of Example 7.45, and so $\operatorname{reg} F(\bar{x},(\bar{z},0)) \ge 2$. Also

$$\operatorname{epi} f^* = \operatorname{cl}^*\operatorname{co}\big\{(\delta_1+\delta_n,-1)\big|\ n\in\mathbb{N}\big\}+\{0\}\times\mathbb{R}_+,\tag{7.154}$$

which follows from the well-known formula for general supremum functions:

$$\text{epi } f^* = \text{cl}^*\text{co} \bigcup_{t\in T} \Big(\text{epi } f_t^*\Big). \tag{7.155}$$

Picking now any $x^* \in D^*F(\bar{x}, (\bar{z}, 0))(S_{(Z\times\mathbb{R})^*})$ and using Theorem 7.42 together with representation (7.155), we arrive at

$$(x^*, \langle x^*, \bar{x}\rangle) \in \text{cl}^*\Big[(1 - \|z^*\|)\,\text{co}\,\big\{(\delta_1 + \delta_n, -1)\big|\; n \in \mathbb{N}\big\}\Big] + (A^*z^*, \langle z^*, \bar{z}\rangle).$$

As in Example 7.45, this gives us $\|x^*\| \geq 1$, and thus (7.149) fails.

The following result provides efficient conditions, which ensure the validity of the major regularity formula (7.149) when $\dim Z = \infty$. The given proof is different from that of (7.147) in Theorem 7.44 with $\dim Z < \infty$. In particular, it doesn't rely on condition (7.129) that may not hold. Indeed, even in the simplest setting of $T = \emptyset$, the left-hand side of (7.129) is $\text{cl}^* S_{Z^*}$, which is obviously not a subset of S_{Z^*} when $\dim Z = \infty$.

Theorem 7.47 (Exact Bound Formula for Finite Inequality and Infinite Equality Constraints). *In the case of arbitrary Banach spaces X and Z in* (7.136)*, assume that the index set T is finite, that*

$$f_t(x) = \langle a_t^*, x\rangle - b_t \;\; \textit{for all} \;\; x \in X,\; t \in T \;\; \textit{with} \;\; (a_t^*, b_t) \in X^* \times \mathbb{R},$$

and that, given $(\bar{x}, (\bar{z}, 0)) \in \text{gph}\, F$*, the constraint mapping F satisfies the standard Slater condition at* $(\bar{z}, 0)$ *with* $\bar{x} \in C$*. Then formula* (7.149) *holds.*

Proof. Letting $T := \{1, \ldots, k\}$, observe that $\text{dom}\, f = X$ and so $C_0 = C$ in the notation of Theorem 7.42. Since we obviously have

$$\text{epi } f_n^* = (a_n^*, b_n) + \{0\} \times \mathbb{R}_+ \;\; \text{and} \;\; \{0\} \times \mathbb{R}_+ + \text{epi}\, \delta^*(\cdot; C) \subset \text{epi}\, \delta^*(\cdot; C)$$

for any $z^* \in \mathbb{B}_{Z^*}$ and $n \in \{1, \ldots, k\}$, it follows that

$$\begin{aligned}\big(1 - \|z^*\|\big)\text{co}\big\{\text{epi } f_t^*\big| \in T\big\} + \text{epi}\, \delta^*(\cdot; C_0) = \big(1 - \|z^*\|\big)&\text{co}\big\{(a_n^*, b_n)\big|1 \leq n \leq k\big\}\\ &+\text{epi}\, \delta^*(\cdot; C).\end{aligned}$$

The latter set is clearly weak* closed in $X^* \times \mathbb{R}$, and hence the set M in Theorem 7.42 is represented by

$$M = \bigcup_{z^*\in\mathbb{B}_{Z^*}} \Big\{(1 - \|z^*\|)\text{co}\big\{(a_n^*, b_n)\big|1 \leq n \leq k\big\} + \text{epi}\, \delta^*(\cdot; C) + (A^*z^*, \langle z^*, \bar{z}\rangle)\Big\}.$$

Invoking now the result in the first part of Theorem 7.44, we find sequences of $x_m^* \in X^*$, $\lambda^m \in \mathbb{R}_+^k$, $(v_m^*, s_m) \in \text{epi}\, \delta^*(\cdot; C)$, and $z_m^* \in \mathbb{B}_{Z^*}$ for all $m \in \mathbb{N}$ such that $\sum_{n=1}^k \lambda_n^m = 1 - \|z_m^*\|$ and

$$\Big(x_m^*, \langle x_m^*, \bar{x}\rangle + m^{-1}\Big) = \sum_{n=1}^{k} \lambda_n^m (a_n^*, b_n) + (v_m^*, s_m) \\ +(A^* z_m^*, \langle z_m^*, \bar{z}\rangle) \tag{7.156}$$

with the upper estimate of the regularity bound

$$\operatorname{reg} F\big(\bar{x}, (\bar{z}, 0)\big) \le \|x_m^*\|^{-1} + o(1) = \Big\| \sum_{n=1}^{k} \lambda_n^m a_n^* + v_m^* + A^* z_m^* \Big\|^{-1} + o(1).$$

It follows from considering the second components in (7.156) that

$$\frac{1}{m} = \langle x_m^*, \bar{x}\rangle + \frac{1}{m} - \langle x_m^*, \bar{x}\rangle = \sum_{n=1}^{k} \lambda_n^m b_n + s_m - \sum_{n=1}^{k} \lambda_n^m \langle a_n^*, \bar{x}\rangle - \langle v_m^*, \bar{x}\rangle$$
$$\ge \sum_{n=1}^{k} \lambda_n^m (b_n - \langle a_n^*, \bar{x}\rangle) + s_m - \langle v_m^*, \bar{x}\rangle \ge \sum_{n=1}^{k} \lambda_n^m (b_n - \langle a_n^*, \bar{x}\rangle) \ge 0.$$

Since $\|\lambda^m\| \le 1$, we suppose that $\lambda^m \to \lambda \in \mathbb{R}_+^k$ as $m \to \infty$ and thus deduce from the above that $\sum_{n=1}^{k} \lambda_n (b_n - \langle a_n^*, \bar{x}\rangle) = 0$ by passing to the limit as $m \to \infty$. Defining further the sequences of

$$\varepsilon_m := \sum_{n=1}^{k} |\lambda_n^m - \lambda_n|, \;\; \eta_m := \sum_{n=1}^{k} \lambda_n + \|z_m^*\|, \;\; \widehat{x}_m^* := \sum_{n=1}^{k} \lambda_n a_n^* + A^* z_m^*,$$

note that $\varepsilon_m = o(1)$ and $\eta_m = 1 - o(1)$. Then Proposition 7.43 tells us that

$$\eta_m^{-1} \widehat{x}_m^* \in D^* F\big(\bar{x}, (\bar{z}, 0)\big)\big(S_{(Z\times\mathbb{R}^k)^*}\big).$$

Moreover, the same arguments as in the proof of the second part of Proposition 7.43 show that $\|w_m^*\| \to 0$. It follows therefore that

$$\|x_m^* - \widehat{x}_m^*\| = \Big\| \sum_{n=1}^{k} (\lambda_n^m - \lambda_n) a_n^* + w_m^* \Big\| \le \varepsilon_m \sup_{1\le n\le k} \|a_n^*\| + \|w_m^*\| = o(1),$$

which implies together with the above estimates of $\operatorname{reg} F\big(\bar{x}, (\bar{z}, 0)\big)$ that

$$\operatorname{reg} F\big(\bar{x}, (\bar{z}, 0)\big) \le \big(\|\widehat{x}_m^*\| + o(1)\big)^{-1} + o(1) \le \big(\eta_m \|\eta_m^{-1} \widehat{x}_m^*\| + o(1)\big)^{-1} + o(1)$$
$$\le \Big[(1 - o(1)) \inf\Big\{\|x^*\| \,\Big|\, (x^*, \langle x^*, \bar{x}\rangle) \in L\Big\} + o(1)\Big]^{-1} + o(1).$$

Letting $m \to \infty$ therein, we arrive at

$$\operatorname{reg} F\big(\bar{x}, (\bar{z}, 0)\big) \le \sup\Big\{\|x^*\|^{-1} \,\Big|\, (x^*, \langle x^*, \bar{x}\rangle) \in L\Big\},$$

which yields (7.149) and thus completes the proof of the theorem. △

7.5 Value Functions in DC Semi-infinite Optimization

In this section, we continue investigating SIPs in general Banach (and partly in Asplund) spaces while considering now the minimization of *DC objectives* subject to infinite convex inequality constraints with arbitrary index sets. As mentioned earlier, the abbreviation "DC" stands for the *difference of convex* functions, which have been recognized as a convenient form for representing various remarkable classes of problems important in optimization and its applications. Our main attention is paid here to the study of subdifferential properties of (nonconvex) *marginal/value functions* in parametric versions of such SIPs. Based on these developments, we present applications to sensitivity analysis and necessary optimality conditions in DC SIPs considered in both nonparametric and parametric settings as well as to bilevel semi-infinite programs with fully convex data in Banach and Asplund spaces.

7.5.1 Optimality Conditions for DC Semi-infinite Programs

Consider first *nonparametric* SIPs with DC objectives and infinite convex constraints and obtain necessary optimality conditions for them (necessary and sufficient for fully convex problems) under weakest qualification conditions. These results of their own interest are instrumental to derive in what follows subdifferential formulas for value functions in parametric versions of such SIPs with subsequent applications to optimality conditions and Lipschitzian stability under perturbations. In this subsection, we study the problem

$$\begin{cases} \text{minimize} \;\; \vartheta(x) - \theta(x) \;\; \text{subject to} \\ \vartheta_t(x) \le 0, \quad t \in T, \quad \text{and} \quad x \in \Theta, \end{cases} \tag{7.157}$$

where T is an arbitrary index set, where $\Theta \subset X$ is a closed and convex subset of a Banach space X, and where the functions $\vartheta, \theta, \vartheta_t \colon X \to \overline{\mathbb{R}}$ are l.s.c. and convex. Being oriented toward minimization, we impose by convention that $\infty - \infty := \infty$ along with the standard operations involving ∞ and $-\infty$. The set of feasible solutions to (7.157) is denoted by

$$\Xi := \Theta \cap \big\{x \in X \big|\, \vartheta_t(x) \le 0 \;\text{ for all }\; t \in T\big\}. \tag{7.158}$$

Using the infinite product notation $\mathbb{R}^T$, $\mathbb{R}^{(T)}$, and $\mathbb{R}^{(T)}_+$ from Subsection 7.1.1, define $\operatorname{supp}\lambda := \{t \in T \mid \lambda_t \ne 0\}$ for any $\lambda \in \mathbb{R}^{(T)}$, and observe that

$$\lambda u := \sum_{t \in T} \lambda_t u_t = \sum_{t \in \operatorname{supp}\lambda} \lambda_t u_t \;\; \text{whenever} \;\; u \in \mathbb{R}^T.$$

Recall next definition (7.30) of the Fenchel conjugate, and introduce the following *dual-space* qualification condition, which plays a crucial role in deriving necessary optimality conditions of the KKT type for (7.157).

Definition 7.48 (Closedness Qualification Condition). *We say that the triple* $(\vartheta, \vartheta_t, \Theta)$ *in problem* (7.157) *satisfies the* CLOSEDNESS QUALIFICATION CONDITION (CQC) *if the set*

$$\operatorname{epi}\vartheta^* + \operatorname{cone}\left\{\bigcup_{t\in T}\operatorname{epi}\vartheta_t^*\right\} + \operatorname{epi}\delta^*(\cdot;\Theta)$$

is weak[*] *closed in the product space* $X^* \times \mathbb{R}$.

Note that the introduced CQC is not a "constraint qualification" since it involves not only constraint but also cost functions, namely, the *plus* part ϑ of the cost in (7.157). The closest constraint qualification to CQC is the following one, where the cost term $\operatorname{epi}\vartheta^*$ in Definition 7.48 is omitted: the set

$$\operatorname{cone}\left\{\bigcup_{t\in T}\operatorname{epi}\vartheta_t^*\right\} + \operatorname{epi}\delta^*(\cdot;\Theta) \ \text{ is weak}^* \text{ closed} \tag{7.159}$$

in $X^* \times \mathbb{R}$. This condition known as the convex *Farkas-Minkowski constraint qualification* (convex FMCQ) reduces to the Farkas-Minkowski property (7.49) for linear infinite systems of type (7.48). The reader can check that FMCQ (7.159) implies CQC in the following two cases: either ϑ is continuous at some feasible point $x \in \Xi$ in (7.158), or the convex conic hull $\operatorname{cone}(\operatorname{dom}\vartheta - \Xi)$ is a closed subspace of X. It has been well recognized in semi-infinite programming that *dual* qualification conditions of the CQC and Farkas-Minkowski type for infinite convex systems strictly improve *primal* ones of the Slater type; see Exercise 7.98 and the corresponding commentaries in Section 7.7.

To proceed, we recall some needed results of convex analysis summarized in the following two lemmas. The first one contains relationship between epigraphical duality and subdifferential calculus.

Lemma 7.49 (Epigraphical and Subdifferential Sum Rules). *Let the functions* $\varphi_1, \varphi_2 : X \to \overline{\mathbb{R}}$ *be l.s.c. and convex, and let* $\operatorname{dom}\varphi_1 \cap \operatorname{dom}\varphi_2 \neq \emptyset$. *Then the following conditions are equivalent:*

(i) *The set* $\operatorname{epi}\varphi_1^* + \operatorname{epi}\varphi_2^*$ *is weak*[*] *closed in* $X^* \times \mathbb{R}$.

(ii) *The conjugate epigraphical rule holds*

$$\operatorname{epi}(\varphi_1 + \varphi_2)^* = \operatorname{epi}\varphi_1^* + \operatorname{epi}\varphi_2^*.$$

Furthermore, we have the subdifferential sum rule

$$\partial(\varphi_1 + \varphi_2)(\bar{x}) = \partial\varphi_1(\bar{x}) + \partial\varphi_2(\bar{x})$$

provided that the aforementioned equivalent conditions are satisfied.

The next result presents an appropriate extension of the Farkas lemma to the case of epigraphical convex systems.

Lemma 7.50 (Generalized Farkas Lemma for Epigraphical Systems). *Given $\alpha \in \mathbb{R}$, the following conditions are equivalent:*

(i) $\vartheta(x) \geq \alpha$ *for all* $x \in \Xi$*;*

(ii) $\big(0, -\alpha\big) \in \text{cl}^*\Big(\text{epi}\,\vartheta^* + \text{cone}\Big[\bigcup_{t\in T} \text{epi}\,\vartheta_t^*\Big] + \text{epi}\,\delta^*(\cdot;\Theta)\Big)$.

Now we are ready to establish necessary optimality conditions for the DC program under consideration in (7.157). Given $\bar{x} \in \Xi \cap \text{dom}\,\theta$, define the set of *active constraint multipliers* by

$$A(\bar{x}) := \big\{\lambda \in \mathbb{R}_+^{(T)} \big|\ \lambda_t \vartheta_t(\bar{x}) = 0 \ \text{ for all } \ t \in \text{supp}\,\lambda\big\}. \tag{7.160}$$

Theorem 7.51 (Necessary Optimality Conditions for DC Semi-infinite Programs). *Let $\bar{x} \in \Xi \cap \text{dom}\,\vartheta$ be a local minimizer for problem* (7.157) *satisfying the CQC requirement. Then we have the inclusion*

$$\partial\theta(\bar{x}) \subset \partial\vartheta(\bar{x}) + \bigcup_{\lambda\in A(\bar{x})} \Big[\sum_{t\in\text{supp}\,\lambda} \lambda_t \partial\vartheta_t(\bar{x})\Big] + N(\bar{x};\Theta). \tag{7.161}$$

Proof. There are two possible cases regarding $\bar{x} \in \Xi \cap \text{dom}\,\vartheta$: either $\bar{x} \notin \text{dom}\,\theta$ or $\bar{x} \in \text{dom}\,\theta$. In the first case, we have $\partial\theta(\bar{x}) = \emptyset$, and hence (7.161) holds automatically. Considering the remaining case of $\bar{x} \in \text{dom}\,\theta$, find by the subdifferential definition of convex analysis such $x^* \in X^*$ that

$$\theta(x) - \theta(\bar{x}) \geq \langle x^*, x - \bar{x}\rangle \ \text{ for all } \ x \in X.$$

This implies that the reference local minimizer $\bar{x}$ for (7.157) is also a local minimizer for the following *convex* SIP:

$$\begin{cases} \text{minimize } \ \widetilde{\vartheta}(x) := \vartheta(x) - \langle x^*, x - \bar{x}\rangle - \theta(\bar{x}) \\ \text{subject to } \ \vartheta_t(x) \leq 0, \quad t \in T, \quad \text{and } \ x \in \Theta. \end{cases} \tag{7.162}$$

Since (7.162) is convex, its local minimizer $\bar{x}$ is its global solution, i.e.,

$$\widetilde{\vartheta}(\bar{x}) \leq \widetilde{\vartheta}(x) \ \text{ for all } \ x \in \Xi.$$

Then Lemma 7.50 tells us that the latter is equivalent to the inclusion

$$\big(0, -\widetilde{\vartheta}(\bar{x})\big) \in \text{cl}^*\Big(\text{epi}\,\widetilde{\vartheta}^* + \text{cone}\Big[\bigcup_{t\in T} \text{epi}\,\vartheta_t^*\Big] + \text{epi}\,\delta^*(\cdot;\Theta)\Big).$$

Observing from the structure of $\widetilde{\vartheta}$ in (7.162) that $\text{epi}\,\widetilde{\vartheta}^* = (-x^*, \theta(\bar{x}) - \langle x^*, \bar{x}\rangle) + \text{epi}\,\vartheta^*$, we get therefore the relationship

$$\begin{aligned}&\big(0,-\widetilde{\vartheta}(\bar{x})\big)\in\big(-x^*,\theta(\bar{x})-\langle x^*,\bar{x}\rangle\big)\\&+\mathrm{cl}^*\Big(\mathrm{epi}\,\vartheta^*+\mathrm{cone}\Big[\bigcup_{t\in T}\mathrm{epi}\,\vartheta_t^*\Big]+\mathrm{epi}\,\delta^*(\cdot;\Theta)\Big).\end{aligned}\tag{7.163}$$

Furthermore, the assumed CQC ensures that (7.163) is equivalent to

$$\begin{aligned}&\big(x^*,-\widetilde{\vartheta}(\bar{x})-\theta(\bar{x})+\langle x^*,\bar{x}\rangle\big)\\&\in\Big(\mathrm{epi}\,\vartheta^*+\mathrm{cone}\Big[\bigcup_{t\in T}\mathrm{epi}\,\vartheta_t^*\Big]+\mathrm{epi}\,\delta^*(\cdot;\Theta)\Big).\end{aligned}\tag{7.164}$$

Now applying the useful representation

$$\mathrm{epi}\,\varphi^*=\bigcup_{\varepsilon\geq 0}\Big\{\big(x^*,\langle x^*,x\rangle+\varepsilon-\varphi(x)\big)\Big|\;x^*\in\partial_\varepsilon\varphi(x)\Big\},\tag{7.165}$$

which is valid for all $x\in\mathrm{dom}\,\varphi$, to the conjugate functions ϑ^*, ϑ_t^*, and $\delta^*(\cdot;\Theta)$ with taking into account the structure of the positive cone $\mathbb{R}_+^{(T)}$ in (7.3) and noting that $-\widetilde{\vartheta}(\bar{x})-\theta(\bar{x})+\langle x^*,\bar{x}\rangle=\langle x^*,\bar{x}\rangle-\vartheta(\bar{x})$, we find

$$\varepsilon,\varepsilon_t,\gamma\geq 0,\;\;u^*\in\partial_\varepsilon\vartheta(\bar{x}),\;\;\lambda\in\mathbb{R}_+^{(T)},\;\;u_t^*\in\partial_{\varepsilon_t}\vartheta_t(\bar{x}),\;\;\text{and}\;\;v^*\in\partial\delta_\gamma(\bar{x};\Theta)$$

satisfying the following two equalities:

$$\begin{cases}x^*=u^*+\displaystyle\sum_{t\in T}\lambda_t u_t^*+v^*,\\ \langle x^*,\bar{x}\rangle-\vartheta(\bar{x})=\langle u^*,\bar{x}\rangle+\varepsilon-\vartheta(\bar{x})+\displaystyle\sum_{t\in T}\lambda_t\Big[\langle u_t^*,\bar{x}\rangle+\varepsilon_t-\langle\vartheta_t^*,\bar{x}\rangle\Big]\\ +\langle v^*,\bar{x}\rangle+\gamma-\delta(\bar{x};\Theta).\end{cases}$$

Since $\bar{x}\in\Theta$, the first equality above allows us reducing the second one to

$$\varepsilon+\sum_{t\in T}\lambda_t\varepsilon_t-\sum_{t\in T}\lambda_t\vartheta_t(\bar{x})+\gamma=0.\tag{7.166}$$

The feasibility of $\bar{x}$ for problem (7.157) and the choice of $(\varepsilon,\lambda_t,\gamma)$ yield

$$\varepsilon\geq 0,\;\;\gamma\geq 0,\;\;\lambda_t\geq 0,\;\;\text{and}\;\;\lambda_t\vartheta_t(\bar{x})\leq 0\;\;\text{for all}\;\;t\in T,$$

and therefore we get from (7.166) that in fact $\varepsilon=0$, $\gamma=0$, $\lambda_t\vartheta_t(\bar{x})=0$, and $\lambda_t\varepsilon_t=0$ for all $t\in T$. Furthermore, the latter implies that $\varepsilon_t=0$ for all $t\in\mathrm{supp}\,\lambda$. Hence we obtain the inclusions

$$u^*\in\partial\vartheta(\bar{x}),\;\;\;u_t^*\in\partial\vartheta_t(\bar{x}),\;\;\;\text{and}\;\;\;v^*\in\partial\delta(\bar{x};\Theta)=N(\bar{x};\Theta),$$

which allow us to conclude from the above that

$$x^*\in\partial\vartheta(\bar{x})+\sum_{t\in\mathrm{supp}\,\lambda}\lambda_t\partial\vartheta_t(\bar{x})+N(\bar{x};\Theta)\;\;\text{with}\;\;\lambda_t\vartheta_t(\bar{x})=0\;\;\text{for}\;\;t\in\mathrm{supp}\,\lambda.$$

This justifies (7.161) and thus completes the proof of the theorem. △

Let us present two useful consequences of Theorem 7.51 concerning subdifferential/normal cone *calculus* for infinite convex systems.

Corollary 7.52 (Subdifferential Sum Rule Involving Convex Infinite Constraints). *Let* $\bar{x} \in \Xi$ *with* $\theta(\bar{x}) = 0$ *and* $\vartheta(\bar{x}) < \infty$*, and let* $(\vartheta, \vartheta_t, \Theta)$ *satisfy all the assumptions of Theorem* 7.51*. Then we have the equality*

$$\partial\big(\vartheta + \delta(\cdot;\Xi)\big)(\bar{x}) = \partial\vartheta(\bar{x}) + \bigcup_{\lambda \in A(\bar{x})}\Big[\sum_{t \in \operatorname{supp}\lambda} \partial\vartheta_t(\bar{x})\Big] + N(\bar{x};\Theta).$$

Proof. The inclusion "$\supset$" in the claimed sum rule can be derived directly from the definitions. To verify the opposite inclusion therein, pick an arbitrary subgradient $x^* \in \partial(\vartheta + \delta(\cdot;\Xi))(\bar{x})$ with $\bar{x} \in \Xi \cap \operatorname{dom}\vartheta$, and get

$$\vartheta(x) - \vartheta(\bar{x}) \ge \langle x^*, x - \bar{x}\rangle \quad \text{whenever} \quad x \in \Xi,$$

which means by the construction of Ξ in (7.158) that $\bar{x}$ is a (global) minimizer for the following *DC program* with infinite constraints:

$$\begin{cases} \text{minimize} \;\; \vartheta(x) - \widetilde{\theta}(x) \;\; \text{with} \;\; \widetilde{\theta}(x) := \langle x^*, x - \bar{x}\rangle + \vartheta(\bar{x}) \\ \text{subject to} \;\; \vartheta_t(x) \le 0 \;\; \text{for all} \;\; t \in T, \quad \text{and} \;\; x \in \Theta. \end{cases} \tag{7.167}$$

Applying Theorem 7.51 to problem (7.167) and taking into account the structure of the linear function $\widetilde{\theta}$ therein, we get from (7.161) that

$$\partial\widetilde{\theta}(\bar{x}) = \big\{x^*\big\} \subset \partial\vartheta(\bar{x}) + \bigcup_{\lambda \in A(\bar{x})}\Big[\sum_{t \in \operatorname{supp}\lambda} \partial\vartheta_t(\bar{x})\Big] + N(\bar{x};\Theta),$$

which justifies the claimed inclusion and thus completes the proof. △

The next corollary provides a calculation of the normal cone to the feasible constraint set Ξ in terms of its initial data of (7.12) and the set of active constraint multipliers (7.160).

Corollary 7.53 (Normal Cone Calculation for Convex Infinite Constraints). *Assume that* ϑ_t *and* Θ *satisfy the assumptions of Theorem* 7.51 *with CQC specified as FMCQ* (7.159)*. Then for any* $\bar{x} \in \Xi$*, we have*

$$N(\bar{x};\Xi) = \bigcup_{\lambda \in A(\bar{x})}\Big[\sum_{t \in \operatorname{supp}\lambda} \partial\vartheta_t(\bar{x})\Big] + N(\bar{x};\Theta).$$

Proof. Follows from Corollary 7.52 by letting $\vartheta(x) \equiv 0$ therein. △

The final result of this subsection concerns the *convex SIP*, which is a specification of (7.157) with $\theta \equiv 0$. We show that in this case the necessary conditions of Theorem 7.51 are also sufficient for (global) optimality.

Theorem 7.54 (Necessary and Sufficient Optimality Conditions for Convex SIPs). *Let $\bar{x} \in \Xi$ be a feasible solution to problem (7.157) with $\theta \equiv 0$ and $\vartheta(\bar{x}) < \infty$, and let the assumptions of Theorem 7.51 be satisfied. Then $\bar{x}$ is optimal to this problem if and only if there is $\lambda \in \mathbb{R}_+^{(T)}$ such that the following generalized KKT condition holds:*

$$0 \in \partial\vartheta(\bar{x}) + \bigcup_{\lambda \in A(\bar{x})} \Big[\sum_{t \in \operatorname{supp}\lambda} \partial\vartheta_t(\bar{x}) \Big] + N(\bar{x}; \Theta). \tag{7.168}$$

Proof. The necessary of (7.168) for optimality in this problem follows immediately from Theorem 7.51 with $\theta(x) \equiv 0$. To justify the sufficiency part, suppose that (7.168) holds with some $\lambda \in A(\bar{x})$; the latter implies, in particular, that $\partial\vartheta_t(\bar{x}) \neq \emptyset$ whenever $t \in \operatorname{supp}\lambda$. Then we find $x^* \in X^*$ satisfying the inclusions $-x^* \in N(\bar{x}; \Theta)$ and

$$x^* \in \partial\vartheta(\bar{x}) + \sum_{t \in \operatorname{supp}\lambda} \partial\vartheta_t(\bar{x}) \subset \partial\Big(\vartheta + \sum_{t \in T} \lambda_t \vartheta_t\Big)(\bar{x}).$$

This tells by the construction of convex subgradients that

$$\vartheta(x) + \sum_{t \in T} \lambda_t \vartheta_t(x) \geq \vartheta(\bar{x}) + \sum_{t \in T} \lambda_t \vartheta_t(\bar{x}) + \langle x^*, x - \bar{x} \rangle \geq 0 \tag{7.169}$$

for all $x \in X$. Since $\lambda_t \vartheta_t(\bar{x}) = 0$ for all $t \in T$ by $\lambda \in A(\bar{x})$ in (7.160) and since $-x^* \in N(\bar{x}; \Theta)$, we get from (7.169) and the normal cone structure that

$$\vartheta(x) + \sum_{t \in T} \lambda_t \vartheta_t(x) - \vartheta(\bar{x}) \geq \langle x^*, x - \bar{x} \rangle \geq 0 \quad \text{for all} \quad x \in \Theta,$$

which yields by (7.158) and (7.160) the inequality

$$\vartheta(x) \geq \vartheta(x) + \sum_{t \in T} \lambda_t \vartheta_t(x) \geq \vartheta(\bar{x}) \quad \text{whenever} \quad x \in \Xi$$

and thus verifies the claimed global optimality of $\bar{x}$. △

7.5.2 Regular Subgradients of Value Functions for DC SIPs

Let us now consider the *parametric* version of the DC semi-infinite program (7.157) formalized, with a bit different notation, as

$$\text{minimize}_y \ \varphi(x, y) - \psi(x, y) \quad \text{subject to } y \in F(x) \cap G(x), \tag{7.170}$$

where the moving (parameterized by x) constraint sets are given by

$$F(x) := \big\{ y \in Y \big| \ (x, y) \in \Omega \big\}, \tag{7.171}$$

$$G(x) := \big\{ y \in Y \big|\ \varphi_t(x, y) \le 0,\ t \in T \big\}. \tag{7.172}$$

In what follows, we assume, unless otherwise stated, that the spaces X and Y are Banach, that T is an arbitrary index set, that the functions $\varphi, \psi, \varphi_t\colon X \times Y \to \overline{\mathbb{R}}$ are l.s.c. and convex, and that the set Ω is closed and convex.

The main object of our study in the rest of this section is the (optimal) *value function* in (7.170) defined by

$$\mu(x) := \inf \big\{ \varphi(x, y) - \psi(x, y) \big|\ y \in F(x) \cap G(x) \big\}, \tag{7.173}$$

which is nonconvex unless $\psi \equiv 0$. The value function (7.173) belongs to the general class of marginal functions whose subdifferential properties have been studied in Section 4.1; see also the corresponding commentaries in Section 4.6. However, the results obtained therein are expressed in terms of the coderivative of the constraint mapping in (7.170), while the major goal of our study here is to derive subdifferential results for (7.173) expressed entirely via the *initial data* of (7.170) with taking into account the infinite inequality constraint nature of (7.172) and the DC structure of the cost function in (7.170).

In this subsection, we concentrate on evaluating the *regular subdifferential* of (7.173), which is defined in Banach spaces exactly as in finite dimensions (1.33). The results obtained are of their own interest, while they also can be considered, together with similar calculations for the ε-enlargements (1.34), as approximating tools for evaluating the limiting (both basic and singular) subdifferentials of the value function, which are the most valuable applications to DC semi-infinite optimization and Lipschitzian stability of (7.170). The necessary optimality conditions for the nonparametric DC version (7.157) obtained in Subsection 7.5.1 play a significant role in our subdifferential device. For brevity, we confine ourselves here to considering only regular subgradients of (7.173) while leaving the ε-case as an exercise for the reader.

In the next theorem and further results below, we use the notation

$$M(x) := \big\{ y \in F(x) \cap G(x) \big|\ \mu(x) = \varphi(x, y) - \psi(x, y) \big\}, \tag{7.174}$$

$$\Gamma := \Omega \cap \big\{ (x, y) \in X \times Y \big|\ \varphi_t(x, y) \le 0 \ \text{ for all }\ t \in T \big\}, \tag{7.175}$$

$$\begin{aligned} \Lambda(\bar{x}, \bar{y}, y^*) := \Big\{ \lambda \in \mathbb{R}^{(T)}_+ \Big|\ & y^* \in \partial_y \varphi(\bar{x}, \bar{y}) + \sum_{t \in \operatorname{supp} \lambda} \lambda_t \partial_y \varphi_t(\bar{x}, \bar{y}) \\ & + N_y\big((\bar{x}, \bar{y}); \Omega\big),\ \lambda_t \varphi_t(\bar{x}, \bar{y}) = 0 \ \text{ as }\ t \in \operatorname{supp} \lambda \Big\}, \end{aligned} \tag{7.176}$$

where $N_y((\bar{x}, \bar{y}); \Omega)$ stands for the subdifferential of the indicator function $y \mapsto \delta((\bar{x}, y); \Omega)$ at $\bar{y}$; the notation $N_x((\bar{x}, \bar{y}); \Omega)$ below is similar.

Theorem 7.55 (Upper Estimate for Regular Subgradients of Value Functions in DC SIPs). *Let* dom $M \ne \emptyset$*, and let CQC from Definition* 7.48 *be satisfied for the* *triple* $(\varphi, \varphi_t, \Omega)$ *in* (7.170). *Then, given any* $(\bar{x}, \bar{y}) \in \operatorname{gph} M \cap \operatorname{dom} \partial\psi$ *and* $\gamma > 0$, *we have the inclusion*

$$\widehat{\partial}\mu(\bar{x}) \subset \bigcap_{(x^*,y^*)\in\partial\psi(\bar{x},\bar{y})} \Big\{\partial_x\varphi(\bar{x}, \bar{y}) - x^* + \bigcup_{\lambda\in\Lambda(\bar{x},\bar{y},y^*)} \Big[\sum_{t\in\operatorname{supp}\lambda} \lambda_t\partial_x\varphi_t(\bar{x}, \bar{y})\Big]\Big\}$$

$$+N_x\big((\bar{x}, \bar{y}); \Omega\big) + \gamma\mathbb{B}^*.$$

Proof. Fix $(\bar{x}, \bar{y}) \in \operatorname{gph} M \cap \operatorname{dom} \partial\psi$, $u^* \in \widehat{\partial}\mu(\bar{x})$, and $(x^*, y^*) \in \partial\psi(\bar{x}, \bar{y})$. Picking an arbitrary positive number γ and using the definition of regular subgradients, find $\eta > 0$ such that

$$\mu(x) - \mu(\bar{x}) - \langle u^*, x - \bar{x}\rangle + \gamma\|x - \bar{x}\| \geq 0 \ \text{ if } \ x \in \bar{x} + \eta\mathbb{B}. \tag{7.177}$$

Since $\mu(\bar{x}) = \varphi(\bar{x}, \bar{y}) - \psi(\bar{x}, \bar{y})$ by $\bar{y} \in M(\bar{x})$ and since $\mu(x) \leq \varphi(x, y) - \psi(x, y)$ for all $(x, y) \in \Gamma$, we get from (7.177) and $(x^*, y^*) \in \partial\psi(\bar{x}, \bar{y})$ that

$$\begin{aligned} 0 &\leq \varphi(x, y) - \varphi(\bar{x}, \bar{y}) - \psi(x, y) + \psi(\bar{x}, \bar{y}) - \langle u^*, x - \bar{x}\rangle + \gamma\|x - \bar{x}\| \\ &\leq \varphi(x, y) - \varphi(\bar{x}, \bar{y}) - \langle u^* + x^*, x - \bar{x}\rangle - \langle y^*, y - \bar{y}\rangle + \gamma\|x - \bar{x}\| \end{aligned}$$

for $(x, y) \in \Omega \cap [(\bar{x} + \eta\mathbb{B}) \times Y]$ with $\varphi_t(x, y) \leq 0$, $t \in T$. Consider the function

$$\vartheta(x, y) := \varphi(x, y) - \varphi(\bar{x}, \bar{y}) - \langle u^* + x^*, x - \bar{x}\rangle - \langle y^*, y - \bar{y}\rangle + \gamma\|x - \bar{x}\|,$$

which is l.s.c. and convex on $X \times Y$. It follows from (7.177) and the construction of ϑ that $(\bar{x}, \bar{y})$ is a solution to the following *nonparametric convex* SIP:

$$\begin{cases} \text{minimize } \ \vartheta(x, y) \ \text{ with respect to both } (x, y) \text{ subject to} \\ \varphi_t(x, y) \leq 0 \ \text{ as } \ t \in T, \quad (x, y) \in \Omega \cap \big[(\bar{x} + \eta\mathbb{B}) \times Y\big]. \end{cases} \tag{7.178}$$

The technical Lemma 7.56, which is presented for convenience after the proof of the theorem, tells us the CQC requirement on $(\varphi, \varphi_t, \Omega)$ imposed in this theorem yields the validity of CQC for (7.178). Applying now the optimality conditions from Theorem 7.54 to (7.178) gives us $\lambda \in \mathbb{R}_+^{(T)}$ such that

$$\begin{aligned} &0 \in \partial\vartheta(\bar{x}, \bar{y}) + \sum_{t\in\operatorname{supp}\lambda} \lambda_t\partial\varphi_t(\bar{x}, \bar{y}) + N\big((\bar{x}, \bar{y}); \Omega \cap [(\bar{x} + \eta\mathbb{B}) \times Y]\big) \\ &\text{with } \ \lambda_t\varphi_t(\bar{x}, \bar{y}) = 0 \ \text{ for all } \ t \in \operatorname{supp}\lambda. \end{aligned}$$

Since $(\bar{x}, \bar{y}) \in \imath[(\bar{x} + \eta\mathbb{B}) \times Y]$, it follows from the classical subdifferential rule of convex analysis and the construction of ϑ that

$$\partial\vartheta(\bar{x}, \bar{y}) = \partial\varphi(\bar{x}, \bar{y}) + (-u^* - x^*, -y^*) + (\gamma\mathbb{B}^*) \times \{0\}.$$

Thus we get by (i)$\Rightarrow$(iii) in Lemma 7.49 applied to the indicator functions $\delta((\bar{x}, \bar{y}); \Omega)$ and $\delta((\bar{x}, \bar{y}); (\bar{x} + \eta\mathbb{B}) \times Y)$ that

$$N\big((\bar{x},\bar{y});\Omega\cap[(\bar{x}+\eta\mathbb{B})\times Y]\big)=N\big((\bar{x},\bar{y});\Omega\big).$$

Substituting this into the above optimality condition for (7.178) with taking into account the well-known relationships

$$\partial\varphi(\bar{x},\bar{y})\subset\partial_x\varphi(\bar{x},\bar{y})\times\partial_y\varphi(\bar{x},\bar{y})\ \text{ and }\ \partial\varphi_t(\bar{x},\bar{y})\subset\partial_x\varphi_t(\bar{x},\bar{y})\times\partial_y\varphi_t(\bar{x},\bar{y})$$

ensures the fulfillment of the two inclusions

$$\begin{aligned}
u^*&\in\partial_x\varphi(\bar{x},\bar{y})-x^*+\sum_{t\in\operatorname{supp}\lambda}\lambda_t\partial_x\varphi_t(\bar{x},\bar{y})+N_x\big((\bar{x},\bar{y});\Omega\big)+\gamma\mathbb{B}^*,\\
y^*&\in\partial_y\varphi(\bar{x},\bar{y})+\sum_{t\in\operatorname{supp}\lambda}\lambda_t\partial_y\varphi_t(\bar{x},\bar{y})+N_y\big((\bar{x},\bar{y});\Omega\big)
\end{aligned}$$

with $\lambda_t\varphi_t(\bar{x},\bar{y})=0$, $t\in\operatorname{supp}\lambda$. This verifies by (7.176) the claimed estimate of $\widehat{\partial}\mu(\bar{x})$ by the construction in (7.176) and Lemma 7.56 justified below. △

Lemma 7.56 (Relationships Between Parametric and Nonparametric CQC). *The validity of CQC for* $(\varphi,\varphi_t,\Omega)$ *imposed in Theorem 7.55 yields the fulfillment of this condition for the nonparametric problem* (7.178).

Proof. In the notation of Theorem 7.55, take $(\bar{x},\bar{y})\in\operatorname{gph}M\cap\operatorname{dom}\partial\psi$ with $(\bar{x},\bar{y})\in\operatorname{dom}\varphi\cap\Gamma$, and define the convex and continuous function

$$\xi(x,y):=-\varphi(\bar{x},\bar{y})-\langle u^*+x^*,x-\bar{x}\rangle-\langle y^*,y-\bar{y}\rangle+\gamma\|x-\bar{x}\|$$

on $X\times Y$ that gives us the representation $\vartheta=\varphi+\xi$. Substituting the latter into the assumed CQC for $(\varphi,\varphi_t,\Omega)$ and using the epigraphical rule from Lemma 7.49 with taking into account that the continuity of $\delta(\cdot;(\bar{x}+\eta\mathbb{B}^*)\times Y)$ at the interior point $(\bar{x},\bar{y})$, we conclude that the corresponding set in the CQC property for (7.178) reduces to

$$\operatorname{epi}\varphi^*+\operatorname{cone}\Big[\bigcup_{t\in T}\operatorname{epi}\varphi_t^*\Big]+\operatorname{epi}\delta^*(\cdot;\Omega)+\operatorname{epi}\big[\xi+\delta\big(\cdot;(\bar{x}+\eta\mathbb{B})\times Y\big)\big]^*.$$

On the other hand, by Lemma 7.49, the CQC requirement for $(\varphi,\varphi_t,\Omega)$ yields

$$\operatorname{epi}\big(\varphi+\delta(\cdot;\Gamma)\big)^*=\operatorname{epi}\varphi^*+\operatorname{cone}\Big[\bigcup_{t\in T}\operatorname{epi}\varphi_t^*\Big]+\operatorname{epi}\delta^*(\cdot;\Omega).$$

Substituting this equality into the aforementioned CQC set for $(\varphi,\varphi_t,\Omega)$, we express the latter set as follows:

$$\operatorname{epi}\big(\varphi+\delta(\cdot;\Gamma)\big)^*+\operatorname{epi}\big[\xi+\delta\big(\cdot;(\bar{x}+\eta\mathbb{B})\times Y\big)\big]^*,$$

which in turn reduces to the form

$$\operatorname{epi}\big[\varphi+\delta(\cdot;\Gamma)+\xi+\delta(\cdot;(\bar{x}+\eta\mathbb{B})\times Y)\big]^*$$

by using Lemma 7.49 and the continuity of the function $\xi + \delta\big(\cdot; (\bar{x} + \eta\mathbb{B}) \times Y\big)$ at $(\bar{x}, \bar{y}) \in \operatorname{dom}\big(\varphi + \delta(\cdot; \Gamma)\big)$. The latter set is weak* closed in $X^* \times Y^* \times \mathbb{R}$ as the epigraph of the conjugate function to $\varphi + \delta(\cdot; \Gamma) + \xi + \delta(\cdot; (\bar{x} + \eta\mathbb{B}) \times Y)$. Thus we are done with the proof of this lemma. $\triangle$

As a consequence of Theorem 7.55, we derive necessary optimality conditions for the parametric DC program (7.170) that are *upper subdifferential* conditions according to the terminology of Section 6.1. Indeed, they involve *all* the upper subgradients of the concave function $-\psi$ at the reference point, which reduce to subgradients of the convex function ψ in the cost of (7.170).

Corollary 7.57 (Upper Subdifferential Conditions for Parametric DC SIPs). *Given a parameter value* $\bar{x} \in \operatorname{dom} M$ *in* (7.174)*, let* $\bar{y}$ *be a* (*global*) *optimal solution to the parametric DC program*

$$\textit{minimize}\ \ \varphi(\bar{x}, y) - \psi(\bar{x}, y)\ \ \textit{subject to}\ \ y \in F(\bar{x}) \cap G(\bar{x}) \tag{7.179}$$

with F *and* G *from* (7.171) *and* (7.172)*, respectively, under the standing assumptions made. Suppose in addition that* $\widehat{\partial}\mu(\bar{x}) \neq \emptyset$ *for the value function* (7.173) *under the CQC property for* $(\varphi, \varphi_t, \Omega)$*. Then for each* $(x^*, y^*) \in \partial\psi(\bar{x}, \bar{y})$ *and* $\gamma > 0$*, there are* $u^* \in X^*$ *and* $\lambda \in \mathbb{R}^{(T)}_+$ *from* (7.3) *such that*

$$\begin{aligned}
&u^* + x^* \in \partial_x\varphi(\bar{x}, \bar{y}) + \sum_{t \in \operatorname{supp}\lambda} \lambda_t \partial_x \varphi_t(\bar{x}, \bar{y}) + N_x\big((\bar{x}, \bar{y}); \Omega\big) + \gamma\mathbb{B}^*,\\
&y^* \in \partial_y\varphi(\bar{x}, \bar{y}) + \sum_{t \in \operatorname{supp}\lambda} \lambda_t \partial_y \varphi_t(\bar{x}, \bar{y}) + N_y\big((\bar{x}, \bar{y}); \Omega\big),\\
&\lambda_t \varphi_t(\bar{x}, \bar{y}) = 0\ \ \textit{for all}\ \ t \in \operatorname{supp}\lambda.
\end{aligned}$$

Proof. Follows directly from the upper estimate in Theorem 7.55 due to $\widehat{\partial}\mu(\bar{x}) \neq \emptyset$ and the construction of the KKT multiplier set in (7.176). $\triangle$

The most restrictive and not easily verifiable assumption in Corollary 7.57 is that of $\widehat{\partial}\mu(\bar{x}) \neq \emptyset$. In the next subsection, we derive improved necessary optimality conditions for (7.170) while replacing the restrictive requirement on $\widehat{\partial}\mu(\bar{x}) \neq \emptyset$ by more natural and verifiable assumptions in the case of Asplund spaces. This comes as a consequence of upper estimates for basic and singular subgradients of the DC value function (7.173) in more general settings.

7.5.3 Limiting Subgradients of Value Functions for DC SIPs

We begin with the constructive evaluation of the basic subdifferential (1.24) of the value function (7.173) and obtain two independent results in this direction under different assumptions and with completely different proofs. Recall from Section 1.5 (see also [522] for more details) that the basic subdifferential of any $\varphi\colon X \to \overline{\mathbb{R}}$ on a Banach space X at $\bar{x} \in \operatorname{dom}\varphi$ is defined by

$$\partial\varphi(\bar{x}) := \mathop{\mathrm{Lim\,sup}}_{x \xrightarrow{\varphi} \bar{x},\, \varepsilon\downarrow 0} \widehat{\partial}_\varepsilon \varphi(x) \tag{7.180}$$

via the sequential outer limit of the ε-subdifferential mappings $\widehat{\partial}_\varepsilon\varphi\colon X \rightrightarrows X^*$ of φ at points nearby. If φ is l.s.c. around $\bar{x}$ and the space X is Asplund, then $\varepsilon > 0$ can be equivalently omitted in (7.180); see [522, Theorem 2.34].

For the first result, we need the following condition on the *minus* term ψ in (7.173), which allows us to derive a tight upper estimate of $\partial\mu(\bar{x})$.

Definition 7.58 (Inner Subdifferential Stability). *We say that a convex function* $\psi\colon X \to \overline{\mathbb{R}}$ *is* INNER SUBDIFFERENTIALLY STABLE *at* $\bar{x} \in \mathrm{dom}\,\psi$ *if*

$$\mathop{\mathrm{Lim\,inf}}_{x \xrightarrow{\mathrm{dom}\,\psi} \bar{x}} \partial\psi(x) \neq \emptyset, \tag{7.181}$$

where Lim inf *stands for the Painlevé-Kuratowski inner limit* (1.20) *with the usage of the weak* sequential convergence on* X^*.

Note that (7.181) reduces to a singleton in the case of general Banach spaces if ψ is Gâteaux differentiable on a neighborhood of $\bar{x}$ and its Gâteaux derivative operator $d\psi\colon X \to X^*$ is continuous with respect to the weak* topology of X^*. The next proposition relaxes the smoothness assumption *around* $\bar{x}$ provided that the closed unit ball $\mathbb{B}^*$ in X^* is *weak* sequentially compact*. This latter property holds for general classes of Banach spaces X, in particular; for those admitting an equivalent norm Gâteaux differentiable at nonzero points (Gâteaux smooth), for weak Asplund spaces that includes every Asplund space and every weakly compactly generated space, every reflexive and every separable space, etc.; see, e.g., [255] for more details.

Proposition 7.59 (Sufficient Conditions for Inner Subdifferential Stability). *Let* X *be a Banach space such that the closed unit ball* $\mathbb{B}^*$ *is weak* sequentially compact in* X^**, and let* ψ *be convex, continuous, and Gâteaux differentiable at* $\bar{x} \in \mathrm{int}(\mathrm{dom}\,\psi)$. *Then* ψ *is inner subdifferentially stable at* $\bar{x}$.

Proof. Take any sequence $x_k \to \bar{x}$ as $k \to \infty$ and suppose that it entirely belongs to some neighborhood $U \subset \mathrm{dom}\,\psi$ of $\bar{x}$. It follows from the continuity of the convex function ψ at $\bar{x}$ that it is actually Lipschitz continuous around $\bar{x}$, and hence its subdifferential mapping $\partial\psi(\cdot)$ is bounded in X^* by the Lipschitz constant of ψ; see Exercises 1.69(i) and 7.102. This implies by using the weak* sequential compactness of the dual ball B^* that every subset of the set

$$V^* := \left\{ x^* \in X^* \,\middle|\, \exists\, x \in U \;\text{ with }\; x^* \in \partial\psi(x) \right\}$$

contains a subsequence converging in the weak* topology of X^*. Then picking any sequence of subgradients $x_k^* \in \partial\psi(x_k)$, we suppose without loss of generality that there is $x^* \in X^*$ such that $x_k^* \xrightarrow{w^*} x^*$ as $k \to \infty$. It follows from convex subdifferential definition (1.35) with $\varepsilon = 0$ that $x^* \in \partial\psi(\bar{x})$. Since ψ is continuous and Gâteaux differentiable at $\bar{x}$, we have from standard convex analysis that

$\partial\psi(\bar{x}) = \{d\psi(\bar{x})\}$, and therefore $x_k^* \xrightarrow{w^*} d\psi(\bar{x})$ as $k \to \infty$. This clearly verifies the inner subdifferential stability (7.181) of ψ at $\bar{x}$. $\triangle$

It is not hard to give various examples of functions, which are not Gâteaux differentiable at the reference point while being inner subdifferentially stable at it. Such functions can be constructed by the following scheme. Take a closed and convex subset Ω of a Gâteaux smooth space X, a point $\bar{x} \in \operatorname{bd}\Omega$, and a function $\theta(x)$ that is convex, continuous, and Gâteaux differentiable on an open set containing $\bar{x}$. Then define $\psi\colon X \to \overline{\mathbb{R}}$ as $\psi(x) := \theta(x)$ on Ω and as $\psi(x) := \emptyset$ otherwise. It follows from Definition 7.58 and Proposition 7.59 that $\operatorname{Lim\,inf}\psi$ in (7.181) reduces to $\{d\theta(\bar{x})\}$, and thus we have the inner subdifferential stability of ψ at $\bar{x}$. Observe that

$$\partial\psi(\bar{x}) = d\theta(\bar{x}) + N(\bar{x}; \Omega)$$

by the subdifferential sum rule from Lemma 7.49 due the assumed continuity of θ. Taking into account our convention on $\infty - \infty = \infty$, we get a boundary domain point $\bar{x} \in \operatorname{bd}(\operatorname{dom}\psi)$, which is a local minimizer for the DC function $\varphi - \psi$ provided that $\operatorname{dom}\varphi \subset \operatorname{dom}\psi$.

Now we are ready to establish the aforementioned tight upper estimate of basic subgradients of the value function (7.173) under the inner subdifferential stability of ψ in (7.170). This result requires also the inner semicontinuity property (1.20) of the argminimum mapping $M(\cdot)$ from (7.174).

Theorem 7.60 (Basic Subgradients of DC Value Functions Under Inner Subdifferential Stability). *Given $(\bar{x}, \bar{y}) \in \operatorname{gph} M$ in (7.170), suppose that $M(\cdot)$ is inner semicontinuous, that ψ is inner subdifferentially stable, and that CQC holds for $(\varphi, \varphi_t, \Omega)$ at this point. Then for any fixed $(x^*, y^*) \in \operatorname*{Lim\,inf}_{(x,y)\xrightarrow{\operatorname{dom}\psi}(\bar{x},\bar{y})} \partial\psi(x, y)$, we have the inclusion*

$$\partial\mu(\bar{x}) \subset \partial_x\varphi(\bar{x}, \bar{y}) - x^* + \bigcup_{\lambda\in\Lambda(\bar{x},\bar{y},y^*)} \Big[\sum_{t\in\operatorname{supp}\lambda} \lambda_t\partial_x\varphi_t(\bar{x}, \bar{y})\Big] + N_x\big((\bar{x}, \bar{y}); \Omega\big)$$

with the set of KKT multipliers $\Lambda(\bar{x}, \bar{y}, y^*)$ defined in (7.176).

Proof. Fix the pair (x^*, y^*) from the theorem formulation, and pick an arbitrary subgradient $u^* \in \partial\mu(\bar{x})$. Then definition (7.180) gives us sequences $\varepsilon_k \downarrow 0$, $x_k \xrightarrow{\mu} \bar{x}$, and $u_k^* \in \widehat{\partial}_{\varepsilon_k}\mu(x_k)$ with $u_k^* \xrightarrow{w^*} u^*$ as $k \to \infty$. Fixing $k \in \mathbb{N}$ and using ε_k-subgradient construction (1.34) for u_k^*, we find $\eta_k > 0$ such that

$$\langle u_k^*, x - x_k\rangle \le \mu(x) - \mu(x_k) + 2\varepsilon_k\|x - x_k\| \text{ if } x \in x_k + \eta_k\mathbb{B}. \tag{7.182}$$

The inner semicontinuity of $M(\cdot)$ at $(\bar{x}, \bar{y})$ allows us to find a sequence of $y_k \in M(x_k)$ that contains a subsequence converging to $\bar{y}$; we suppose that $y_k \to \bar{y}$ for all $k \to \infty$. By the choice of (x^*, y^*), there is a sequence of subgradients $(x_k^*, y_k^*) \in \partial\psi(x_k, y_k)$ with $(x_k^*, y_k^*) \xrightarrow{w^*} (x^*, y^*)$ as $k \to \infty$. It follows from (7.174) and (7.182) that

$$\begin{aligned}&\langle u_k^*, x - x_k\rangle \le \varphi(x,y) - \psi(x,y) - \varphi(x_k,y_k) + \psi(x_k,y_k)\\ &+2\varepsilon_k\big(\|x-x_k\| + \|y-y_k\|\big) \le \varphi(x,y) - \varphi(x_k,y_k) - \langle x_k^*, x - x_k\rangle - \langle y_k^*, y - y_k\rangle\\ &+2\varepsilon_k\big(\|x-x_k\| + \|y-y_k\|\big)\ \text{ for all }\ (x,y)\in\Gamma\cap\big((x_k,y_k)+\eta_k\mathbb{B}\big).\end{aligned}$$

The latter implies in turn that the inequality

$$\begin{aligned}\langle u_k^* + x_k^*, x - x_k\rangle + \langle y_k^*, y - y_k\rangle \le \varphi(x,y) - \varphi(x_k,y_k)&\\ +2\varepsilon_k\big(\|x-x_k\| + \|y-y_k\|\big)&\end{aligned}$$

valid for all such (x, y) can be written as the ε-subdifferentials inclusion

$$(u_k^* + x_k^*, y_k^*) \in \widehat{\partial}_{2\varepsilon_k}\big(\varphi + \delta(\cdot;\Gamma)\big)(x_k,y_k)\ \text{ for all }\ k\in\mathbb{N}.$$

Passing now to the limit as $k\to\infty$ and taking into account the weak* convergence $(u_k^* + x_k^*, y_k^*) \stackrel{w^*}{\to} (u^* + x^*, y^*)$, we get from definition (7.180) that

$$(u^* + x^*, y^*) \in \partial\big(\varphi + \delta(\cdot;\Gamma)\big)(\bar{x},\bar{y}). \tag{7.183}$$

Since the function $\varphi+\delta(\cdot;\Gamma)$ is convex on $X\times Y$, the basic subdifferential in (7.183) reduces to the one of convex analysis. Thus applying to (7.183) the subdifferential sum rule for infinite systems from Corollary 7.52, which holds under the imposed CQC, gives us the inclusion

$$\partial\big(\varphi + \delta(\cdot;\Gamma)\big)(\bar{x},\bar{y}) \subset \partial\varphi(\bar{x},\bar{y}) + \bigcup_{\lambda\in A(\bar{x},\bar{y})}\Big[\sum_{t\in\operatorname{supp}\lambda}\lambda_t\partial\varphi_t(\bar{x},\bar{y})\Big] + N\big((\bar{x},\bar{y});\Omega\big)$$

with $A(\bar{x},\bar{y}) = \{\lambda\in\mathbb{R}_+^{(T)} \mid \lambda_t\varphi_t(\bar{x},\bar{y}) = 0\ \text{ for all }\ t\in\operatorname{supp}\lambda\}$. Substituting it into (7.183) and taking into account the aforementioned relationships between the full and partial subdifferentials of convex functions, we arrive at

$$\begin{cases} u^* \in \partial_x\varphi(\bar{x},\bar{y}) - x^* + \displaystyle\sum_{t\in\operatorname{supp}\lambda}\lambda_t\partial_x\varphi_t(\bar{x},\bar{y}) + N_x\big((\bar{x},\bar{y});\Omega\big),\\ y^* \in \partial_y\varphi(\bar{x},\bar{y}) + \displaystyle\sum_{t\in\operatorname{supp}\lambda}\lambda_t\partial_y\varphi_t(\bar{x},\bar{y}) + N_y\big((\bar{x},\bar{y});\Omega\big)\end{cases}$$

for some $\lambda\in A(\bar{x},\bar{y})$. This completes the proof of the theorem. △

As discussed above, the inner subdifferential stability of the minus term ψ in (7.170) required in Theorem 7.60 is a rather restrictive requirement. In the next theorem, we replace it by a much more flexible assumption on ψ that holds, in particular, for any *continuous* convex functions. The upper estimate for basic subgradients of (7.173) obtained under the following assumption is less precise in comparison with Theorem 7.60 while being sufficient for the majority of applications including those in this book.

Definition 7.61 (Subdifferential Boundedness). *We say that a convex function* $\psi\colon X \to \overline{\mathbb{R}}$ *is* SUBDIFFERENTIALLY BOUNDED *around* $\bar{x} \in \operatorname{dom}\psi$ *if for any sequences* $\varepsilon_k \downarrow 0$ *and* $x_k \xrightarrow{\operatorname{dom}\psi} \bar{x}$ *as* $k \to \infty$ *there is a sequence of* $x_k^* \in \partial_{\varepsilon_k}\psi(x_k)$, $k \in \mathbb{N}$, *such that the set* $\{x_k^* |\ k \in \mathbb{N}\}$ *is bounded in* X^*.

As mentioned, this property holds for a broad class of convex functions.

Proposition 7.62 (Sufficient Condition for Subdifferential Boundedness of Convex Functions). *Let* $\psi\colon X \to \overline{\mathbb{R}}$ *be a convex function continuous at* $\bar{x} \in \operatorname{int}(\operatorname{dom}\psi)$. *Then* ψ *is subdifferentially bounded around this point.*

Proof. As well known in convex analysis (see Exercise 7.102), the continuity of a convex function ψ at the reference point $\bar{x} \in \operatorname{int}(\operatorname{dom}\psi)$ yields that ψ is locally Lipschitzian around $\bar{x}$. On the other hand, the local Lipschitz continuity of any (not only convex) function ensures the uniform boundedness of subgradients around the point in question; see Exercise 1.69. Furthermore, $\partial\psi(x) \subset \partial_\varepsilon\psi(x)$ for any $\varepsilon > 0$. Taking now arbitrary sequences $\varepsilon_k \downarrow 0$ and $x_k \xrightarrow{\operatorname{dom}\psi} \bar{x}$ as $k \to \infty$, we have $x_k^* \in \partial_{\varepsilon_k}\psi(x_k)$ for any sequence of subgradients $x_k^* \in \partial\psi(x_k)$. This justifies the subdifferential boundedness of ψ. △

The following theorem provides a result independent of Theorem 7.60. Its proof involves the classical Brøndsted-Rockafellar theorem on subdifferential density in convex analysis, which is a predecessor and convex counterpart of the fundamental Ekeland's variational principle.

Theorem 7.63 (Basic Subgradients of Value Functions in DC Programs Under Subdifferential Boundedness). *Suppose that for both spaces* X *and* Y *the dual unit balls are sequentially weak* compact, that the argminimum mapping* (7.24) *is inner semicontinuous at some point* $(\bar{x}, \bar{y}) \in \operatorname{gph} M$, *that* ψ *in* (7.173) *is subdifferentially bounded around* $(\bar{x}, \bar{y})$, *and that CQC holds for* $(\varphi, \varphi_t, \Omega)$. *Then we have the upper estimate*

$$\partial\mu(\bar{x}) \subset \partial_x\varphi(\bar{x}, \bar{y}) + \bigcup_{(x^*, y^*)\in\partial\psi(\bar{x},\bar{y})} \Big\{ -x^* + \bigcup_{\lambda\in\Lambda(\bar{x},\bar{y},y^*)} \Big[\sum_{t\in\operatorname{supp}\lambda} \lambda_t\partial_x\varphi_t(\bar{x}, \bar{y})\Big]\Big\}$$
$$+N_x\big((\bar{x}, \bar{y}); \Omega\big).$$

Proof. Pick any $u^* \in \partial\mu(\bar{x})$, and similar to the proof of Theorem 7.60, find sequences $\varepsilon_k \downarrow 0$, $x_k \xrightarrow{\mu} \bar{x}$, and $u_k^* \in \widehat{\partial}_{\varepsilon_k}\mu(x_k)$ satisfying $u_k^* \xrightarrow{w^*} u^*$ as $k \to \infty$. Then we get $\eta_k \downarrow 0$ such that inequality (7.182) holds and, by the assumed inner semicontinuity of $M(\cdot)$, obtain a sequence of $y_k \in M(x_k)$ converging to $\bar{y}$ as $k \to \infty$. Select further $\nu_k > 0$ with $2\sqrt{\nu_k} < \eta_k$ and, by taking into account that $\nu_k \downarrow 0$ and $(x_k, y_k) \to (\bar{x}, \bar{y})$ and by employing the subdifferential boundedness of ψ, find a sequence of $(x_k^*, y_k^*) \in \partial_{\nu_k}\psi(x_k, y_k)$ such that the set $\{(x_k^*, y_k^*) \in X^* \times Y^* |\ k \in \mathbb{N}\}$ is bounded. It follows from the structure of the ε-subdifferential mapping (7.101) that $(x^*, y^*) \in \partial\psi(\bar{x}, \bar{y})$. Similar to the proof of Theorem 7.60, we derive from (7.182) the inequality

$$\begin{aligned}\langle u_k^* + x_k^*, x - x_k\rangle &+\langle y_k^*, y - y_k\rangle - \nu_k \le \varphi(x,y) - \varphi(x_k,y_k)\\ &+2\varepsilon_k\big(\|x - x_k\| + \|y - y_k\|\big)\end{aligned}$$

held for all $(x,y) \in \Gamma \cap \big((x_k,y_k) + \eta_k\mathbb{B}\big)$. This implies that

$$(u_k^* + x_k^*, y_k^*) \in \partial_{\nu_k}\vartheta_k(x_k,y_k), \quad k \in \mathbb{N}, \tag{7.184}$$

in terms of the ε-subdifferentials (with $\varepsilon := \nu_k$) of the convex l.s.c. functions $\vartheta_k(\cdot)$ given in the summation form

$$\begin{aligned}\vartheta_k(x,y) :&= \varphi(x,y) + \delta\big((x,y); \Gamma \cap [(x_k,y_k) + \eta_k\mathbb{B}]\big)\\ &-\varphi(x_k,y_k) + 2\varepsilon_k\big(\|x - x_k\| + \|y - y_k\|\big).\end{aligned} \tag{7.185}$$

Applying now to the elements in (7.184) the Brøndsted-Rockafellar density theorem, we find pairs $(\widetilde{x}_k, \widetilde{y}_k) \in \operatorname{dom}\vartheta_k$ and $(\widetilde{x}_k^*, \widetilde{y}_k^*) \in \partial\vartheta_k(\widetilde{x}_k, \widetilde{y}_k)$ satisfying for all $k \in \mathbb{N}$ the following inequalities:

$$\begin{aligned}&\|\widetilde{x}_k - x_k\| + \|\widetilde{y}_k - y_k\| \le \sqrt{\nu_k} \ \text{ and}\\ &\|\widetilde{x}_k^* - (u_k^* + x_k^*)\| + \|\widetilde{y}_k^* - y_k^*\| \le \sqrt{\nu_k}.\end{aligned} \tag{7.186}$$

They imply by the constructions above and the choice of ν_k that

$$\begin{aligned}&\langle \widetilde{x}_k^*, x - \widetilde{x}_k\rangle + \langle \widetilde{y}_k^*, y - \widetilde{y}_k\rangle \le \vartheta_k(x,y) - \vartheta_k(\widetilde{x}_k,\widetilde{y}_k) \le \varphi(x,y) - \varphi(\widetilde{x}_k,\widetilde{y}_k)\\ &+2\varepsilon_k\big(\|x - x_k\| + \|y - y_k\|\big) - 2\varepsilon_k\big(\|\widetilde{x}_k - x_k\| + \|\widetilde{y}_k - y_k\|\big)\\ &\le \varphi(x,y) - \varphi(\widetilde{x}_k,\widetilde{y}_k) + 2\varepsilon_k\big(\|x - \widetilde{x}_k\| + \|y - \widetilde{y}_k\|\big)\end{aligned}$$

for all $(x,y) \in \Gamma \cap \big((x_k,y_k) + \eta_k\mathbb{B}\big)$, which yields in turn the inclusions

$$(\widetilde{x}_k^*, \widetilde{y}_k^*) \in \widehat{\partial}_{2\varepsilon_k}\big(\varphi + \delta(\cdot;\Gamma)\big)(\widetilde{x}_k,\widetilde{y}_k), \quad k \in \mathbb{N}. \tag{7.187}$$

It easily follows from the convergence $(x_k,y_k) \to (\bar{x},\bar{y})$, $(u_k^* + x_k^*, y_k^*) \xrightarrow{w^*} (u^* + x^*, y^*)$ and from the norm estimates in (7.186) that

$$(\widetilde{x}_k,\widetilde{y}_k) \to (\bar{x},\bar{y}) \ \text{ and } \ (\widetilde{x}_k^*,\widetilde{y}_k^*) \xrightarrow{w^*} (u^* + x^*, y^*) \ \text{ as } \ k \to \infty.$$

Thus passing to the limit in (7.187) as $k \to \infty$ and using construction (7.180) of the basic subdifferential, we arrive at inclusion (7.183) as in the proof of Theorem 7.60, where the basic subdifferential agrees with the subdifferential of convex analysis for the convex function $\varphi + \delta(\cdot;\Gamma)$. Proceeding finally as in the proof of Theorem 7.60 by employing the subdifferential sum rule from Corollary 7.52, we complete the proof of the theorem. △

Our next results concern the singular subdifferential $\partial^\infty\mu(\bar{x})$ of the DC value function (7.173). According to (1.38) and Exercise 1.68, the singular subdifferential of any l.s.c. $\varphi\colon X \to \overline{\mathbb{R}}$ on a Banach space X is defined by

$$\partial^\infty\varphi(\bar{x}) := \operatorname*{Lim\,sup}_{\substack{x \xrightarrow{\varphi} \bar{x} \\ \lambda,\,\varepsilon\downarrow 0}} \lambda\widehat{\partial}_\varepsilon\varphi(x) \tag{7.188}$$

via the sequential outer limit, where $\varepsilon > 0$ can be omitted if X is Asplund.

Theorem 7.64 (Singular Subgradients of Value Functions in DC Programs). *Suppose that the assumptions of Theorem 7.63 are satisfied with replacing CQC for* $(\varphi, \varphi_t, \Omega)$ *by the corresponding FMCQ (7.159) for* (φ_t, Ω) *in (7.170). Assume in addition that* $\Gamma \subset \operatorname{dom}\varphi$ *for the set of feasible solutions (7.175). Then we have the upper estimate*

$$\partial^\infty\mu(\bar{x}) \subset \bigcup_{\lambda\in\Lambda^\infty(\bar{x},\bar{y})} \Big[\sum_{t\in\operatorname{supp}\lambda} \lambda_t\partial_x\varphi_t(\bar{x},\bar{y})\Big] + N_x\big((\bar{x},\bar{y});\Omega\big), \tag{7.189}$$

where the set of singular multipliers is defined by

$$\Lambda^\infty(\bar{x},\bar{y}) := \Big\{\lambda\in\mathbb{R}^{(T)}_+ \Big|\ 0\in\sum_{t\in\operatorname{supp}\lambda}\lambda_t\partial_y\varphi_t(\bar{x},\bar{y}) + N_y\big((\bar{x},\bar{y});\Omega\big),$$
$$\lambda_t\varphi_t(\bar{x},\bar{y}) = 0 \ \textit{for all}\ \ t\in\operatorname{supp}\lambda\Big\}.$$

Proof. Pick any $u^*\in\partial^\infty\mu(\bar{x})$, and by (7.188), find sequences

$$\lambda_k\downarrow 0,\ \ \varepsilon_k\downarrow 0,\ \ x_k\xrightarrow{\mu}\bar{x},\ \ u_k^*\in\widehat{\partial}_{\varepsilon_k}\mu(x_k)\ \text{ with }\ \lambda_k u_k^*\xrightarrow{w^*}u^*\ \text{as}\ \ k\to\infty.$$

Following the proof of Theorem 7.63, select sequences

$$\nu_k\downarrow 0\ \text{ as }\ k\to\infty,\quad y_k\in M(x_k),\ \text{ and }\ (x_k^*,y_k^*)\in\partial_{\nu_k}\psi(x_k,y_k),\quad k\in\mathbb{N},$$

such that $\{(x_k^*,y_k^*)\}$ weak* converges in $X^*\times Y^*$ to some $(x^*,y^*)\in\partial\psi(\bar{x},\bar{y})$. Further, the application of the Brøndsted-Rockafellar theorem to the function $\vartheta_k(x,y)$ from (7.185) gives us sequences of $(\widetilde{x}_k,\widetilde{y}_k)\in\operatorname{dom}\vartheta_k$ and $(\widetilde{x}_k^*,\widetilde{y}_k^*)\in\partial\vartheta_k(\widetilde{x}_k,\widetilde{y}_k)$ satisfying the estimates in (7.186) and the subdifferential inclusions (7.187) for all $k\in\mathbb{N}$. Using the convexity of $\varphi+\delta(\cdot;\Gamma)$ and the assumption on $\Gamma\subset\operatorname{dom}\varphi$ allows us to rewrite (7.187) as

$$\langle\widetilde{x}_k^*, x-\widetilde{x}_k\rangle + \langle\widetilde{y}_k^*, y-\widetilde{y}_k\rangle \le \varphi(x,y)-\varphi(\widetilde{x}_k,\widetilde{y}_k)+2\varepsilon_k\big(\|x-\widetilde{x}_k\|+\|y-\widetilde{y}_k\|\big)$$

for all $(x,y)\in\Gamma$ and $k\in\mathbb{N}$. This implies, by picking any $\gamma>0$ and employing the lower semicontinuity of φ around $(\bar{x},\bar{y})$, that

$$\lambda_k\big[\langle\widetilde{x}_k^*, x-\widetilde{x}_k\rangle + \langle\widetilde{y}_k^*, y-\widetilde{y}_k\rangle\big] \le \lambda_k\big[\varphi(x,y)-\varphi(\widetilde{x}_k,\widetilde{y}_k)+2\varepsilon_k\big(\|x-\widetilde{x}_k\|$$
$$+\|y-\widetilde{y}_k\|\big)\big] \le \lambda_k\big[\varphi(x,y)-\varphi(\bar{x},\bar{y})+\gamma+2\varepsilon_k\big(\|x-\widetilde{x}_k\|+\|y-\widetilde{y}_k\|\big)\big]$$

for all $(x,y)\in\Gamma$ and all $k\in\mathbb{N}$ sufficiently large. Passing now to the limit as $k\to\infty$ and taking into account that the sequence $\{\widetilde{y}_k^*\}$ is bounded in Y^*, that $\lambda_k\downarrow 0$, and that $\lambda_k\widetilde{x}_k^*\xrightarrow{w^*}u^*$ by (7.186), we get the relationship

$$\langle u^*, x - \bar{x}\rangle \le 0 \text{ for all } (x, y) \in \Gamma,$$

which can be rewritten as $(u^*, 0) \in N((\bar{x}, \bar{y}); \Gamma)$. Applying the normal cone calculus for infinite systems from Corollary 7.53 gives us

$$(u^*, 0) \in \bigcup_{\lambda \in A(\bar{x},\bar{y})} \Big[\sum_{t \in \operatorname{supp}\lambda} \lambda_t \partial\varphi_t(\bar{x}, \bar{y})\Big] + N\big((\bar{x}, \bar{y}); \Omega\big)$$

with $A(\bar{x}, \bar{y}) = \{\lambda \in \mathbb{R}^{(T)}_+ |\ \lambda_t\varphi_t(\bar{x}, \bar{y}) = 0,\ t \in \operatorname{supp}\lambda\}$. The latter yields (7.189) and thus completes the proof of the theorem. △

The next theorem presents applications of the upper estimates for both basic and singular subdifferentials of the value function (7.173) established in Theorems 7.63 and 7.64 to derive efficient conditions ensuring the local Lipschitz continuity of (7.173) via the initial data as well as necessary optimality conditions for local optimality in the parametric DC semi-infinite program under consideration. The obtained results essentially use the *Asplund* property of the *parameter* space X; this is not required for the decision space Y.

Recall that characterizing the local Lipschitz continuity of any l.s.c. function φ on an Asplund space presented in Exercise 4.34(ii) involves both the triviality condition $\partial^\infty\varphi(\bar{x}) = \{0\}$ for the singular subdifferential and the SNEC property of φ at the reference point in the case of infinite dimensions. While the condition $\partial^\infty\mu(\bar{x}) = \{0\}$ for the value function (7.173) is straightforward from Theorem 7.64, it is not the case for SNEC, which is fully independent from the above triviality condition. Nevertheless, the following lemma of its own interest shows that for the general class of *marginal/value functions*, including the one in (7.173), the SNEC property holds under natural assumptions on the initial problem data.

Lemma 7.65 (SNEC Property of Marginal Functions). *Let*

$$\mu(x) := \inf\big\{\phi(x, y)\big|\ y \in \Phi(x)\big\}, \quad x \in X, \tag{7.190}$$

where X is Asplund, where the argminimum map

$$x \mapsto S(x) := \big\{y \in \Phi(x)\big|\ \phi(x, y) = \mu(x)\big\}$$

is inner semicontinuous at some point $(\bar{x}, \bar{y}) \in \operatorname{gph} S$ and where ϕ is locally Lipschitzian around this point. Then (7.190) is SNEC at $\bar{x}$ provided that it is l.s.c. around $\bar{x}$ and that the mapping Φ therein is Lipschitz-like around $(\bar{x}, \bar{y})$.

Proof. To verify the SNEC property of (7.190) at $\bar{x}$, we use its subdifferential characterization presented in Exercise 2.50. Based on this, take any sequences $\lambda_k \downarrow 0$, $x_k \stackrel{\mu}{\to} \bar{x}$, and $x_k^* \in \lambda_k\widehat{\partial}\mu(x_k)$ with $x_k^* \stackrel{w^*}{\to} 0$, and then show that $\|x_k^*\| \to 0$ along some subsequence. To proceed, employ the inner semicontinuity of $S(\cdot)$ at $(\bar{x}, \bar{y})$ and select a sequence of $y_k \in S(x_k)$ whose subsequence converges (with no relabeling) to $\bar{y}$. Take $\widetilde{x}_k^* \in \widehat{\partial}\mu(x_k)$ such that $x_k^* = \lambda_k\widetilde{x}_k^*$. Since $\widetilde{x}_k^*$ is a regular subgradient of φ at x_k, for any $\eta > 0$, there is $\gamma > 0$ such that

$$\langle\widetilde{x}_k^*, x - x_k\rangle \le \mu(x) - \mu(x_k) + \eta\|x - x_k\| \;\text{ whenever }\; x \in x_k + \gamma\mathbb{B}.$$

Considering the extended-real-valued function

$$\xi(x, y) := \phi(x, y) + \delta\big((x, y); \operatorname{gph}\Phi\big) \;\text{ for all }\; (x, y) \in X \times Y,$$

we easily conclude from the above that

$$\big\langle(\widetilde{x}_k^*, 0), (x - x_k, y - y_k)\big\rangle \le \xi(x, y) - \xi(x_k, y_k) + \eta\big(\|x - x_k\| + \|y - y_k\|\big)$$

whenever $(x, y) \in (x_k, y_k) + \gamma\mathbb{B}$, which means that $(\widetilde{x}_k^*, 0) \in \widehat{\partial}\xi(x_k, y_k)$.

Fix now an arbitrary sequence $\varepsilon_k \downarrow 0$ as $k \to \infty$. Since ξ is locally Lipschitzian around $(\bar{x}, \bar{y})$, while X and Y are Asplund, we apply the fuzzy sum rule from Exercise 2.42 to the summation function ξ at (x_k, y_k) and thus find, by taking into account the convergence above, sequences

$$\begin{array}{l}(x_{1k}, y_{1k}) \stackrel{\phi}{\to} (\bar{x}, \bar{y}),\; (x_{2k}, y_{2k}) \stackrel{\operatorname{gph}\Phi}{\longrightarrow} (\bar{x}, \bar{y}) \;\text{ as }\; k \to \infty,\\ (x_{1k}^*, y_{1k}^*) \in \widehat{\partial}\phi(x_{1k}, y_{1k}), \;\text{ and }\; (x_{2k}^*, y_{2k}^*) \in \widehat{N}\big((x_{2k}, y_{2k}); \operatorname{gph}\Phi\big)\end{array}$$

such that $\lambda_k\|(x_{1k}^*, y_{1k}^*)\| \to (0, 0)$ with the estimates

$$\|\widetilde{x}_k^* - x_{1k}^* - x_{2k}^*\| \le \varepsilon_k \;\text{ and }\; \|y_{1k}^* + y_{2k}^*\| \le \varepsilon_k \;\text{ as }\; k \in \mathbb{N}. \tag{7.191}$$

This implies that $\lambda_k\|y_{2k}^*\| \to 0$ as $k \to \infty$. Taking now into account that

$$\big(\lambda_k x_{2k}^*, \lambda_k y_{2k}^*\big) \in \widehat{N}\big((x_{2k}, y_{2k}); \operatorname{gph}\Phi) \Longleftrightarrow \lambda_k x_{2k}^* \in \widehat{D}^*\Phi(x_{2k}, y_{2k})(-\lambda_k y_{2k}^*)$$

and that Φ is Lipschitz-like around $(\bar{x}, \bar{y})$ with some modulus $\ell > 0$, we get from the coderivative estimate for Lipschitz-like mappings (see implication (a)$\Rightarrow$(b) of Exercise 3.41, which holds in any Banach space) that

$$\|\lambda_k x_{2k}^*\| \le \ell\|\lambda_k y_{2k}^*\| \;\text{ for large }\; k \in \mathbb{N}.$$

This clearly yields $\lambda_k\|x_{2k}^*\| \to 0$. Combining the latter with (7.191) and with $x_k^* = \lambda_k\widetilde{x}_k^*$, we conclude that $\|x_k^*\| \to 0$ as $k \to \infty$ and thus justify the SNEC property of μ at $\bar{x}$ claimed in the lemma. $\triangle$

Now we are ready to establish the aforementioned major theorem.

Theorem 7.66 (Lipschitz Continuity of Value Functions and Optimality Conditions for Parametric DC SIPs). *Let the parameter space X be Asplund in the assumptions of Theorem 7.64 and suppose in addition that*

$$\Big\{\bigcup_{\lambda\in\Lambda^\infty(\bar{x},\bar{y})}\Big[\sum_{t\in\operatorname{supp}\lambda}\lambda_t\partial_x\varphi_t(\bar{x}, \bar{y})\Big] + N_x\big((\bar{x}, \bar{y}); \Omega\big)\Big\} = \{0\}. \tag{7.192}$$

Then the value function $\mu(\cdot)$ is locally Lipschitzian around $\bar{x}$ provided that it is l.s.c. around this point (which is ensured by the inner semicontinuity of $M(\cdot)$ around

$(\bar{x}, \bar{y})$) in each of the following two cases: either **(a)** $\dim X < \infty$ *or* **(b)** *both* φ *and* ψ *are continuous at* $(\bar{x}, \bar{y})$, *and the constraint mapping* $x \mapsto F(x) \cap G(x)$ *is Lipschitz-like around* $(\bar{x}, \bar{y})$.

If furthermore CQC holds for $(\varphi, \varphi_t, \Omega)$, *then we have the following necessary optimality conditions for the* (*global*) *minimizer* $\bar{y}$ *of the DC program* (7.179)*: there are* $(x^*, y^*) \in \partial\psi(\bar{x}, \bar{y})$, $u^* \in X^*$, *and* $\lambda \in \mathbb{R}^{(T)}_+$ *satisfying*

$$\begin{cases} u^* + x^* \in \partial_x \varphi(\bar{x}, \bar{y}) + \sum\limits_{t \in \operatorname{supp} \lambda} \lambda_t \partial_x \varphi_t(\bar{x}, \bar{y}) + N_x\big((\bar{x}, \bar{y}); \Omega\big), \\ y^* \in \partial_y \varphi(\bar{x}, \bar{y}) + \sum\limits_{t \in \operatorname{supp} \lambda} \lambda_t \partial_y \varphi_t(\bar{x}, \bar{y}) + N_y\big((\bar{x}, \bar{y}); \Omega\big), \\ \lambda_t \varphi_t(\bar{x}, \bar{y}) = 0 \ \text{ for all } \ t \in \operatorname{supp} \lambda. \end{cases} \tag{7.193}$$

Proof. If (7.192) holds, then $\partial^\infty \mu(\bar{x}) = \{0\}$ by Theorem 7.64. Further, it is easy to derive directly from the definitions that the lower semicontinuity of $\mu(\cdot)$ around $\bar{x}$ follows from the inner semicontinuity of $M(\cdot)$ around $(\bar{x}, \bar{y})$. Thus the local Lipschitz continuity of $\mu(\cdot)$ around $\bar{x}$ is a consequence of Theorem 1.22 in the case (a) where X is finite-dimensional.

In case (b), recall that the continuity of the convex functions φ and ψ at $(\bar{x}, \bar{y})$ implies their Lipschitz continuity around this point, and thus $\mu(\cdot)$ is SNEC at $\bar{x}$ due to Lemma 7.65. This verifies the first part of the theorem.

To justify the second part on the necessary optimality conditions, observe that any $\bar{y} \in M(\bar{x})$ under the consideration in this theorem is a *global* solution to (7.179). It follows from the local Lipschitz continuity of μ around $\bar{x}$ that $\partial\mu(\bar{x}) \neq \emptyset$; see Exercise 2.32(ii). Thus using the upper estimate of $\partial\mu(\bar{x})$ in Theorem 7.63 under the assumed CQC for $(\varphi, \varphi_t, \Omega)$, we conclude that the set on the right-hand side of this estimate is nonempty as well. This yields the claimed necessary optimality conditions (7.193) by construction (7.176) of the KKT multiplier set $\Lambda(\bar{x}, \bar{y}, y^*)$. △

Note that, in contrast to the necessary optimality conditions of Corollary 7.57, the results of (7.193) give us *lower* subdifferential optimality conditions in the enhanced form (with $\gamma = 0$ instead of $\gamma > 0$ in Corollary 7.57) under different while easily verifiable assumptions. Note also that the results of Sections 7.1 and 7.3 provide *characterizations* of the Lipschitz-like property of the infinite constraint inequality system in (7.179) entirely via the functions φ_t for the cases of linear, block-perturbed, and convex structures.

Convex ($\psi \equiv 0$) and concave ($\varphi \equiv 0$) SIPs are particular cases of the DC programs under consideration, and so the obtained results for general DC SIPs can be directly applied to these important cases with the corresponding specifications. Furthermore, the convex case allows us to derive new results, which cannot be deduced from those for general DC SIPs obtained above. The next theorem establishes a *precise formula* (equality, not inclusion) for calculating the subdifferential of the convex value function in such SIPs.

Theorem 7.67 (Calculating Subgradients of Value Functions in Convex SIPs). *Consider the value function* $\mu(\cdot)$ *from* (7.173) *with* $\psi \equiv 0$, *and suppose that CQC* *holds for the convex triple* $(\varphi, \varphi_t, \Omega)$ *in general Banach spaces. Then* $\mu(\cdot)$ *is convex, and its subdifferential at* $\bar{x} \in \operatorname{dom} \mu$ *is calculated by*

$$\partial\mu(\bar{x}) = \Big\{x^* \in X^* \Big| \ (x^*, 0) \in \partial\varphi(\bar{x}, \bar{y}) + \bigcup_{\lambda \in A(\bar{x},\bar{y})} \Big[\sum_{t \in \operatorname{supp}\lambda} \lambda_t \partial\varphi_t(\bar{x}, \bar{y})\Big] + N\big((\bar{x}, \bar{y}); \Omega\big)\Big\} \ \textit{for any} \ \ \bar{y} \in M(\bar{x}),$$

where $A(\bar{x}, \bar{y}) := \big\{\lambda \in \mathbb{R}_+^{(T)} \big| \ \lambda_t \varphi_t(\bar{x}, \bar{y}) = 0, \ t \in \operatorname{supp}\lambda\big\}$.

Proof. The convexity of the value function (7.173) with $\psi \equiv 0$ and all the convex data easily follows from its definition and the convexity assumptions. To verify first the inclusion "$\subset$" in the claimed formula for $\partial\mu(\bar{x})$, we proceed as in the proof of Theorem 7.55 by taking $\gamma = 0$ and $\eta = \infty$.

To justify the opposite inclusion, pick any x^* from the right-hand side therein and thus find $\lambda \in A(\bar{x}, \bar{y})$, $(u^*, v^*) \in \partial\varphi(\bar{x}, \bar{y})$, $(u_t^*, v_t^*) \in \partial\varphi_t(\bar{x}, \bar{y})$, and $(\widetilde{u}^*, \widetilde{v}^*) \in N((\bar{x}, \bar{y}); \Omega)$ satisfying the equality

$$(x^*, 0) = (u^*, v^*) + \sum_{t \in \operatorname{supp}\lambda} \lambda_t (u_t^*, v_t^*) + (\widetilde{u}^*, \widetilde{v}^*).$$

It follows from the construction of $A(\bar{x}, \bar{y})$ that for the chosen pairs (u^*, v^*), (u_t^*, v_t^*), and $(\widetilde{u}^*, \widetilde{v}^*)$, we have the relationships

$$\begin{cases} \varphi(x, y) - \mu(\bar{x}) = \varphi(x, y) - \varphi(\bar{x}, \bar{y}) \ge \langle u^*, x - \bar{x}\rangle + \langle v^*, y - \bar{y}\rangle, \\ 0 \ge \lambda_t \varphi_t(x, y) - \lambda_t \varphi_t(\bar{x}, \bar{y}) \ge \lambda_t \langle u_t^*, x - \bar{x}\rangle + \lambda_t \langle v_t^*, y - \bar{y}\rangle, \quad t \in \operatorname{supp}\lambda, \\ 0 \ge \langle \widetilde{u}^*, x - \bar{x}\rangle + \langle \widetilde{v}^*, y - \bar{y}\rangle \ \text{ whenever } \ (x, y) \in \Gamma, \end{cases}$$

which imply together with the above equality that

$$\varphi(x, y) + \delta\big((x, y); \Gamma\big) - \mu(\bar{x}) \ge \langle x^*, x - \bar{x}\rangle \ \text{ for all } \ (x, y) \in X \times Y.$$

The latter shows in turn that $\mu(x) - \mu(\bar{x}) \ge \langle x^*, x - \bar{x}\rangle$ for all $x \in X$ and hence completes the proof of the theorem. △

7.5.4 Bilevel Semi-infinite Programs with Convex Data

In this subsection, we return to optimistic bilevel programs studied in Chapter 6 for the case of finitely many inequality constraints at the lower level described by smooth as well as by locally Lipschitzian functions on finite-dimensional spaces. Here we consider *fully convex* bilevel programs in arbitrary *Banach* spaces with *infinite* constraints and derive for them necessary optimality conditions, which cannot be deduced from the results of Chapter 6 even in the case finitely many constraints in $\mathbb{R}^n$. Developing the value function approach allows us to reduce the bilevel programs under consideration to single-level DC SIPs and then apply the results obtained above in Section 7.5.

Consider the optimistic bilevel program

$$\begin{cases} \text{minimize} \ \ f(x,y) \ \ \text{subject to} \\ y \in M(x) := \big\{ y \in G(x) \big| \ \varphi(x,y) = \mu(x) \big\}, \end{cases} \tag{7.194}$$

where $M(x)$ is the set of optimal solutions to the lower-level problem

$$\text{minimize } \varphi(x,y) \text{ subject to } y \in G(x) := \big\{ y \in Y \big| \ \varphi_t(x,y) \le 0, \ t \in T \big\}$$

with an arbitrary index set T, and where $\mu(\cdot)$ is the optimal value function of the parametric lower-level problem defined by

$$\mu(x) := \inf \big\{ \varphi(x,y) \big| \ y \in G(x) \big\}. \tag{7.195}$$

The standing assumption of this section is that the bilevel problem (7.194) is *fully convex* on the Banach spaces X, Y meaning that all the functions there are l.s.c. and convex with respect to both variables.

To evaluate subgradients of the value function (7.195) and derive necessary optimality conditions for (7.194), we proceed via penalization under partial calmness. Observe that all the results of Subsection 6.2.3 apply to problem (7.194) with no change. Based on them, we get that any partially calm feasible solution $(\bar{x}, \bar{y})$ to (7.194) is a local optimal solution to the single-level program:

$$\begin{cases} \text{minimize} \ \ \kappa^{-1} f(x,y) + \varphi(x,y) - \mu(x) \\ \text{subject to} \ \ \varphi_t(x,y) \le 0, \ \ t \in T, \end{cases} \tag{7.196}$$

where $\kappa > 0$ is the constant of partial calmness, provided that the upper-level objective f is continuous at $(\bar{x}, \bar{y})$. Let us first efficiently evaluate the convex subdifferential of the value function (7.195) in the lower-level program.

Theorem 7.68 (Subgradients of Value Functions in Convex Bilevel Programs). *Let $(\bar{x}, \bar{y})$ be a partially calm feasible solution to the fully convex bilevel program (7.194). Suppose that CQC holds for (φ, φ_t) and that f is continuous at $(\bar{x}, \bar{y})$. Then there is a number $\kappa > 0$ such that*

$$\partial \mu(\bar{x}) \times \{0\} \subset \kappa^{-1} \partial f(\bar{x}, \bar{y}) + \partial \varphi(\bar{x}, \bar{y}) + \bigcup_{\lambda \in A(\bar{x}, \bar{y})} \Big[\sum_{t \in \operatorname{supp} \lambda} \lambda_t \partial \varphi_t(\bar{x}, \bar{y}) \Big],$$

where the set $A(\bar{x}, \bar{y})$ of active constraint multipliers is defined in Theorem 7.67. In particular, we have the upper estimate

$$\partial \mu(\bar{x}) \subset \kappa^{-1} \partial_x f(\bar{x}, \bar{y}) + \partial_x \varphi(\bar{x}, \bar{y}) + \bigcup_{\lambda \in A(\bar{x}, \bar{y})} \Big[\sum_{t \in \operatorname{supp} \lambda} \lambda_t \partial_x \varphi_t(\bar{x}, \bar{y}) \Big].$$

Proof. The second inclusion in the theorem clearly follows from the first one; so we verify the latter. The assumptions made ensure that $(\bar{x}, \bar{y})$ a local minimizer of the penalized problem (7.196), which is a DC SIP of type (7.157) described by the l.s.c. convex functions

$$\vartheta(x, y) := \kappa^{-1} f(x, y) + \varphi(x, y), \quad \theta(x, y) := \mu(x), \ \vartheta_t(x, y) := \varphi_t(x, y)$$

with $\Theta = X \times Y$ in (7.11). Let us check that the imposed CQC for (φ, φ_t) yields the validity of CQC for (ϑ, ϑ_t). Using the structure of the feasible set

$$\Xi := \Big\{(x, y) \in X \times Y \Big| \ \varphi_t(x, y) \le 0 \ \text{ for all } \ t \in T\Big\}$$

in (7.196), the well-known conjugate representation from convex analysis

$$\text{epi}\,(\varphi_1 + \varphi_2)^* = \text{cl}^*(\text{epi}\,\varphi_1^* + \text{epi}\,\varphi_2^*), \tag{7.197}$$

which is valid for any l.s.c. convex functions such that $\text{dom}\,\varphi_1 \cap \text{dom}\,\varphi_2 \neq \emptyset$ with omitting the weak* closure if one of the functions is continuous at some point $\bar{x} \in \text{dom}\,\varphi_1 \cap \text{dom}\,\varphi_2$, and then employing the imposed CQC give us

$$\begin{aligned} \text{epi}\,\vartheta^* + \text{cone}\Big[\bigcup_{t\in T} \text{epi}\,\vartheta_t^*\Big] &= \text{epi}\,\big(\kappa^{-1} f\big)^* + \text{epi}\,\varphi^* + \text{cone}\Big[\bigcup_{t\in T} \text{epi}\,\varphi_t^*\Big] \\ &= \text{epi}\,\big(\kappa^{-1} f\big)^* + \text{epi}\,\big(\varphi + \delta(\cdot;\Xi)\big)^* = \text{epi}\,\big(\vartheta + \delta(\cdot;\Xi)\big)^*. \end{aligned}$$

Applying further (7.197) without the closure operation to the above sum function ϑ with the continuous term f implies that

$$\begin{aligned} \text{epi}\,\vartheta^* + \text{cone}\Big[\bigcup_{t\in T} \text{epi}\,\vartheta_t^*\Big] &= \text{epi}\,\big(\kappa^{-1} f\big)^* + \text{epi}\,\varphi^* + \text{cone}\Big[\bigcup_{t\in T} \text{epi}\,\varphi_t^*\Big] \\ &= \text{epi}\,\big(\kappa^{-1} f\big)^* + \text{epi}\,\big(\varphi + \delta(\cdot;\Xi)\big)^* = \text{epi}\,\big(\vartheta + \delta(\cdot;\Xi)\big)^* \end{aligned}$$

and thus allows us to conclude that the set

$$\text{epi}\,\vartheta^* + \text{cone}\Big[\bigcup_{t\in T} \text{epi}\,\vartheta_t^*\Big] \ \text{ is weak* closed in } \ X^* \times Y^* \times \mathbb{R}.$$

This is exactly the CQC property needed for the application of Theorem 7.51 to (7.196). Employing the latter result and the subdifferential sum rule

$$\partial\vartheta(\bar{x}, \bar{y}) = \partial\big(\kappa^{-1} f + \varphi\big)(\bar{x}, \bar{y}) = \kappa^{-1}\partial f(\bar{x}, \bar{y}) + \partial\varphi(\bar{x}, \bar{y}),$$

which holds by the continuity of f, we arrive at the first inclusion claimed in the theorem and thus complete the whole proof. △

Next we establish the main result of this subsection providing necessary optimality conditions for the fully convex bilevel programs with an arbitrary (finite or infinite) number of inequality constraints.

Theorem 7.69 (Necessary Optimality Condition for Fully Convex Bilevel SIPs). *Let $(\bar{x}, \bar{y})$ be a partially calm optimal solution to the fully convex bilevel program (7.194). Suppose that CQC holds for the lower-level program in (7.194), that the upper-level objective f is continuous at $(\bar{x}, \bar{y})$, and that $\partial\mu(\bar{x}) \neq \emptyset$ for the convex value function (7.195). Then for each $\widetilde{y} \in M(\bar{x})$ from the argminimum set in (7.194), there exist a number $\kappa > 0$ and multipliers $\lambda = (\lambda_t) \in \mathbb{R}_+^{(T)}$ and $\beta = (\beta_t) \in \mathbb{R}_+^{(T)}$ from the positive cone in (7.3) such that we have the following relationships:*

$$
\begin{aligned}
&0 \in \partial_x f(\bar{x}, \bar{y}) + \kappa\big[\partial_x\varphi(\bar{x}, \bar{y}) - \partial_x\varphi(\bar{x}, \widetilde{y})\big] + \sum_{t\in\operatorname{supp}\lambda} \lambda_t \partial_x\varphi_t(\bar{x}, \bar{y})\\
&\quad -\kappa \sum_{t\in\operatorname{supp}\beta} \beta_t \partial_x\varphi_t(\bar{x}, \widetilde{y}),\\
&0 \in \partial_y f(\bar{x}, \bar{y}) + \kappa\partial_y\varphi(\bar{x}, \bar{y}) + \sum_{t\in\operatorname{supp}\lambda} \lambda_t \partial_y\varphi_t(\bar{x}, \bar{y}),\\
&0 \in \partial_y\varphi(\bar{x}, \widetilde{y}) + \sum_{t\in\operatorname{supp}\beta} \beta_t \partial_y\varphi_t(\bar{x}, \widetilde{y}),\\
&\lambda_t\varphi_t(\bar{x}, \bar{y}) = \beta_t\varphi_t(\bar{x}, \widetilde{y}) = 0 \ \text{ for all } \ t \in T.
\end{aligned}
$$

Proof. Since $\partial\mu(\bar{x}) \neq \emptyset$, we take $x^* \in \partial\mu(\bar{x})$ and by Theorem 7.68 find $\kappa > 0$ and $\lambda \in \mathbb{R}_+^{(T)}$ satisfying the inclusion

$$
\kappa(x^*, 0) \in \partial f(\bar{x}, \bar{y}) + \kappa\partial\varphi(\bar{x}, \bar{y}) + \sum_{t\in\operatorname{supp}\lambda} \lambda_t \partial\varphi_t(\bar{x}, \bar{y}) \tag{7.198}
$$

with $\lambda_t\varphi_t(\bar{x}, \bar{y}) = 0$ as $t \in \operatorname{supp}\lambda$. On the other hand, picking $\widetilde{y} \in M(\bar{x})$ and applying to $x^* \in \partial\mu(\bar{x})$ the result of Theorem 7.67 give us $\beta \in \mathbb{R}_+^{(T)}$ such that

$$
x^* \in \partial_x\varphi(\bar{x}, \widetilde{y}) + \sum_{t\in\operatorname{supp}\beta} \partial_x\varphi_t(\bar{x}, \widetilde{y}), \quad 0 \in \partial_y\varphi(\bar{x}, \widetilde{y}) + \sum_{t\in\operatorname{supp}\beta} \partial_y\varphi_t(\bar{x}, \widetilde{y}),
$$

and $\beta_t\varphi_t(\bar{x}, \widetilde{y}) = 0$ for all $t \in \operatorname{supp}\beta$. Substituting this into (7.198) leads us to the claimed necessary optimality conditions. △

As an immediate consequence of Theorem 7.69, we get the following necessary optimality conditions for the bilevel SIP (7.194) involving only the reference optimal solution $(\bar{x}, \bar{y})$.

Corollary 7.70 (Specification of Necessary Optimality Conditions for Bilevel SIPs). *Let $(\bar{x}, \bar{y})$ be an optimal solution to (7.194) under the assumptions of Theorem 7.69. Then there are $\kappa > 0$ and $\lambda, \beta \in \mathbb{R}_+^{(T)}$ such that*

$$
\begin{aligned}
&0 \in \partial_x f(\bar{x}, \bar{y}) + \kappa\Big[\partial_x \varphi(\bar{x}, \bar{y}) - \partial_x \varphi(\bar{x}, \bar{y})\Big] + \sum_{t \in T}\Big[\big(\lambda_t - \kappa\beta_t\big)\partial_x \varphi_t(\bar{x}, \bar{y})\Big],\\
&0 \in \partial_y f(\bar{x}, \bar{y}) + \kappa\partial_y \varphi(\bar{x}, \bar{y}) + \sum_{t \in T} \lambda_t \partial_y \varphi_t(\bar{x}, \bar{y}),\\
&0 \in \partial_y \varphi(\bar{x}, \bar{y}) + \sum_{t \in T} \beta_t \partial_y \varphi_t(\bar{x}, \bar{y}),\\
&\lambda_t \varphi_t(\bar{x}, \bar{y}) = \beta_t \varphi_t(\bar{x}, \bar{y}) = 0 \ \textit{for all} \ t \in T.
\end{aligned}
$$

Proof. Follows directly from Theorem 7.69 by putting $\widetilde{y} = \bar{y} \in M(\bar{x})$ in the necessary optimality conditions obtained therein. △

It has been well recognized in convex analysis that the subdifferentiability assumption $\partial\mu(\bar{x}) \neq \emptyset$ imposed in Theorem 7.69 and Corollary 7.70 is not restrictive. In particular, it holds in the Banach space setting of (7.195) under certain primal and dual qualification conditions; see Exercise 7.110.

7.6 Exercises for Chapter 7

Exercise 7.71 (Dual Description of Consistency for Infinite Linear Inequality Systems). Verify the equivalence in (7.7) by using convex separation. *Hint:* Compare it with the proof of [210, Theorem 3.1].

Exercise 7.72 (Interiority Conditions for Infinite Linear Systems). Prove the following statements for infinite inequality systems $\mathcal{F}$ in (7.1):

(i) If $\operatorname{gph}\mathcal{F} \neq \emptyset$ and the set $\{a_t^* |\ t \in T\}$ is bounded, then $\operatorname{int}(\operatorname{gph}\mathcal{F}) \neq \emptyset$. *Hint:* Proceed similarly to the proof of implication (iii)⇒(v) in Theorem 7.2.

(ii) $\operatorname{int}(\operatorname{dom}\mathcal{F}) \neq \emptyset$ if $\operatorname{gph}\mathcal{F} \neq \emptyset$ without the boundedness assumption.

Exercise 7.73 (Extended Farkas Lemma). Verify Proposition 7.3. *Hint:* Compare it with the proof in [210, Lemma 2.4].

Exercise 7.74 (Distance Function Representation of the Exact Lipschitzian Bound). Verify formula (7.26). *Hint:* Employ the equivalent between the Lipschitz-like property of F and the metric regularity one for F^{-1} established in Theorem 3.2(ii) with the exact bound relationship therein, and then proceed by using Definition 3.1(b) of the exact regularity bound for F^{-1}.

Exercise 7.75 (Existence of Best Approximations). Justify the existence of solutions to the optimization problem (7.28). *Hint:* Use the Alaoglu-Bourbaki theorem and the continuity of the mapping $x^* \mapsto \langle x^*, x\rangle$ in the weak* topology of X^*.

Exercise 7.76 (Fenchel Conjugates). Given a proper function $\varphi\colon X \to \overline{\mathbb{R}}$, verify the convexity and lower semicontinuity of the Fenchel conjugate (7.30).

Exercise 7.77 (Fenchel Conjugates for Suprema of Linear Functions). Prove the representations in (7.40). *Hint:* Compare it with [121] and [297].

Exercise 7.78 (Coderivative Calculation for Infinite Linear Inequality Systems). Calculate the coderivative for the general linear inequality system given in (7.48). *Hint:* Proceed as in the proof of Theorem 7.5.

Exercise 7.79 (Farkas-Minkowski Property for Infinite Linear Inequalities). Give sufficient conditions for the validity of the Farkas-Minkowski property (7.49) for the infinite linear system (7.48).

Exercise 7.80 (Equivalent Descriptions of the Strong Slater Condition for the Infinite Linear Inequality Systems). Formulate and prove a counterpart of Theorem 7.2 for the infinite linear constraint systems defined in (7.48).

Exercise 7.81 (Farkas-Minkowski Property from Strong Slater Condition).

(i) Verify that (7.58) implies the Farkas-Minkowski property in finite dimensions provided that the set $\text{co}\{-c_t^*, a_t^*, b_t)|\ t \in T\}$ is compact, and clarify whether the latter condition is essential for this statement.

(ii) Does it hold in infinite-dimensional spaces?

(iii) Does it hold in infinite dimensions if the set on the right-hand side of (7.58) is replaced by its weak* closure?

(iv) Does the strong Slater condition (7.57) for infinite linear systems always imply the Farkas-Minkowski property in finite-dimensional spaces?

Exercise 7.82 (Nonempty Graphical Interior for Infinite Linear Systems). Let X and P be arbitrary Banach spaces in (7.48).

(i) Show that SSC (7.57) and the boundedness of the set $\{(a_t^*, c_t^*)|\ t \in T\}$ in $X^* \times P^*$ imply that $\text{int}(\text{gph}\,\mathcal{F}) \neq \emptyset$.

(ii) Is either of these conditions necessary to have $\text{int}(\text{gph}\,\mathcal{F}) \neq \emptyset$?

(iii) Is either of these conditions essential to have $\text{int}(\text{gph}\,\mathcal{F}) \neq \emptyset$?

Exercise 7.83 (Lower Subdifferential Optimality Conditions in the KKT Form). Formulate and prove a lower subdifferential counterpart of Corollary 7.19.

Exercise 7.84 (Coderivatives of Block-Perturbed Infinite Linear Systems). Give a detailed proof of Proposition 7.24.

Exercise 7.85 (Characterization of SSC for Block-Perturbed Linear Systems). Give a detailed proof of the equivalence (iii)⇔(iv) in Proposition 7.25. *Hint:* Consider first the case of the maximum partition $\mathcal{J} = \mathcal{J}_{\max}$, and compare it with the proof in [298, Theorem 6.1].

Exercise 7.86 (Distance Function for Maximum Partition).

(i) Given a direct proof of assertion (7.91).

(ii) Prove that SSC for $\sigma(0)$ is equivalent to the inner/lower semicontinuity of $\mathcal{F}_{\max}$ (cf. [211, Theorem 5.1]), and deduce from it the property in (7.91).

Exercise 7.87 (Characteristic Set for Infinite Convex Inequalities). Obtain the characteristic set representation for convex inequality systems in (7.98) from that in (7.81) for block-perturbed linear systems.

Exercise 7.88 (Calculation of the Coderivative Norm for Convex Systems).

(i) Give an example when the equality holds in the setting of Theorem 7.31, while the set $\bigcup_{j\in J} \text{dom}\,\varphi_j^*$ is unbounded.

(ii) Is the reflexivity of X necessary for the equalities in Theorem 7.31?

(iii) Is the reflexivity of X essential for the equalities in Theorem 7.31?

Exercise 7.89 (Coderivative Criterion for Lipschitzian Stability of Convex Systems). Formulate and prove a convex counterpart of Proposition 7.25.

Exercise 7.90 (Metric Regularity from Lipschitzian Stability for Infinite Convex Inequality Systems). Derive a characterization of metric regularity for infinite convex inequality systems from the equality formula for the exact Lipschitzian bound obtained in Theorem 7.31.

Exercise 7.91 (Optimality Conditions for SIPs with Block-Perturbed Linear Constraints). Derive upper and lower subdifferential optimality conditions for minimizing extended-real-valued function subject to the infinite linear block-perturbed inequality constraints (7.77) in Banach and Asplund spaces, respectively.

Exercise 7.92 (Necessary Optimality Conditions for SIPs with Convex Inequality Constraints). Derive upper and lower subdifferential optimality conditions for minimizing extended-real-valued function subject to the infinite convex inequality constraints (7.94) in Banach and Asplund spaces, respectively.

Exercise 7.93 (Sum Rule for ε-Subgradients of Convex Functions). Given convex functions $\varphi_1, \varphi_2\colon X \to \overline{\mathbb{R}}$ one of which is continuous at $\bar{x} \in \mathrm{dom}\,\varphi_1 \cap \mathrm{dom}\,\varphi_2$, justify the ε-subdifferential sum rule (7.102). *Hint:* Modify known proofs of the classical Moreau-Rockafellar theorem for the case of $\varepsilon > 0$ in (7.102); compare it, e.g., with the proof given in [757, Theorem 2.8.7].

Exercise 7.94 (Optimality Conditions in DC Programming). Consider the DC program defined in (7.107).

(i) Give a proof of the characterization of global minimizers in Lemma 7.33, and compare it with the one in [348].

(ii) Is the subdifferential inclusion formulated in Lemma 7.33 necessary for the local optimality of $\bar{x}$ in (7.107)?

(iii) Verify the sufficient condition for local minimizers in Lemma 7.34. *Hint:* Compare it with the proof in [235] given under the Lipschitz continuity of φ_2 around $\bar{x}$, and check that the latter assumption is equivalent to the continuity of φ_2 at $\bar{x}$.

(vi) Is the condition of Lemma 7.34 necessary for the local optimality in (7.107)?

Exercise 7.95 (Conditions for Calculating the Exact Regularity). Verify the relationships in (7.135), and show that the inclusion therein is generally strict.

Exercise 7.96 (Fenchel Conjugates for Suprema of Convex Functions).

(i) Given a direct proof of representation (7.154).

(ii) Verify formula (7.155) for the supremum of convex functions $f(x) := \sup_{t \in T} f_t(x)$. *Hint:* Deduce this, e.g., from [352, Vol. 2, Theorem 2.4.4].

Exercise 7.97 (Relationships Between CQC and FMCQ for Infinite Convex Systems). Consider the DC optimization problem (7.157), its feasible set Ξ (7.158), and the qualification conditions CQC (7.48) and FMCQ (7.159).

(i) Show that FMCQ$\Rightarrow$CQC if ϑ in (7.157) is continuous at some $x \in \Xi$.

(ii) Show that FMCQ$\Rightarrow$CQC if $\mathrm{cone}(\mathrm{dom}\,\vartheta - \Xi)$ is a closed subspace of X.

(iii) Give examples showing that CQC and FMCQ are generally independent.

Exercise 7.98 (Slater Constraint Qualification for Infinite Convex Systems). The convex inequality system $\{\vartheta_t(x) \le 0,\ t \in T \subset \mathbb{R}^m,\ x \in \mathbb{R}^n\}$ satisfies the *Slater qualification condition* (SCQ) if T is compact, the mapping $(t, x) \mapsto \vartheta_t(x)$ is continuous on $T \times \mathbb{R}^n$, and there is $x_0 \in \mathbb{R}^n$ such that $\vartheta_t(x_0) < 0$ for all $t \in T$.

(i) Show that SCQ$\Rightarrow$FMCQ if the set Ξ in (7.158) with $\Theta = \mathbb{R}^n$ is bounded.

(ii) Give an example of an infinite convex inequality system with $n = 2$ and $m = 1$ for which the converse implication in (i) is violated.

Exercise 7.99 (Conjugate Epigraphical and Subdifferential Sum Rules).

(i) Give a detailed proof of Lemma 7.49 and compare it with [131].

(ii) Construct an example showing that the subdifferential sum rule doesn't imply the epigraphical one therein.

(iii) Compare the equivalent epigraphical qualification conditions for the subdifferential sum rule given in Lemma 7.49 with other qualification conditions for this rule well recognized in convex and variational analysis in both finite and infinite dimensions; see [667, 757] and also the singular subdifferential condition (2.34) from Theorem 2.19 and Exercise 2.54(i).

Exercise 7.100 (Epigraphical Farkas Lemma).

(i) Give a detailed proof of Lemma 7.50 and compare it with [212].

(ii) Under which assumptions the weak* closure in Lemma 7.50(ii) can be replaced by the norm closure and when any closure operation can be omitted therein?

Exercise 7.101 (Epigraphs of Conjugate Functions via ε-Subdifferentials). Give a proof of representation (7.165) and compare it with [387].

Exercise 7.102 (Local Lipschitz Continuity of Convex Functions). Show that any convex function, which is continuous at some interior point of its domain, is locally Lipschitzian around this point.

Exercise 7.103 (Estimates for ε-Subgradients of Value Functions in DC SIPs). Derive a counterpart of Theorem 7.55 for ε-subgradients (1.34) of (7.173).

Exercise 7.104 (Basic Subgradients of DC Value Functions Under Extended Inner Semicontinuity). Using the definition of μ-inner semicontinuity given in Exercise 4.21, perform the following:

(i) Prove extended versions of Theorems 7.60, 7.63, and 7.64 with replacing the inner semicontinuity of the mapping $M(\cdot)$ therein by its μ-inner semicontinuity.

(ii) Construct examples showing the extensions obtained in this way are strictly better than the original formulations.

Exercise 7.105 (Closed-Graph Property of Subdifferential Mappings for l.s.c. Convex Functions on Banach Spaces).

(i) Let $\varphi\colon X \to \overline{\mathbb{R}}$ is a l.s.c. convex function on a Banach space. Prove that the graph of $x \mapsto \partial_\varepsilon\varphi(x)$ is closed in $X \times X^*$ for any $\varepsilon \geq 0$.

(ii) Show that $(x^*, y^*) \in \partial\psi(\bar{x}, \bar{y})$ in the proof of Theorem 7.63.

Exercise 7.106 (Relationships Between Subdifferential Upper Estimates for DC Value Functions). Let $\mu(\cdot)$ be the DC value function (7.173).

(i) Give an example showing that the upper estimate of $\partial\mu(\bar{x})$ from Theorem 7.60 may be better than the one in Theorem 7.63.

(ii) Investigate the possibilities to obtain upper estimates for $\partial\mu(\bar{x})$ by passing to the limit from that for regular subgradients in Theorem 7.55 in the case of Asplund (in particular, finite-dimensional) spaces and from the corresponding counterpart of Theorem 7.55 for the ε-enlargements $\widehat{\partial}_\varepsilon\mu(\cdot)$ in more general Banach space settings.

(iii) Clarify the same issues as in (ii) for the singular subdifferential $\partial^\infty\mu(\bar{x})$.

Exercise 7.107 (Lipschitz-Like Property of Feasible Solution Maps for Parameterized Versions of DC SIPs).

(i) Show that the Lipschitz-like property of the feasible solution map $x \mapsto F(x) \cap G(x)$ in the framework of Theorem 7.66 is essential for the validity of both stability and optimality conclusions of this theorem.

(ii) Based on characterizing the Lipschitz-like property of the infinite inequality systems in (7.172) obtained in Sections 7.1 and 7.3, impose appropriate assumptions on the constraint set Ω in (7.171) ensuring the feasible solution map $x \to F(x) \cap G(x)$ is Lipschitz-like at the reference point.

Exercise 7.108 (Upper Subdifferential Estimate for Value Functions in Convex SIPs). Give a detailed proof of the upper estimate of $\partial\mu(\bar{x})$ in Theorem 7.67.

Exercise 7.109 (Conjugate Epigraphical Representations). Verify representation (7.197), and show that the weak* closure can be omitted therein if one of the functions is continuous at some common point of the domains $\operatorname{dom}\varphi_i$, $i = 1, 2$. *Hint:* Compare it with the corresponding results and proofs in [116, 757].

Exercise 7.110 (Subdifferentiation of Value Functions for Convex Programs).

(i) Find appropriate qualification conditions ensuring subdifferentiability of value functions for convex programs with finitely many constraints in both finite and infinite dimensions. Are

Slater-type and subdifferential Mangasarian-Fromovitz constraint qualifications sufficient for this property?

(ii) Find appropriate qualification conditions ensuring subdifferentiability of value functions for convex SIPs in Banach spaces. *Hint:* Proceed first with dual constraint qualifications of the FMCQ and CQC types and then with primal ones of the Slater type; compare this with [210].

Exercise 7.111 (Value Functions and Optimality Conditions for Fully Convex SIPs with Upper-Level Constraints). Extend the results of Subsection 7.5.4 to bilevel SIPs with convex constraints at the upper level.

Exercise 7.112 (Comparison Between Lipschitzian and DC Approaches to Convex Bilevel Programs). Compare the necessary optimality conditions for fully convex bilevel programs containing finitely many inequality constraints that follow from Lipschitzian problems (see Theorems 6.21 and 6.23 and Exercise 6.46) with those obtained in Theorem 7.69 and Corollary 7.70 when the index set T is finite.

7.7 Commentaries to Chapter 7

Sections 7.1–7.3. *Semi-infinite programs* constitute a remarkable class of optimization problems that are intrinsically *infinite-dimensional* even in the case of linear inequality constraints on *finite-dimensional* decision variables. Their systematic study has started in the 1960s for SIPs with *linear* inequality systems and *compact* index sets being mainly motivated by applications to approximation theory, linear optimal control, and practical optimization models; see more information in [15, 298, 345] and their references. Then the study and applications have been extended to *convex* and also nonconvex while *differentiable* inequality systems over compact index sets as, e.g., in [96, 137, 394, 395, 396, 418, 442, 696, 783]. Note that the index set compactness was very essential in the obtained methods and results in these and related studies. More recently, further developments have been done for linear and convex systems with *arbitrary* index sets by using different techniques; see [139, 140, 141, 142, 210, 211, 212, 261, 299, 331, 464], among other publications. The major issues addressed in the SIP literature concerned well-posedness and ill-posedness properties, qualitative/topological and quantitative/Lipschitz-type stability analysis of parameterized feasible and optimal solution sets, necessary and sufficient optimality conditions, numerical methods, as well as various applications.

The material presented in Sections 7.1–7.3 is based on the author's joint papers with Cánovas, López, and Parra [140, 141, 142] dealing with robust Lipschitzian stability of parameterized infinite systems of linear, block-perturbed, and convex inequalities, necessary optimality conditions for minimizing nonsmooth functionals constrained by such systems, and some applications to water resource optimization. As seen above, methods and results of variational analysis and generalized differentiation presented in the previous chapters played a crucial role in these developments.

Section 7.4. This section is based on the author's joint paper with Nghia [548]. Note that, while the approach of [140] led us to complete qualitative and quantitative characterizations of the Lipschitz-like property of solution sets to linear infinite inequalities under adequate assumptions, its extension [142] to convex infinite inequalities via linear block perturbations and Fenchel duality ended up with a rather restrictive boundedness condition in the case of nonreflexive spaces; see Theorem 7.31 and Example 7.32. The latter condition was dismissed for a larger setting of perturbed infinite convex inequality and linear equality systems as a consequence of more general results on metric regularity of convex-graph multifunctions between arbitrary Banach spaces. The novel approach of [548] reduced the study of metric regularity for such mappings to the *unconstrained minimization* of *DC functions* and brought us to precise calculation of the exact regularity bounds of convex-graph multifunctions and infinite constraint systems via ε-coderivative and coderivative norms. Lemma 7.33 from global DC optimization was established by Hiriart-Urruty [348], while its local counterpart in Lemma 7.34 was obtained by Dür [235].

Corollary 7.37, summarizing the previous developments of this section, presents a major result of [548] allowing us to *precisely calculate* the *exact covering bound* of a general convex-graph multifunction between Banach spaces without additional assumptions. It implies, in particular, the regularity formula (7.149) for infinite convex constraint systems under the bounded SSC introduced in [548]. Note that another proof of (7.149) is given, in a different form under a certain uniform boundedness condition on the functions f_t, in the parallel study [373] based on the previous developments in [377] on *perfect regularity* for convex-graph multifunctions. However, there is a mistake in the proof of the aforementioned result in [373] due to the incorrect application on p. 1025 therein of the classical Sion's minimax theorem [691] whose assumptions fail to fulfill in the setting under consideration in [373].

Section 7.5. This section is mainly based on the author's joint paper with Dinh and Nghia [215] and is devoted to the subdifferentiation of the optimal value functions in DC SIPs with various applications. Note that the optimal value/marginal function for such problems is generally nonconvex, while evaluating its both basic and singular limiting subdifferentials gives us a crucial information concerning sensitivity analysis, optimality conditions, and their applications in finite and infinite dimensions. An important role in our analysis is played by the closedness qualification conditions from Definition 7.48, introduced and comprehensively studied by the same team [214] in the general LCTV space setting. In the latter paper the reader can find more discussions on the genesis of CQC and its relationships with the Farkas-Minkowski property as well as with other well-recognized constraint qualifications for finite and infinite convex systems of both primal and dual types; cf. also [116, 120, 121, 212, 213, 303, 479, 757] and the references therein. Lemma 7.49, taken from Burachik and Jeyakumar [131], provides probably the weakest conditions for the validity of the convex subdifferential sum rule in Banach spaces. Note that the equivalence between assertions (i) and (ii) in this result follows from the well-known formula (7.197). Lemma 7.50 established by Dinh et al. [212] is yet another extension of the classical Farkas lemma to infinite convex constraint systems; see the recent survey [213] on more results and discussions in this direction. Lemma 7.65 of its own interest is taken from the author's paper with Nam [532].

The last subsection of Section 7.5 implements the value function approach described in Chapter 6, together with the subdifferential results obtained above in this section, to the case of fully convex bilevel semi-infinite programs in Banach spaces indexed by arbitrary sets. Observe that in this way, we are able to significantly improve the results presented in Chapter 6, while specified to the fully convex setting, even for finitely many inequality constraints in finite-dimensional spaces.

Section 7.6. This section contains various exercises with different levels of difficulties concerning all the basic material presented in Chapter 7. As usual, we provide hints and references for the most difficult exercises. Similarly to the results of Chapter 6 on bilevel programs with finitely many constraints, relaxing the partial calmness assumption remains a challenging issue. It seems also that the pessimistic version of bilevel SIPs is Terra incognita in bilevel optimization.

Chapter 8
Nonconvex Semi-infinite Optimization

In this chapter we continue the study of SIPs in infinite-dimensional spaces while considering now problems without any convexity assumptions. A major goal is to develop effective *calculus rules* to deal with *infinite operations* (i.e., calculating normals to infinite set intersections), which are definitely of their own interest besides just applications to nonconvex SIPs. Developing various strategies in this direction, we begin with systems described by differentiable functions and then proceed with Lipschitzian and more general ones.

8.1 Optimization of Infinite Differentiable Systems

The optimization framework of our study is the class of constrained SIPs:

$$\begin{cases} \text{minimize } \ \varphi(x) \ \text{ subject to} \\ \varphi_t(x) \le 0 \ \text{ with } \ t \in T \ \text{ and } \ h(x) = 0, \end{cases} \tag{8.1}$$

where $\varphi, \varphi_t \colon X \to \overline{\mathbb{R}}$ with an arbitrary index set T, and where $h\colon X \to Y$ is a mapping between Banach spaces. Consider the *infinite system*

$$\Omega := \big\{ x \in X \big|\ h(x) = 0,\ \varphi_t(x) \le 0 \ \text{ as } \ t \in T \big\}, \tag{8.2}$$

which is the set of feasible solutions to (8.1). In this section we mainly focus on the *precise calculations*, entirely via the initial data of (8.2), of regular and basic *normal cones* to the nonconvex set Ω given by the infinite intersection under certain differentiability assumptions on φ_t and h. To achieve this major goal, we introduce new *constraint qualifications* and compare them with those studied in Chapter 7 for linear and convex SIPs as well as with conventional ones known for nonconvex differentiable systems. Then the obtained calculus results easily apply to deriving various necessary optimality conditions for SIPs (8.1) with nonsmooth objectives and differentiable constraint functions under the developed constraint qualifications.

B. S. Mordukhovich, *Variational Analysis and Applications*, Springer Monographs in Mathematics, https://doi.org/10.1007/978-3-319-92775-6_8

8.1.1 Qualification Conditions for Infinite Systems

Our standing assumptions on the data of (8.2) imposed throughout the whole section, unless otherwise stated, are as follows:

(SA) Given $\bar{x} \in \Omega$, the functions φ_t are Fréchet differentiable at $\bar{x}$ with the bounded derivative set $\{\nabla\varphi_t(\bar{x})|\ t \in T\}$ while h is strictly differentiable at $\bar{x}$.

In addition to (SA), we may also impose some stronger requirements on the inequality constraint functions φ_t that postulate a certain uniformity of their behavior with respect to the index parameter $t \in T$. We say that the functions $\{\varphi_t\}_{t\in T}$ are *uniformly Fréchet differentiable* at $\bar{x}$ if

$$s(\eta) := \sup_{t\in T} \sup_{\substack{x\in B_\eta(\bar{x}) \\ x\neq\bar{x}}} \frac{|\varphi_t(x) - \varphi_t(\bar{x}) - \langle\nabla\varphi_t(\bar{x}), x - \bar{x}\rangle|}{\|x - \bar{x}\|} \to 0 \text{ as } \eta \downarrow 0. \quad (8.3)$$

Similarly, the functions $\{\varphi_t\}_{t\in T}$ are *uniformly strictly differentiable* at $\bar{x}$ if condition (8.3) above is replaced by the stronger one as $\eta \downarrow 0$:

$$r(\eta) := \sup_{t\in T} \sup_{\substack{x,x'\in B_\eta(\bar{x}) \\ x\neq x'}} \frac{|\varphi_t(x) - \varphi_t(x') - \langle\nabla\varphi_t(\bar{x}), x - x'\rangle|}{\|x - x'\|} \to 0. \quad (8.4)$$

Easily verifiable conditions for the validity of *all* the assumptions imposed in (SA), (8.3), and (8.4) on the inequality constraint functions are as follows.

Proposition 8.1 (Uniform Differentiability Assumptions on Compact Index Sets). *Let T be a compact metric space, let φ_t in* (8.2) *be Fréchet differentiable around $\bar{x}$ for each $t \in T$, and let the mapping $(x,t) \in X \times T \mapsto \nabla\varphi_t(x) \in X^*$ be continuous on $B_\eta(\bar{x}) \times T$ for some $\eta > 0$. Then the standing assumptions* (SA) *as well as* (8.3) *and* (8.4) *are satisfied.*

Proof. It is easy to see that the standing assumptions (SA) hold since the function $t \mapsto \|\nabla\varphi_t(\bar{x})\|$ is continuous on the compact space T. Let us now show the validity of (8.4), which obviously yields (8.3). Arguing by contradiction, suppose that (8.4) fails and then find $\varepsilon > 0$, sequences $\{t_k\} \subset T$, $\{\eta_k\} \downarrow 0$, and $\{x_k\}$, $\{x'_k\} \subset B_{\eta_k}(\bar{x})$ with $x_k \neq x'_k$ such that for all large $k \in \mathbb{N}$, we have

$$\frac{|\varphi_{t_k}(x_k) - \varphi_{t_k}(x'_k) - \langle\nabla\varphi_{t_k}(\bar{x}), x_k - x'_k\rangle|}{\|x_k - x'_k\|} \geq \varepsilon - \frac{1}{k}. \quad (8.5)$$

The compactness of T gives us a subsequence of $\{t_k\}$ converging (without relabeling) to some $\bar{t} \in T$. Applying the classical mean value theorem to (8.5), we find $\theta_k \in [x_k, x'_k] := \text{co}\,\{x_k, x'_k\}$ such that

$$\frac{\varepsilon}{2} < \frac{|\langle\nabla\varphi_{t_k}(\theta_k), x_k - x'_k\rangle - \langle\nabla\varphi_{t_k}(\bar{x}), x_k - x'_k\rangle|}{\|x_k - x'_k\|} \leq \|\nabla\varphi_{t_k}(\theta_k) - \nabla\varphi_{t_k}(\bar{x})\|$$

for all large $k \in \mathbb{N}$. This contradicts the assumed continuity of the mapping $(x, t) \in X \times T \mapsto \nabla\varphi_t(x)$ on $B_\eta(\bar{x}) \times T$ and thus completes the proof. $\triangle$

The following constraint qualification condition has been well recognized in the area of nonconvex SIPs with smooth data and compact index sets.

Definition 8.2 (Extended Mangasarian-Fromovitz Constraint Qualification). *System* (8.2) *satisfies the* EXTENDED MANGASARIAN-FROMOVITZ CONSTRAINT QUALIFICATION (EMFCQ) *at* $\bar{x} \in \Omega$ *if the operator* $\nabla h(\bar{x})\colon X \to Y$ *is surjective and there is* $\widetilde{x} \in X$ *with* $\nabla h(\bar{x})\widetilde{x} = 0$ *and*

$$\langle \nabla\varphi_t(\bar{x}), \widetilde{x}\rangle < 0 \ \textit{for all} \ \ t \in T(\bar{x}) := \big\{t \in T \big|\ \varphi_t(\bar{x}) = 0\big\}. \tag{8.6}$$

If the index set T is finite, EMFCQ reduces to the classical MFCQ in nonlinear programming (NLP). Similarly to MFCQ in NLP, the main application of EMFCQ is supporting the KKT necessary conditions for local minimizers $\bar{x}$ in SIP (8.1) with $X = \mathbb{R}^n$ and $Y = \mathbb{R}^m$ given in the form: there are multipliers $\lambda_t \in \mathbb{R}^{(T)}_+$, $t \in T$, and $\mu_j \in \mathbb{R}$, $j = 1, \ldots, m$, such that

$$0 = \nabla\varphi(\bar{x}) + \sum_{t \in T(\bar{x})} \lambda_t \nabla\varphi_t(\bar{x}) + \sum_{j=1}^{n} \mu_j \nabla h_j(\bar{x}) \tag{8.7}$$

provided that T is a compact set and that the mapping $(x, t) \mapsto \nabla\varphi_t(x)$ is continuous. The following example shows that the KKT condition (8.7) may fail for nonconvex SIPs with noncompact index sets.

Example 8.3 (Violation of KKT for Nonconvex SIPs with Countable Index Sets Under EMFCQ). Consider the SIP problem (8.1) with countably many inequality constraints given by

$$\begin{cases} \text{minimize} \ \ (x_1 + 1)^2 + x_2 \ \ \text{with} \ \ (x_1, x_2) \in \mathbb{R}^2 \ \ \text{subject to} \\ x_1 + 1 \le 0, \ \ \dfrac{1}{3k} x_1^3 - x_2 \le 0 \ \ \text{for all} \ \ k \in \mathbb{N} \setminus \{1\}. \end{cases}$$

With $X := \mathbb{R}^2$, $Y := \{0\}$, $\varphi(x_1, x_2) := (x_1 + 1)^2 + x_2$, $T := \mathbb{N}$, $\varphi_1(x_1, x_2) := x_1 + 1$, and $\varphi_k(x_1, x_2) := \dfrac{1}{3k} x_1^3 - x_2$ as $k \in \mathbb{N} \setminus \{1\}$ in (8.1), observe that $\bar{x} := (-1, 0)$ is a global minimizer for this problem and that $T(\bar{x}) = \{1\}$ for the active index set (8.6). It is easy to check that EMFCQ holds at $\bar{x}$ while there is no Lagrange multiplier $\lambda \in \mathbb{R}_+$ satisfying the KKT condition (8.7) at $\bar{x}$. Indeed, we have $\langle \nabla\varphi_1(\bar{x}), (-1, 0)\rangle = -1 < 0$, which shows that the following equation doesn't admit any solution:

$$(0, 0) = \nabla\varphi(\bar{x}) + \lambda\nabla\varphi_1(\bar{x}) = (0, 1) + (\lambda, 0).$$

The next version of MFCQ for infinite systems is more appropriate for the study of (8.2) and the subsequent applications to general SIPs of type (8.1).

Definition 8.4 (Perturbed Mangasarian-Fromovitz Constraint Qualification). *Given $\bar{x} \in \Omega$, we say that system (8.2) satisfies the* PERTURBED MANGASARIAN-FROMOVITZ CONSTRAINT QUALIFICATION (PMFCQ) *at $\bar{x}$ if the derivative operator $\nabla h(\bar{x}) \colon X \to Y$ is surjective and if there is $\widetilde{x} \in X$ such that $\nabla h(\bar{x})\widetilde{x} = 0$ and that*

$$\inf_{\varepsilon>0} \sup_{t \in T_\varepsilon(\bar{x})} \langle \nabla\varphi_t(\bar{x}), \widetilde{x}\rangle < 0 \ \text{ with } \ T_\varepsilon(\bar{x}) := \big\{t \in T \,\big|\, \varphi_t(\bar{x}) \geq -\varepsilon\big\}. \tag{8.8}$$

In contrast to EMFCQ, the PMFCQ condition involves the ε-active index set $T_\varepsilon(\bar{x})$ and taking the infimum over $\varepsilon > 0$ in (8.8); that is where the name "perturbed" comes from. Since $T(\bar{x}) \subset T_\varepsilon(\bar{x})$ for all $\varepsilon > 0$, PMFCQ is stronger than EMFCQ while being, as we'll see below, much more appropriate for applications to SIPs with arbitrary (including compact) index sets.

Let us present some assumptions on the initial data of (8.2) ensuring the equivalence between PMFCQ and EMFCQ.

Proposition 8.5 (PMFCQ from EMFCQ). *Let T be a compact metric space, and let $\bar{x} \in \Omega$ in (8.2). Assume that the function $t \mapsto \varphi_t(\bar{x})$ is u.s.c. on T, that the derivative operator $\nabla h(\bar{x}) \colon X \to Y$ is surjective, and that there is $\widetilde{x} \in X$ with the following properties: $\nabla h(\bar{x})\widetilde{x} = 0$, the function $t \mapsto \langle \nabla\varphi_t(\bar{x}), \widetilde{x}\rangle$ is u.s.c. on T, and $\langle \nabla\varphi_t(\bar{x}), \widetilde{x}\rangle < 0$ for all $t \in T(\bar{x})$. Then PMFCQ holds at $\bar{x}$ and is equivalent to EMFCQ at this point.*

Proof. Arguing by contradiction, suppose that PMFCQ fails at $\bar{x}$. Then it follows from (8.8) that there exist sequences $\varepsilon_k \downarrow 0$ and $\{t_k\} \subset T$ such that

$$t_k \in T_{\varepsilon_k}(\bar{x}) \ \text{ and } \ \langle \nabla\varphi_{t_k}(\bar{x}), \widetilde{x}\rangle \geq -k^{-1} \ \text{ for all } \ k \in \mathbb{N}.$$

Since T is a compact metric space, we find a subsequence of t_k (no relabeling) converging to some $\bar{t} \in T$. Observe that the upper semicontinuity assumptions imposed in the proposition imply that

$$\varphi_{\bar{t}}(\bar{x}) \geq \limsup_{k\to\infty} \varphi_{t_k}(\bar{x}) = 0 \ \text{ and } \langle \nabla\varphi_{\bar{t}}(\bar{x}), \widetilde{x}\rangle \geq \limsup_{k\to\infty} \langle \nabla\varphi_{t_k}(\bar{x}), \widetilde{x}\rangle \geq 0.$$

Thus we get $\bar{t} \in T(\bar{x})$ and $\langle \nabla\varphi_{\bar{t}}(\bar{x}), \widetilde{x}\rangle \geq 0$, a contradiction. △

The following example shows that EMFCQ doesn't imply PMFCQ even for simple frameworks of nonconvex SIPs in $\mathbb{R}^2$ with compact index sets.

Example 8.6 (EMFCQ Doesn't Imply PMFCQ for Infinite Systems with Compact Index Sets). Let $X = \mathbb{R}^2$, $T = [0, 1]$ in (8.2) with $h = 0$ and

$$\varphi_0(x) := x_1 + 1 \leq 0, \quad \varphi_t(x) := t x_1 - x_2^3 \leq 0 \ \text{ for } \ t \in T \setminus \{0\}.$$

It is easy to check that the functions $\varphi_t(x)$, $t \in T$, satisfy (SA) and (8.4) at $\bar{x} = (-1, 0)$. Observe further that $T(\bar{x}) = \{0\}$, that $T_\varepsilon(\bar{x}) = [0, \varepsilon]$ as $\varepsilon \in (0, 1)$, and that EMFCQ holds at $\bar{x}$. However, for any $d = (d_1, d_2) \in \mathbb{R}^2$, we have

$$\inf_{\varepsilon>0} \sup_{t\in T_\varepsilon(\bar{x})} \langle \nabla\varphi_t(\bar{x}), d\rangle = \inf_{\varepsilon>0} \sup \Big\{ d_1, \sup \big\{ td_1 \big| \, t \in (0,\varepsilon] \big\} \Big\} \ge 0,$$

which shows that PMFCQ fails at $\bar{x}$. Note that the u.s.c. assumptions with respect to t in Propositions 8.5 don't hold in this example.

It has been well recognized in NLP theory (i.e., when T is a finite set in (8.2)) that MFCQ is equivalent to the Slater constraint qualification (SCQ) if the functions φ_t are smooth and convex while h is linear. The next proposition shows that a similar equivalence holds for SIPs with replacing MFCQ by PMFCQ and SCQ by the strong Slater condition (SSC) used in Chapter 7 in the case of infinite inequality constraint systems. This result *alone* indicates that PMFCQ, rather than EMFCQ, is the most natural extension of SSC to nonconvex infinite systems with arbitrary index sets.

Proposition 8.7 (Equivalence Between PMFCQ and SSC for Differentiable Convex Systems). *Assume that in* (8.2) *all the functions $\varphi_t(x)$ are convex and uniformly Fréchet differentiable at $\bar{x}$ and that $h(x) := Ax$ is a surjective continuous linear operator. Then PMFCQ is equivalent to the following* STRONG SLATER CONDITION (SSC) *for* (8.2)*: there is $\widehat{x} \in X$ such that $A\widehat{x} = 0$ and $\sup_{t\in T} \varphi_t(\widehat{x}) < 0$.*

Proof. Suppose first that SSC holds at $\bar{x}$, i.e., there are $\widehat{x} \in X$ and $\delta > 0$ such that $A\widehat{x} = 0$ and $\varphi_t(\widehat{x}) < -2\delta$ for all $t \in T$. This implies, together with the imposed assumptions, that for each $\varepsilon \in (0,\delta)$ and $t \in T_\varepsilon(\bar{x})$ we have

$$\langle \nabla\varphi_t(\bar{x}), \widehat{x} - \bar{x}\rangle \le \varphi_t(\widehat{x}) - \varphi_t(\bar{x}) \le -2\delta + \varepsilon \le -\delta.$$

Define further $\widetilde{x} := \widehat{x} - \bar{x}$ and get $A\widetilde{x} = A\widehat{x} - A\bar{x} = 0$ with $\langle \nabla\varphi_t(\bar{x}), \widetilde{x}\rangle \le -\delta$ for all $t \in T_\varepsilon(\bar{x})$ and $\varepsilon \in (0,\delta)$. This clearly yields PMFCQ.

Conversely, suppose that PMFCQ holds at $\bar{x}$. Then there are $\varepsilon, \eta > 0$ and $\widetilde{x} \in X$ such that $\langle \nabla\varphi_t(\bar{x}), \widetilde{x}\rangle \le -\eta$ for all $t \in T_\varepsilon(\bar{x})$ and that $A\widetilde{x} = 0$. It follows from the imposed uniform Fréchet differentiability (8.3) of φ_t at $\bar{x}$ with using the function $s(\cdot)$ defined therein that for each $\lambda > 0$ we have

$$\varphi_t(\bar{x} + \lambda\widetilde{x}) \le \varphi_t(\bar{x}) + \lambda\langle \nabla\varphi_t(\bar{x}), \widetilde{x}\rangle + \lambda\|\widetilde{x}\| s\big(\lambda\|\widetilde{x}\|\big),$$

which readily implies that $\varphi_t(\bar{x} + \lambda\widetilde{x}) \le \lambda\big(-\eta + \|\widetilde{x}\| s(\lambda\|\widetilde{x}\|)\big)$ as $t \in T_\varepsilon(\bar{x})$. For $t \notin T_\varepsilon(\bar{x})$ it follows from the above that

$$\varphi_t(\bar{x} + \lambda\widetilde{x}) \le -\varepsilon + \lambda \sup_{\tau\in T} \|\nabla\varphi_\tau(\bar{x})\| \cdot \|\widetilde{x}\| + \lambda\|\widetilde{x}\| s\big(\lambda\|\widetilde{x}\|\big),$$

which yields the existence of $\lambda_0 > 0$ so small that $\sup_{t\in T} \varphi_t(\widehat{x}) < 0$ with $\widehat{x} := \bar{x} + \lambda_0\widetilde{x}$. Furthermore, it is easy to see that $A\widehat{x} = 0$. This justifies SSC at $\widehat{x}$ and hence completes the proof of the proposition. △

Next we introduce yet another qualification condition for the nonlinear constraint system (8.2) the versions of which have been exploited in Chapter 7 for linear and convex infinite inequality systems.

Definition 8.8 (Nonlinear Farkas-Minkowski Constraint Qualification). *Given $\bar{x} \in \Omega$ in (8.2) with $h \equiv 0$, we say that the* NONLINEAR FARKAS-MINKOWSKI CONSTRAINT QUALIFICATION (NFMCQ) *holds at $\bar{x}$ if the cone*

$$\text{cone}\left\{\big(\nabla\varphi_t(\bar{x}), \langle\nabla\varphi_t(\bar{x}), \bar{x}\rangle - \varphi_t(\bar{x})\big)\big|\ t \in T\right\}$$

is weak closed in the product space $X^* \times \mathbb{R}$.*

Let us compare the introduced NFMCQ with the other constrained qualifications discussed in this section in the case of infinite inequality constraints.

Proposition 8.9 (Sufficient Conditions for NFMCQ). *Consider the infinite system (8.2) with $h \equiv 0$ therein. Then NFMCQ is satisfied at $\bar{x} \in \Omega$ in each of the following three settings:*

(i) *The index T is finite and MFCQ holds at $\bar{x}$.*

(ii) $\dim X < \infty$, *the set $\{(\nabla\varphi_t(\bar{x}), \langle\nabla\varphi_t(\bar{x}), \bar{x}\rangle - \varphi_t(\bar{x}))|\ t \in T\}$ is compact, and PMFCQ is fulfilled at $\bar{x}$.*

(iii) *The index T is a compact metric space,* $\dim X < \infty$, *the mappings $t \in T \mapsto \varphi_t(\bar{x})$ and $t \in T \mapsto \nabla\varphi_t(\bar{x})$ are continuous, and EMFCQ holds at $\bar{x}$.*

Proof. Define $\widetilde{\varphi}_t(x) := \langle\nabla\varphi_t(\bar{x}), x - \bar{x}\rangle + \varphi_t(\bar{x})$ for all $x \in X$. To verify (i), suppose that T is finite and that MFCQ holds at $\bar{x}$ for the inequality system in (8.2). It is clear that the functions $\widetilde{\varphi}_t$ also satisfy MFCQ at $\bar{x}$. Since these functions are linear, we observe from Proposition 8.7 that there is $\widehat{x} \in X$ such that $\widetilde{\varphi}_t(\widehat{x}) = \langle\nabla\varphi_t(\bar{x}), \widehat{x} - \bar{x}\rangle + \varphi_t(\bar{x}) < 0$ for all $t \in T$. It is not hard to check that the latter condition yields the validity of FMCQ at $\bar{x}$.

Considering next case (ii) with $X = \mathbb{R}^d$, suppose that PMFCQ holds at $\bar{x}$ and the set $\{(\nabla\varphi_t(\bar{x}), \langle\nabla\varphi_t(\bar{x}), \bar{x}\rangle - \varphi_t(\bar{x}))|\ t \in T\}$ is compact in $\mathbb{R}^d \times \mathbb{R}$. Note that the functions $\widetilde{\varphi}_t$ defined above also satisfy PMFCQ at $\bar{x}$ and then apply Proposition 8.7 to these functions. This gives us $\widehat{x} \in \mathbb{R}^d$ satisfying

$$\sup_{t\in T}\widetilde{\varphi}_t(\widehat{x}) = \sup_{t\in T}\left\{\langle\nabla\varphi_t(\bar{x}), \widehat{x} - \bar{x}\rangle + \varphi_t(\bar{x})\right\} < 0. \tag{8.9}$$

We now claim that $(0, 0) \notin \text{co}\{(\nabla\varphi_t(\bar{x}), \langle\nabla\varphi_t(\bar{x}), \bar{x}\rangle - \varphi_t(\bar{x}))|\ t \in T\}$. Indeed, otherwise ensures the existence of $\lambda \in \mathbb{R}^{(T)}_+$ with $\sum_{t\in T}\lambda_t = 1$ such that

$$(0, 0) = \sum_{t\in T}\lambda_t\big(\nabla\varphi_t(\bar{x}), \langle\nabla\varphi_t(\bar{x}), \bar{x}\rangle - \varphi_t(\bar{x})\big).$$

Combining the latter with (8.9) tells us that

$$0 = \sum_{t\in T}\lambda_t\langle\nabla\varphi_t(\bar{x}), \widehat{x}\rangle - \sum_{t\in T}\lambda_t\big(\langle\nabla\varphi_t(\bar{x}), \bar{x}\rangle - \varphi_t(\bar{x})\big) \le \sup_{t\in T}\widetilde{\varphi}_t(\widehat{x}) < 0,$$

a contradiction. Based on the claimed condition and the result of Exercise 8.89, we show that the convex conic hull

$$\text{cone}\left\{(\nabla\varphi_t(\bar{x}), \langle\nabla\varphi_t(\bar{x}), \bar{x}\rangle - \varphi_t(\bar{x}))\big|\ t \in T\right\}\ \text{ is closed in }\ \mathbb{R}^{d+1},$$

which justifies (ii). Finally, (iii) follows from (ii) due to Proposition 8.5. △

We conclude this subsection by showing NFMCQ and PMFCQ are generally *independent* for infinite inequality systems even in finite dimensions.

Example 8.10 (Independence of NFMCQ and PMFCQ). It is easy to check that for the constraint inequality system from Example 8.6, we have NFMCQ satisfied at $\bar{x} = (-1, 0)$ since the corresponding conic hull

$$\begin{aligned}&\text{cone}\left\{(\nabla\varphi_t(\bar{x}), \langle\nabla\varphi_t(\bar{x}), \bar{x}\rangle - \varphi_t(\bar{x}))\,\middle|\, t \in T\right\}\\&= \text{cone}\left\{(1, 0, -1), (t, 0, 0)\,\middle|\, t \in (0, 1]\right\}\end{aligned}$$

is closed in $\mathbb{R}^3$. On the other hand, Example 8.6 shows that PMFCQ doesn't hold for this infinite inequality system at $\bar{x}$.

To demonstrate that NFMCQ doesn't generally follow from PMFCQ (and even from EMFCQ), consider the countable system in $\mathbb{R}^2$ from Example 8.3. When $\bar{x} = (-1, 0)$, we get $T_\varepsilon(\bar{x}) = \{k \in \mathbb{N} \setminus \{1\} |\ k \geq (3\varepsilon)^{-1}\} \cup \{1\}$ for the perturbed active index set in (8.8). It shows that PMFCQ and hence EMFCQ hold at $\bar{x}$. On the other hand, the convex conic hull

$$\begin{aligned}&\text{cone}\left\{(\nabla\varphi_t(\bar{x}), \langle\nabla\varphi_t(\bar{x}), \bar{x}\rangle - \varphi_t(\bar{x}))\,\middle|\, t \in T\right\}\\&= \text{cone}\left[(1, 0, -1) \cup \left\{\left(\frac{1}{k}, -1, -\frac{2}{3k}\right)\,\middle|\, k \neq 1\right\}\right]\end{aligned}$$

is not closed in $\mathbb{R}^3$; see Fig. 8.1; i.e., NFMCQ is not satisfied at this point.

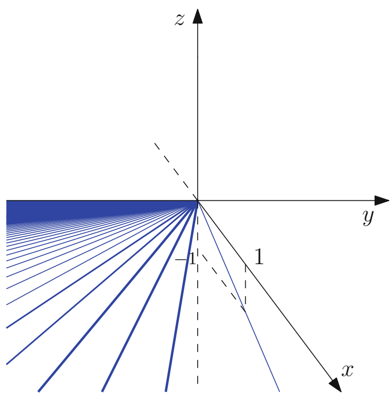


Fig. 8.1 Failure of NFMCQ.

8.1.2 Normal Cones to Nonconvex Infinite Constraint Sets

This subsection is mainly devoted to *precise calculating* both regular and basic normal cones to the constraint set Ω in (8.2), which is given via the *infinite intersection* of nonconvex sets. Based on the constraint qualification conditions discussed in Subsection 8.1.1, we derive various calculation formulas for regular and basic normals to the set Ω entirely in terms of its initial data of (8.2) in arbitrary Banach spaces.

Let us first present the following useful result from functional analysis.

Lemma 8.11 (Weak* Closed Images of Adjoint Operators). *Let $A : X \to Y$ be a surjective continuous linear operator. Then the image of its adjoint operator A^*Y^* is a weak* closed subspace of X^*.*

Proof. Define $C := A^*Y^* \subset X^*$ and pick any $k \in \mathbb{N}$. We claim that the set $A_k := C \cap k\mathbb{B}^*$ is weak* closed in X^*. Considering a net $\{x_\nu^*\}_{\nu\in\mathcal{N}} \subset A_k$ weak* converging to $x^* \in X^*$ and taking into account that the dual ball $\mathbb{B}^*$ is weak* compact in X^* give us $x^* \in k\mathbb{B}^*$. The above construction shows that there is a net $\{y_\nu^*\}_{\nu\in\mathcal{N}} \subset Y^*$ satisfying $x_\nu^* = A^*y_\nu^*$ whenever $\nu \in \mathcal{N}$. Furthermore, it follows from the surjectivity of A that

$$\|x_\nu^*\| = \|A^*y_\nu^*\| \geq \kappa\|y_\nu^*\| \ \text{ for all } \ \nu \in \mathbb{N},$$

where $\kappa := \inf\{\|A^*y^*\| \text{ over } \|y^*\| = 1\} \in (0,\infty)$; see Exercise 1.53. Hence $\|y_\nu^*\| \leq k\kappa^{-1}$ for all $\nu \in \mathcal{N}$. By passing to a subnet, suppose that y_ν^* weak* converges to some element $y^* \in Y^*$ for which $x^* = A^*y^* \in A_k$. This thus verifies that the set $A_k = C \cap k\mathbb{B}_{X^*}$ is weak* closed for all $k \in \mathbb{N}$. The classical Banach-Dieudonné-Krein-Šmulian theorem yields therefore that the image set $C = A^*Y^*$ is weak* closed in X^*. △

Now we are ready to establish the main result of this subsection.

Theorem 8.12 (Regular and Basic Normals to Infinite Systems). *Let $\bar{x} \in \Omega$ for the infinite constraint set (8.2), and let PMFCQ hold at $\bar{x}$ under the validity of (8.3). Then the regular normal cone to Ω at $\bar{x}$ is calculated by*

$$\widehat{N}(\bar{x};\Omega) = \bigcap_{\varepsilon>0} \text{cl}^*\text{cone}\left\{\nabla\varphi_t(\bar{x})\,\middle|\, t \in T_\varepsilon(\bar{x})\right\} + \nabla h(\bar{x})^*Y^*. \tag{8.10}$$

If furthermore the functions φ_t, $t \in T$, satisfy (8.4), then the basic normal cone to Ω at $\bar{x}$ is calculated by the same formula

$$N(\bar{x};\Omega) = \bigcap_{\varepsilon>0} \text{cl}^*\text{cone}\left\{\nabla\varphi_t(\bar{x})\,\middle|\, t \in T_\varepsilon(\bar{x})\right\} + \nabla h(\bar{x})^*Y^*. \tag{8.11}$$

Proof. To verify first the inclusion "$\supset$" in (8.10), observe from the definition of PMFCQ that there are $\widetilde{\varepsilon} > 0$, $\delta > 0$, and $\widetilde{x} \in X$ such that $\nabla h(\bar{x})\widetilde{x} = 0$ and

$$\sup_{t\in T_\varepsilon(\bar{x})} \langle\nabla\varphi_t(\bar{x}), \widetilde{x}\rangle < -\delta \ \text{ for all } \varepsilon \leq \widetilde{\varepsilon}. \tag{8.12}$$

Fix $\varepsilon \in (0,\widetilde{\varepsilon})$, pick x^* from the set on the right-hand side of (8.10), and then find a net $(\lambda_\nu)_{\nu\in\mathcal{N}} \subset \mathbb{R}_+^{(T)}$ and a dual element $y^* \in Y^*$ satisfying

$$x^* = w^* - \lim_{\nu \in \mathcal{N}} \sum_{t \in T_\varepsilon(\bar{x})} \lambda_{t\nu} \nabla \varphi_t(\bar{x}) + \nabla h(\bar{x})^* y^*.$$

Combining this with (8.12) gives us the estimate

$$\langle x^*, \widetilde{x} \rangle = \lim_{\nu \in \mathcal{N}} \sum_{t \in T_\varepsilon(\bar{x})} \lambda_{t\nu} \langle \nabla \varphi_t(\bar{x}), \widetilde{x} \rangle + \langle \nabla h(\bar{x})^* y^*, \widetilde{x} \rangle \le -\delta \limsup_{\nu \in \mathcal{N}} \sum_{t \in T_\varepsilon(\bar{x})} \lambda_{t\nu}.$$

It follows further that for each $\eta > 0$ and $x \in \Omega \cap B_\eta(\bar{x})$, we have

$$\begin{aligned}
&\langle x^*, x - \bar{x} \rangle = \lim_{\nu \in \mathcal{N}} \sum_{t \in T_\varepsilon(\bar{x})} \lambda_{t\nu} \langle \nabla \varphi_t(\bar{x}), x - \bar{x} \rangle + \langle \nabla h(\bar{x})^* y^*, x - \bar{x} \rangle \\
&\le \limsup_{\nu \in \mathcal{N}} \sum_{t \in T_\varepsilon(\bar{x})} \lambda_{t\nu} \Big(\varphi_t(x) - \varphi_t(\bar{x}) + \|x - \bar{x}\| s(\eta) \Big) + \langle y^*, \nabla h(\bar{x})(x - \bar{x}) \rangle \\
&\le \limsup_{\nu \in \mathcal{N}} \sum_{t \in T_\varepsilon(\bar{x})} \lambda_{t\nu} \Big(\varepsilon + \|x - \bar{x}\| s(\eta) \Big) + \|y^*\| \Big(\|h(x) - h(\bar{x})\| + o(\|x - \bar{x}\|) \Big) \\
&\le \Big(\varepsilon + \|x - \bar{x}\| s(\eta) \Big) \limsup_{\nu \in \mathcal{N}} \sum_{t \in T_\varepsilon(\bar{x})} \lambda_{t\nu} + \|y^*\| o(\|x - \bar{x}\|).
\end{aligned}$$

Taking now the above estimate of $\langle x^*, \widetilde{x} \rangle$ into account yields

$$\langle x^*, x - \bar{x} \rangle \le -\frac{\langle x^*, \widetilde{x} \rangle}{\delta} \Big(\varepsilon + \|x - \bar{x}\| s(\eta) \Big) + o(\|x - \bar{x}\|) \|y^*\|,$$

which implies in turn due to $\varepsilon, \eta \downarrow 0$ that

$$\limsup_{x \xrightarrow{\Omega} \bar{x}} \frac{\langle x^*, x - \bar{x} \rangle}{\|x - \bar{x}\|} \le 0,$$

which means that $x^* \in \widehat{N}(\bar{x}; \Omega)$ and thus justifies the inclusion "$\supset$" in (8.10).

Next we prove the inclusion "$\subset$" in (8.11) under the assumption that φ_t are uniformly strictly differentiable at $\bar{x}$. This immediately implies the one "$\subset$" in (8.10) under the latter assumption, while we note that similar arguments justify the inclusion "$\subset$" in (8.10) under merely the uniform Fréchet differentiability of φ_t at $\bar{x}$. To proceed with proving "$\subset$" in (8.11), define

$$A_\varepsilon := \mathrm{cl}^* \mathrm{cone} \big\{ \nabla \varphi_t(\bar{x}) \big|\ t \in T_\varepsilon(\bar{x}) \big\} + \nabla h(\bar{x})^* Y^* \quad \text{for} \quad \varepsilon > 0.$$

Arguing by contradiction, pick $x^* \in N(\bar{x}; \Omega) \setminus \{0\}$ and suppose that $x^* \notin A_\varepsilon$ for some $\varepsilon \in (0, \widetilde{\varepsilon})$. We first claim that the set A_ε is weak* closed in X^* for all $\varepsilon \le \widetilde{\varepsilon}$ by showing that $\mathrm{cl}^* B_\varepsilon \subset A_\varepsilon$, where

$$B_\varepsilon := \mathrm{cone} \big\{ \nabla \varphi_t(\bar{x}) \big|\ t \in T_\varepsilon(\bar{x}) \big\} + \nabla h(\bar{x})^* (Y^*).$$

To justify it, pick any $u^* \in \mathrm{cl}^* B_\varepsilon$ and therefore find some nets $(\lambda_\nu)_{\nu \in \mathbb{N}} \subset \mathbb{R}_+^{(T)}$ and $(y_\nu^*)_{\nu \in \mathbb{N}} \subset Y^*$ for which

$$u_\nu^* = \sum_{t\in T_\varepsilon(\bar{x})} \lambda_{t\nu}\nabla\varphi_t(\bar{x}) + \nabla h(\bar{x})^* y_\nu^* \xrightarrow{w^*} u^*.$$

Similarly to the proof of the estimate for $\langle x^*, \widetilde{x}\rangle$ above, we derive the inequality

$$\langle u^*, \widetilde{x}\rangle \le -\delta \limsup_{\nu\in\mathcal{N}} \sum_{t\in T_\varepsilon(\bar{x})} \lambda_{t\nu}$$

and get furthermore the dual norm estimate

$$\|u_\nu^* - \nabla h(\bar{x})^* y_\nu^*\| = \|\sum_{t\in T_\varepsilon(\bar{x})} \lambda_{t\nu}\nabla\varphi_t(\bar{x})\| \le \sup_{\tau\in T_\varepsilon(\bar{x})} \|\nabla\varphi_\tau(\bar{x})\| \sum_{t\in T_\varepsilon(\bar{x})} \lambda_{t\nu},$$

which verifies the boundedness of the net $\{u_\nu^* - \nabla h(\bar{x})^* y_\nu^*\}_{\nu\in\mathcal{N}}$ in X^*. The Alaoglu-Bourbaki theorem tells us that there is a subnet of $\{u_\nu^* - \nabla h(\bar{x})^* y_\nu^*\}$ (without relabeling) weak* converging to some $v^* \in \mathrm{cl}^*\mathrm{cone}\big\{\nabla g_t(\bar{x})\big|\ t \in T_\varepsilon(\bar{x})\big\}$. Thus the net $\{\nabla h(\bar{x})^* y_\nu^*\}$ weak* converges to $u^* - v^*$. By Lemma 8.11 we find $y^* \in Y^*$ such that $u^* - v^* = \nabla h(\bar{x})^* y^*$, which yields that $u^* = v^* + \nabla h(\bar{x})^* y^* \in A_\varepsilon$ and hence ensures that A_ε is weak* closed in X^*. Since $x^* \notin A_\varepsilon$, we deduce from the classical separation theorem that there are $x_0 \in X$ and $c > 0$ satisfying the inequalities

$$\langle x^*, x_0\rangle \ge 2c > 0 \ge \langle \nabla\varphi_t(\bar{x}), x_0\rangle + \langle y^*, \nabla h(\bar{x})x_0\rangle \tag{8.13}$$

for all $t \in T_\varepsilon(\bar{x})$ and $y^* \in Y^*$; hence $\nabla h(\bar{x})x_0 = 0$. Define further

$$\widehat{x} := x_0 + \frac{c}{\|x^*\|\cdot\|\widetilde{x}\|}\widetilde{x}$$

and observe that $\nabla h(\bar{x})\widehat{x} = 0$. It follows from (8.13) and PMFCQ that

$$\langle x^*, \widehat{x}\rangle = \langle x^*, x_0 + \frac{c}{\|x^*\|\cdot\|\widetilde{x}\|}\widetilde{x}\rangle \ge 2c + \frac{c}{\|x^*\|\cdot\|\widetilde{x}\|}\langle x^*, \widetilde{x}\rangle \ge c, \tag{8.14}$$

$$\langle \nabla\varphi_t(\bar{x}), \widehat{x}\rangle = \langle \nabla\varphi_t(\bar{x}), x_0\rangle + \frac{c}{\|x^*\|\cdot\|\widetilde{x}\|}\langle \nabla\varphi_t(\bar{x}), \widetilde{x}\rangle \le -\widetilde{\delta} \tag{8.15}$$

for all $t \in T_\varepsilon(\bar{x})$ with $\widetilde{\delta} := \delta c(\|x^*\|\cdot\|\widetilde{x}\|)^{-1} > 0$. Noting that $\widehat{x} \ne 0$ by (8.15), suppose without loss of generality that $\|\widehat{x}\| = 1$. Furthermore, we get from construction (1.58) of basic normals in Banach spaces that there are sequences $\varepsilon_k \downarrow 0$, $\eta_k \downarrow 0$, $x_k \xrightarrow{\Omega} \bar{x}$, and $x_k^* \xrightarrow{w^*} x^*$ as $k \to \infty$ with

$$\langle x_k^*, x - x_k\rangle \le \varepsilon_k\|x - x_k\| \ \text{ for all } \ x \in B_{\eta_k}(x_k) \cap \Omega, \quad k \in \mathbb{N}. \tag{8.16}$$

Since the mapping h is strictly differentiable at $\bar{x}$ with the surjective derivative $\nabla h(\bar{x})$, it follows from the Lyusternik-Graves theorem held in general Banach spaces (see Corollary 3.8 and the corresponding commentaries in Section 3.5) that h is *metrically regular* around $\bar{x}$, i.e., there are neighborhoods U of $\bar{x}$ and V of

$0 = h(\bar{x})$ and a constant $\mu > 0$ such that

$$\operatorname{dist}\big(x; h^{-1}(y)\big) := \inf\big\{\|x - z\| \;\big|\; z \in h^{-1}(y)\big\} \le \mu\|y - h(x)\|$$

for all $x \in U$ and $y \in V$. Using $h(x_k) = 0$ and $\nabla h(\bar{x})\widehat{x} = 0$, we have

$$\|h(x_k + t\widehat{x})\| = \|h(x_k + t\widehat{x}) - h(x_k) - \nabla h(\bar{x})(t\widehat{x})\| = o(t) \;\text{ for small }\; t > 0.$$

Using metric regularity, for any small $t > 0$, we get $x_t \in h^{-1}(0)$ with $\|x_k + t\widehat{x} - x_t\| = o(t)$ if $x_k \in U$. This allows us to find $\widetilde{\eta}_k < \eta_k$ and $\widetilde{x}_k := x_{\widetilde{\eta}_k} \in h^{-1}(0)$ satisfying $\widetilde{\eta}_k + o(\widetilde{\eta}_k) \le \eta_k$ and $\|x_k + \widetilde{\eta}_k\widehat{x} - \widetilde{x}_k\| = o(\widetilde{\eta}_k)$. Note that

$$\begin{aligned}\|x_k - \widetilde{x}_k\| &\le \widetilde{\eta}_k\|\widehat{x}\| + \|x_k + \widetilde{\eta}_k\widehat{x} - \widetilde{x}_k\| = \widetilde{\eta}_k + o(\widetilde{\eta}_k) \le \eta_k,\\ \|x_k - \widetilde{x}_k\| &\ge \widetilde{\eta}_k\|\widehat{x}\| - \|x_n + \widetilde{\eta}_k\widehat{x} - \widetilde{x}_k\| = \widetilde{\eta}_k - o(\widetilde{\eta}_k).\end{aligned}$$

By the classical uniform boundedness principle , there is a constant M such that $M > \|x_k^*\|$ for all $k \in \mathbb{N}$ due to $x_k^* \xrightarrow{w^*} x^*$ as $k \to \infty$. It follows from (8.14) that $\langle x_k^*, \widehat{x}\rangle > 0$ for large $k \in \mathbb{N}$. Then we have

$$\begin{aligned}\frac{\langle x_k^*, \widetilde{x}_k - x_k\rangle}{\|\widetilde{x}_k - x_k\|} &= \frac{\langle x_k^*, \widetilde{x}_k - \widetilde{\eta}_k\widehat{x} - x_k\rangle}{\|\widetilde{x}_k - x_k\|} + \frac{\langle x_k^*, \widetilde{\eta}_k\widehat{x}\rangle}{\|\widetilde{x}_k - x_k\|}\\ &\ge -M\frac{\|\widetilde{x}_k - \widetilde{\eta}_k\widehat{x} - x_k\|}{\|\widetilde{x}_k - x_k\|} + \widetilde{\eta}_k\frac{\langle x_k^*, \widehat{x}\rangle}{\|\widetilde{x}_k - x_k\|}\\ &\ge -M\frac{o(\widetilde{\eta}_k)}{\widetilde{\eta}_k - o(\widetilde{\eta}_k)} + \frac{\widetilde{\eta}_k}{\widetilde{\eta}_k + o(\widetilde{\eta}_k)}\langle x_k^*, \widehat{x}\rangle.\end{aligned}$$

Passing now to the limit as $k \to \infty$ and using $o(\widetilde{\eta}_k)/\widetilde{\eta}_k \to 0$ yields

$$\liminf_{k\to\infty}\frac{\langle x_k^*, \widetilde{x}_k - x_k\rangle}{\|\widetilde{x}_k - x_k\|} \ge \langle x^*, \widehat{x}\rangle,$$

which shows that $\widetilde{x}_k \notin \Omega$ for large $k \in \mathbb{N}$ by (8.14) and (8.16).

Define now $u_k := x_k + \widetilde{\eta}_k\widehat{x} - \widetilde{x}_k$ and get $\|u_k\| = o(\widetilde{\eta}_k)$ with $\|\widetilde{x}_k + u_k - x_k\| = \widetilde{\eta}_k$ by the arguments above. It follows from (SA), (8.4), and (8.15) that for each $t \in T_\varepsilon(\bar{x})$, we have the relationships

$$\begin{aligned}-\widetilde{\delta} \ge \frac{\langle\nabla\varphi_t(\bar{x}), \widetilde{\eta}_k\widehat{x}\rangle}{\widetilde{\eta}_k} &= \frac{\langle\nabla\varphi_t(\bar{x}), \widetilde{x}_k - x_k\rangle}{\|\widetilde{x}_k + u_k - x_k\|} + \frac{\langle\nabla\varphi_t(\bar{x}), u_k\rangle}{\|\widetilde{x}_k + u_k - x_k\|}\\ &\ge \frac{\langle\nabla\varphi_t(\bar{x}), \widetilde{x}_k - x_k\rangle}{\|\widetilde{x}_k - x_k\|}\frac{\|\widetilde{x}_k - x_k\|}{\|\widetilde{x}_k + u_k - x_k\|} + \frac{\langle\nabla\varphi_t(\bar{x}), u_k\rangle}{\|\widetilde{x}_k + u_k - x_k\|}\\ &\ge \Big(\frac{\varphi_t(\widetilde{x}_k) - \varphi_t(x_k)}{\|\widetilde{x}_k - x_k\|} - r(\widehat{\eta}_k)\Big)\frac{\|\widetilde{x}_k - x_k\|}{\|\widetilde{x}_k + u_k - x_k\|} - \sup_{\tau\in T_\varepsilon(\bar{x})}\|\nabla\varphi_\tau(\bar{x})\|\frac{o(\widetilde{\eta}_k)}{\widetilde{\eta}_k}\\ &\ge \Big(\frac{\varphi_t(\widetilde{x}_k)}{\|\widetilde{x}_k - x_k\|} - r(\widehat{\eta}_k)\Big)\frac{\|\widetilde{x}_k - x_k\|}{\|\widetilde{x}_k + u_k - x_k\|} - \sup_{\tau\in T}\|\nabla\varphi_\tau(\bar{x})\|\frac{o(\widetilde{\eta}_k)}{\widetilde{\eta}_k},\end{aligned}$$

where $\widehat{\eta}_k := \max\{\|x_k - \bar{x}\| \text{ and } \|\widetilde{x}_k - \bar{x}\|\} \to 0$ as $k \to \infty$. Note that

$$\frac{\widetilde{\eta}_k - o(\widetilde{\eta}_k)}{\widetilde{\eta}_k} \le \frac{\|\widetilde{x}_k - x_k\|}{\|\widetilde{x}_k + u_k - x_k\|} \le \frac{\widetilde{\eta}_k + o(\widetilde{\eta}_k)}{\widetilde{\eta}_k}, \quad \text{and so} \quad \frac{\|\widetilde{x}_k - x_k\|}{\|\widetilde{x}_k + u_k - x_k\|} \to 1$$

as $k \to \infty$. Furthermore, since $r(\widehat{\eta}_k) \to 0$ and $o(\widetilde{\eta}_k/)\widetilde{\eta}_k \to 0$, we have $\varphi_t(\widetilde{x}_k) \le -(\widetilde{\delta}/2)\|\widetilde{x}_k - x_k\| \le 0$ for each $t \in T_\varepsilon(\bar{x})$ when $k \in \mathbb{N}$ is large. In the remaining case of $t \notin T_\varepsilon(\bar{x})$, it follows directly that

$$\begin{aligned}\varphi_t(\widetilde{x}_k) &\le \varphi_t(\bar{x}) + \langle \nabla\varphi_t(\bar{x}), \widetilde{x}_k - \bar{x}\rangle + \|\widetilde{x}_k - \bar{x}\| r(\widehat{\eta}_k)\\ &\le -\varepsilon + \sup_{\tau \in T} \|\nabla\varphi_\tau(\bar{x})\|\widehat{\eta}_k + \widehat{\eta}_k r(\widehat{\eta}_k).\end{aligned}$$

Hence $\varphi_t(\widetilde{x}_k) \le 0$, $t \in T$, and $h(\widetilde{x}_k) = 0$ for large $k \in \mathbb{N}$. This means that $\widetilde{x}_k \in \Omega$ for such k, which contradicts the conclusion achieved above. Thus we get that $N(\bar{x}; \Omega) \subset A_\varepsilon$ for all $\varepsilon \in (0, \widetilde{\varepsilon})$. To complete the proof of inclusion "$\subset$" in (8.11), we only need to verify that

$$\bigcap_{\varepsilon>0} A_\varepsilon \subset \bigcap_{\varepsilon>0} \Big[\text{cl}^*\text{cone}\big\{\nabla\varphi_t(\bar{x}) \big| \, t \in T_\varepsilon(\bar{x})\big\}\Big] + \nabla h(\bar{x})^* Y^*. \tag{8.17}$$

Let us take any u^* in the left-hand side of (8.17). This means that for any $\varepsilon > 0$ we can find $x^*_\varepsilon \in C_\varepsilon := \text{cl}^*\text{cone}\big\{\nabla\varphi_t(\bar{x}) \big| \, t \in T_\varepsilon(\bar{x})\big\}$ and $y^*_\varepsilon \in Y^*$ such that $u^* = x^*_\varepsilon + \nabla h(\bar{x})^* y^*_\varepsilon$. Standard arguments similar to those used above show that the net $u^* - \nabla h(\bar{x})^* y^*_\varepsilon = x^*_\varepsilon$ is uniformly bounded. Then the Alaoglu-Bourbaki theorem gives us a subnet of $\{\varepsilon\}$ labeled as $\{\varepsilon_\nu\}$ such that $x^*_{\varepsilon_\nu} \xrightarrow{w^*} x^*$. It follows that $u^* - x^* \in \text{cl}^*(\nabla h(\bar{x})^* Y^*)$. Lemma 8.11 tells us that there is $y^* \in Y^*$ satisfying $u^* - x^* = \nabla h(\bar{x})^* y^*$. Note further that $\varepsilon_\nu \to 0$ and so for any $\alpha > 0$ we get $w^* - \lim_\nu x^*_{\varepsilon_\nu} \in \text{cl}^* C_\alpha = C_\alpha$. It follows that $x^* \in \bigcap_{\alpha>0} C_\alpha$. This implies that $u^* = x^* + \nabla h(\bar{x})^* y^*$ belongs to the right-hand side of (8.17) and thus completes the proof of the theorem. $\triangle$

Let us show next that PMFCQ is essential for the validity of both normal cone representations in (8.10) and (8.11). Moreover, this condition cannot be replaced by its weaker EMFCQ version.

Example 8.13 (Violation of the Normal Cone Representations in the Absence of PMFCQ). Consider the infinite system in $\mathbb{R}^2$ given in Example 8.6. It is shown there that EMFCQ holds at $\bar{x} = (-1, 0)$ but PMFCQ fails. We can easily check that in this case $\widehat{N}(\bar{x}; \Omega) = N(\bar{x}; \Omega) = \mathbb{R}_+ \times \mathbb{R}_-$ while

$$\text{cl cone}\big\{\nabla\varphi_t(\bar{x}) \big| \, t \in T_\varepsilon(\bar{x})\big\} = \text{cl cone}\big\{(1,0) \cup \{(t,0) \big| \, t \in (0,\varepsilon)\big\} = \mathbb{R}_+ \times \{0\}.$$

i.e., the inclusions "$\subset$" in (8.10) and (8.11) are violated.

The next example demonstrates that the perturbed index set $T_\varepsilon(\bar{x})$ cannot be replaced by its unperturbed counterpart $T(\bar{x})$ in representations (8.10) and (8.11) for both regular and basic normal cones.

Example 8.14 (Perturbation of the Active Index Set Is Essential). Let us revisit the nonlinear infinite system in SIP of Example 8.3:

$$\varphi_1(x) = x_1 + 1 \le 0, \quad \varphi_k(x) = \frac{1}{3k}x_1^3 - x_2 \le 0 \ \text{ for } \ k \in \mathbb{N} \setminus \{1\},$$

where $x = (x_1, x_2) \in \mathbb{R}^2$ and $T := \mathbb{N}$. It is easy to see that this inequality system satisfies our standing assumptions and that the functions $\varphi_t(x)$ are uniformly strictly differentiable at $\bar{x} = (-1, 0)$. Observe further that $\widehat{N}(\bar{x}; \Omega) = N(\bar{x}; \Omega) = \mathbb{R}_+ \times \mathbb{R}_-$. As shown above, both PMFCQ and EMFCQ hold at $\bar{x}$. However, $T(\bar{x}) = \{1\}$ and

$$N(\bar{x}; \Omega) \ne \text{cone}\left\{\nabla\varphi_t(\bar{x})\middle|\ t \in T(\bar{x})\right\} = \text{cone}\left\{\nabla\varphi_1(\bar{x})\right\} = \mathbb{R}_+ \times \{0\},$$

which shows the violation of the unperturbed counterparts of (8.10), (8.11).

Now we derive several corollaries of Theorem 8.12, which are of their own interest. The first one concerns the case where the set $\{\nabla\varphi_t(\bar{x})|\ t \in T\}$ may *not be bounded* in X^* as in our standing assumptions. It follows that this case can be reduced to the basic case of Theorem 8.12 with some modifications.

Corollary 8.15 (Normal Cone Representation for Infinite Systems with Unbounded Gradients). *Considering* (8.2)*, assume the following:*

(a) *The functions φ_t, $t \in T$, are Fréchet differentiable at the point $\bar{x}$ with $\|\nabla\varphi_t(\bar{x})\| > 0$ for all $t \in T$, and the mapping h is strictly differentiable at $\bar{x}$.*

(b) *We have that $\lim_{\eta\downarrow 0}\widetilde{r}(\eta) = 0$, where $\widetilde{r}(\eta)$ is defined by*

$$\widetilde{r}(\eta) := \sup_{t\in T}\ \sup_{\substack{x,x'\in B_\eta(\bar{x})\\ x\ne x'}} \frac{|\varphi_t(x) - \varphi_t(x') - \langle\nabla\varphi_t(\bar{x}), x - x'\rangle|}{\|\nabla\varphi_t(\bar{x})\| \cdot \|x - x'\|} \ \ \textit{for all} \ \ \eta > 0.$$

(c) *The operator $\nabla h(\bar{x})\colon X \to Y$ is surjective, and for some $\varepsilon > 0$, there are $\widetilde{x} \in X$ and $\sigma > 0$ such that $\nabla h(\bar{x})\widetilde{x} = 0$ and that*

$$\langle\nabla g_t(\bar{x}), \widetilde{x} + x\rangle \le 0 \ \ \textit{if} \ \ \|x\| \le \sigma, \ \ t \in \widetilde{T}_\varepsilon(\bar{x}) := \{t \in T|\ g_t(\bar{x}) \ge -\varepsilon\|\nabla g_t(\bar{x})\|\},$$

which can be treated as an updated version of PMFCQ in the unbounded setting. Then the basic normal cone to Ω at $\bar{x}$ is calculated by formula (8.11)*.*

Proof. Define $\widetilde{\varphi}_t(x) := \varphi_t(x)\|\nabla\varphi_t(\bar{x})\|^{-1}$ for all $x \in X,\ t \in T$ and observe that the feasible set Ω from (8.2) admits the representation

$$\Omega = \left\{x \in X\middle|\ \widetilde{\varphi}_t(x) \le 0,\ h(x) = 0\right\}.$$

Replacing φ_t by $\widetilde{\varphi}_t$ in Theorem 8.12, we have that the functions $\widetilde{\varphi}_t$ and h satisfy the standing assumptions (SA) as well as condition (8.4) with taking $\widetilde{r}(\eta)$ instead of $r(\eta)$. Furthermore, it follows from (c) that for some $\varepsilon > 0$ there are $\widetilde{x} \in X$ and $\sigma > 0$ satisfying $\nabla h(\bar{x})\widetilde{x} = 0$ and such that

$$\langle \nabla\widetilde{\varphi}_t(\bar{x}), \widetilde{x}\rangle \le -\sup_{x\in B_\sigma(\bar{x})} \langle \nabla\widetilde{\varphi}_t(\bar{x}), x\rangle = -\sigma\|\nabla\widetilde{\varphi}_t(\bar{x})\| \;\text{ if }\; t \in \widetilde{T}_\varepsilon(\bar{x}),$$

which turns into $\langle \nabla\widetilde{\varphi}_t(\bar{x}), \widetilde{x}\rangle \le -\sigma$ for all $t \in \widetilde{T}_\varepsilon(\bar{x}) = \{t \in T|\ \widetilde{\varphi}_t(\bar{x}) \ge -\varepsilon\}$. Hence PMFCQ holds for $(\widetilde{\varphi}_t, h)$ at $\bar{x}$. It follows from (8.11) for $(\widetilde{\varphi}_t, h)$ that

$$\begin{aligned} N(\bar{x};\Omega) &= \bigcap_{\varepsilon>0} \mathrm{cl}^*\mathrm{cone}\left\{\nabla\widetilde{\varphi}_t(\bar{x})\middle|\ t\in\widetilde{T}_\varepsilon(\bar{x})\right\} + \nabla h(\bar{x})^*Y^* \\ &= \bigcap_{\varepsilon>0} \mathrm{cl}^*\mathrm{cone}\left\{\nabla\varphi_t(\bar{x})\,\|\nabla\varphi_t(\bar{x})\|^{-1}\middle|\ t\in\widetilde{T}_\varepsilon(\bar{x})\right\} + \nabla h(\bar{x})^*Y^* \\ &= \bigcap_{\varepsilon>0} \mathrm{cl}^*\mathrm{cone}\left\{\nabla\varphi_t(\bar{x})\middle|\ t\in\widetilde{T}_\varepsilon(\bar{x})\right\} + \nabla h(\bar{x})^*Y^*, \end{aligned}$$

which justifies (8.11) for (φ_t, h) under the assumptions made. △

The next consequence of Theorem 8.12 concerns SIPs in finite dimensions and justifies simplified representations of the normal cones to infinite constraints without the closure operations in (8.10) and (8.11) and with the replacement of the ε-active index set $T_\varepsilon(\bar{x})$ by that of $T(\bar{x})$ from (8.6).

Corollary 8.16 (Normal Cone Representations for Infinite Systems with Compact Index Sets). *In the setting of* (8.2)*, suppose that* $\dim Y < \dim X < \infty$*, that* T *is a compact metric space, that the function* $t \mapsto \varphi_t(\bar{x})$ *is u.s.c. on* T*, that the mapping* $t \mapsto \nabla\varphi_t(\bar{x})$ *is continuous on* T*, and that PMFCQ holds at* $\bar{x}$*. Then letting*

$$\widetilde{N}(\bar{x};\Omega) := \mathrm{cone}\left\{\nabla\varphi_t(\bar{x})\middle|\ t\in T(\bar{x})\right\} + \nabla h(\bar{x})^*Y^*, \tag{8.18}$$

we have $\widetilde{N}(\bar{x};\Omega) = \widehat{N}(\bar{x};\Omega)$ *when the functions* φ_t *are uniformly Fréchet differentiable at* $\bar{x}$ *and* $\widetilde{N}(\bar{x};\Omega) = N(\bar{x};\Omega)$ *when* φ_t *are uniformly strictly differentiable at* $\bar{x}$*. In particular, if we assume in addition that* $t \mapsto \varphi_t(\bar{x})$ *and* $(x,t) \mapsto \nabla\varphi_t(x)$ *are continuous on* T *and* $X \times T$*, respectively, then we also have* (8.18) *for* $\widetilde{N}(\bar{x};\Omega) = N(\bar{x};\Omega)$ *provided that merely EMFCQ holds at* $\bar{x}$*.*

Proof. Let $X = \mathbb{R}^d$ for some $d \in \mathbb{N}$. It follows from Proposition 8.5 that φ_t, $t \in T$, and h satisfy the standing assumptions (SA). Since system (8.2) satisfies PMFCQ at $\bar{x}$, there are $\widetilde{\varepsilon} > 0$, $\delta > 0$, and $\widetilde{x} \in X$ such that $\langle \nabla\varphi_t(\bar{x}), \widetilde{x}\rangle < -\delta$ for all $t \in T_\varepsilon(\bar{x})$ and $\varepsilon \in (0, \widetilde{\varepsilon})$. Observe that the perturbed active index set $T_\varepsilon(\bar{x})$ is compact in T for all $\varepsilon > 0$ due to the u.s.c. assumption imposed on $t \mapsto \varphi_t(\bar{x})$. The continuity of $t \mapsto \nabla\varphi_t(\bar{x})$ ensures that $\{\nabla\varphi_t(\bar{x})|\ t \in T_\varepsilon(\bar{x})\}$ is a compact subset of $\mathbb{R}^d$. We now claim that $0 \notin \mathrm{co}\,\{\nabla\varphi_t(\bar{x})|\ t \in T_\varepsilon(\bar{x})\}$. Indeed, it follows that

$$\sum_{t\in T_\varepsilon(\bar{x})} \lambda_t\langle \nabla\varphi_t(\bar{x}), \widetilde{x}\rangle \le -\sum_{t\in T_\varepsilon(\bar{x})} \lambda_t\delta = -\delta < 0 \text{ as } \lambda \in \widetilde{\mathbb{R}}_+^{T_\varepsilon(\bar{x})},\ \sum_{t\in T_\varepsilon(\bar{x})} \lambda_t = 1,$$

which yields $0 \neq \sum_{t \in T_\varepsilon(\bar{x})} \lambda_t \nabla\varphi_t(\bar{x})$, i.e., $0 \notin \mathrm{co}\{\nabla\varphi_t(\bar{x})|\ t \in T_\varepsilon(\bar{x})\}$. The result of Exercise 8.89 tells us that the convex conic hull of the set $\{\nabla\varphi_t(\bar{x})|\ t \in T_\varepsilon(\bar{x})\}$ is closed in $\mathbb{R}^d$. By Theorem 8.12 it suffices to show that

$$\bigcap_{\varepsilon>0} \mathrm{cone}\left\{\nabla\varphi_t(\bar{x})\big|\ t \in T_\varepsilon(\bar{x})\right\} = \mathrm{cone}\left\{\nabla\varphi_t(\bar{x})\big|\ t \in T(\bar{x})\right\},$$

where the inclusion "$\supset$" is obvious due to $T(\bar{x}) \subset T_\varepsilon(\bar{x})$ as $\varepsilon > 0$. To justify the converse inclusion, pick any x^* from the set on the left-hand side therein and by the classical Carathéodory theorem find $\lambda_k = (\lambda_{k_1}, \ldots \lambda_{k_{d+1}}) \in \mathbb{R}^{d+1}_+$ and $\nabla\varphi_{t_{k_1}}(\bar{x}), \ldots, \nabla\varphi_{t_{k_{d+1}}}(\bar{x}) \in \{\nabla\varphi_t(\bar{x})|\ t \in T_{k^{-1}}(\bar{x})\} \subset \mathbb{R}^d$ satisfying

$$x^* = \sum_{m=1}^{d+1} \lambda_{k_m} \nabla\varphi_{t_{k_m}}(\bar{x})$$

for all large $k \in \mathbb{N}$. This yields in turn the estimate

$$\langle x^*, \widetilde{x}\rangle = \sum_{m=1}^{d+1} \lambda_{k_m} \langle \nabla\varphi_{t_{k_m}}(\bar{x}), \widetilde{x}\rangle \le -\sum_{m=1}^{d+1} \lambda_{k_m}\delta.$$

Since the sequence $\{\lambda_k\}$ is bounded in $\mathbb{R}^{d+1}$ and so is the one in $\{\lambda_k(\nabla\varphi_{t_{k_1}}(\bar{x}), \ldots, \nabla\varphi_{t_{k_{d+1}}}(\bar{x}))\}$, the compactness of the latter set together with that of T allows us to select sequences $\{\lambda_{k_m}\}$ and $\{t_{k_m}\}$ converging to some $\bar{\lambda}_m$ and $\bar{t}_m \in T$ for each $1 \le m \le d+1$. Note that $0 \ge \varphi_{t_{k_m}}(\bar{x}) \ge -k^{-1}$ for large $k \in \mathbb{N}$, which gives us $0 = \varphi_{\bar{t}_m}(\bar{x})$ whenever $1 \le m \le d+1$. Combining this with the above representation of x^* shows that

$$x^* = \sum_{m=1}^{d+1} \bar{\lambda}_m \nabla\varphi_{\bar{t}_m}(\bar{x}) \in \mathrm{cone}\left\{\nabla\varphi_t(\bar{x})\big|\ t \in T(\bar{x})\right\},$$

which justifies representation (8.18) for both normal cones under the corresponding assumptions on φ_t imposed in Theorem 8.12. The EMFCQ part of this corollary follows from Propositions 8.1 and 8.5. △

The next question addressed in this subsection is about the possibility of obtaining normal cone representations of the "unperturbed" type as in Corollary 8.16 while without any finite dimensionality, compactness, and continuity assumptions made above. The following theorem shows that this can be done when PMFCQ is accompanied by NFMCQ from Definition 8.8. Note that the latter condition is imposed only on the inequality constraint part of (8.2).

Theorem 8.17 (Unperturbed Representations of Normal Cones in General Settings). *Let the functions φ_t, $t \in T$, be uniformly Fréchet differentiable at $\bar{x}$, and let both PMFCQ and NFMCQ hold for (8.2) at $\bar{x}$. Then we have*

$$\widehat{N}(\bar{x};\Omega)=\text{cone}\left\{\nabla\varphi_t(\bar{x})\,\middle|\; t\in T(\bar{x})\right\}+\nabla h(\bar{x})^*Y^*.$$

If in addition the functions φ_t, $t\in T$, satisfy (8.4) at $\bar{x}$, then

$$N(\bar{x};\Omega)=\text{cone}\left\{\nabla\varphi_t(\bar{x})\,\middle|\; t\in T(\bar{x})\right\}+\nabla h(\bar{x})^*Y^*.$$

Proof. First we claim that the set $\bigcap\limits_{\varepsilon>0}\text{cl}^*\text{cone}\,\{\nabla\varphi_t(\bar{x})|\; t\in T_\varepsilon(\bar{x})\}$ belongs to the collections of $x^*\in X^*$ satisfying the inclusion

$$(x^*,\langle x^*,\bar{x}\rangle)\in\text{cl}^*\text{cone}\left\{(\nabla\varphi_t(\bar{x}),\langle\nabla\varphi_t(\bar{x}),\bar{x}\rangle-\varphi_t(\bar{x}))\,\middle|\; t\in T\right\}. \tag{8.19}$$

It follows from PMFCQ at $\bar{x}$ that $\nabla h(\bar{x})$ is surjective and there are $\widetilde{\varepsilon},\delta>0$ and $\widetilde{x}\in X$ such that $\nabla h(\bar{x})\widetilde{x}=0$ and that $\langle\nabla\varphi_t(\bar{x}),\widetilde{x}\rangle<-\delta$ for all $\varepsilon\le\widetilde{\varepsilon}$ and $t\in T_\varepsilon(\bar{x})$. To justify the claimed inclusion, pick $x^*\in\cap_{\varepsilon>0}\text{cl}^*\text{cone}\,\{\nabla\varphi_t(\bar{x})|\; t\in T_\varepsilon(\bar{x})\}$ and for $\varepsilon\in(0,\widetilde{\varepsilon})$ find a net $(\lambda_\nu)_{\nu\in\mathcal{N}}\subset\mathbb{R}^{(T)}_+$ with

$$x^*=w^*-\lim_{\nu\in\mathcal{N}}\sum_{t\in T_\varepsilon(\bar{x})}\lambda_{t\nu}\nabla\varphi_t(\bar{x})$$

from which we deduce the relationships

$$\begin{aligned}\langle x^*,\widetilde{x}\rangle&=\lim_{\nu\in\mathcal{N}}\sum_{t\in T_\varepsilon(\bar{x})}\lambda_{t\nu}\langle\nabla\varphi_t(\bar{x}),\widetilde{x}\rangle\le-\delta\limsup_{\nu\in\mathcal{N}}\sum_{t\in T_\varepsilon(\bar{x})}\lambda_{t\nu},\\ \langle x^*,\bar{x}\rangle&=\lim_{\nu\in\mathcal{N}}\sum_{t\in T_\varepsilon(\bar{x})}\lambda_{t\nu}(\langle\nabla\varphi_t(\bar{x}),\bar{x}\rangle-\varphi_t(\bar{x})+\varphi_t(\bar{x})).\end{aligned}$$

The obtained conditions imply in turn that

$$0\ge\langle x^*,\bar{x}\rangle-\limsup_{\nu\in\mathcal{N}}\sum_{t\in T_\varepsilon(\bar{x})}\lambda_{t\nu}\big(\langle\nabla\varphi_t(\bar{x}),\bar{x}\rangle-\varphi_t(\bar{x})\big)\ge\frac{\varepsilon}{\delta}\langle x^*,\widetilde{x}\rangle.$$

Passing to a subnet and combining it with the representation of x^* yield

$$\begin{aligned}(x^*,\langle x^*,\bar{x}\rangle)\in\text{cl}^*\text{cone}&\left\{\big(\nabla\varphi_t(\bar{x}),\langle\nabla\varphi_t(\bar{x}),\bar{x}\rangle-\varphi_t(\bar{x})\big)\,\middle|\; t\in T\right\}\\&+\{0\}\times[\varepsilon\delta^{-1}\langle x^*,\widetilde{x}\rangle,0]\end{aligned}$$

for all $\varepsilon\in(0,\widetilde{\varepsilon})$, which implies that x^* belongs to the set in (8.19) by letting $\varepsilon\downarrow 0$. Based on NFMCQ, we claim now that

$$\bigcap_{\varepsilon>0}\text{cl}^*\text{cone}\left\{\nabla\varphi_t(\bar{x})\,\middle|\; t\in T_\varepsilon(\bar{x})\right\}=\text{cone}\left\{\nabla\varphi_t(\bar{x})\,\middle|\; t\in T(\bar{x})\right\}. \tag{8.20}$$

The inclusion "$\supset$" in (8.11) is obvious since $T(\bar{x})\subset T_\varepsilon(\bar{x})$ for all $\varepsilon>0$. To justify the converse inclusion in (8.20), pick any element x^* belonging to the set on the left-hand side of (8.20). Then NFMCQ allows us to deduce from (8.19) that there exists $\lambda\in\mathbb{R}^{(T)}_+$ such that

$$(x^*, \langle x^*, \bar{x}\rangle) = \sum_{t\in T} \lambda_t \big(\nabla\varphi_t(\bar{x}), \langle \nabla\varphi_t(\bar{x}), \bar{x}\rangle - \varphi_t(\bar{x})\big).$$

Thus we arrive at the following equalities:

$$0 = \sum_{t\in T} \lambda_t \langle \nabla\varphi_t(\bar{x}), \bar{x}\rangle - \sum_{t\in T} \lambda_t \big(\langle \nabla\varphi_t(\bar{x}), \bar{x}\rangle - \varphi_t(\bar{x})\big) = \sum_{t\in T} \lambda_t \varphi_t(\bar{x}).$$

Since $\varphi_t(\bar{x}) \le 0$, it implies that $\lambda_t\varphi_t(\bar{x}) = 0$ for all $t \in T$ and therefore yields $x^* \in \operatorname{cone}\{\nabla\varphi_t(\bar{x}) |\ t \in T(\bar{x})\}$, which verifies the inclusion "$\subset$" in (8.20). To complete the proof of the theorem, it remains to combine the obtained equality (8.20) with the results of Theorem 8.12. △

The next example shows that PMFCQ *cannot* be replaced by EMFCQ in Theorem 8.17 to ensure the "unperturbed" normal cone representations in the presence of NFMCQ even in finite-dimensional settings.

Example 8.18 (EMFCQ Combined with NFMCQ Doesn't Ensure the Unperturbed Normal Cone Representations). We revisit the infinite constraint system in Example 8.3. It is shown there that this system satisfied EMFCQ but not PMFCQ at $\bar{x} = (-1, 0)$. It is also shown in Example 8.10 that NFMCQ holds at $\bar{x}$. Observe however that both representations in Theorem 8.17 are not satisfied for this system. Indeed, we have

$$\widehat{N}(\bar{x}; \Omega) = N(\bar{x}; \Omega) \neq \operatorname{cone}\big\{\nabla\varphi_t(\bar{x})\big|\ t \in T(\bar{x})\big\} = \operatorname{cone}\{(1, 0)\} = \mathbb{R}_+ \times \{0\}$$

as depicted in Fig. 8.2.

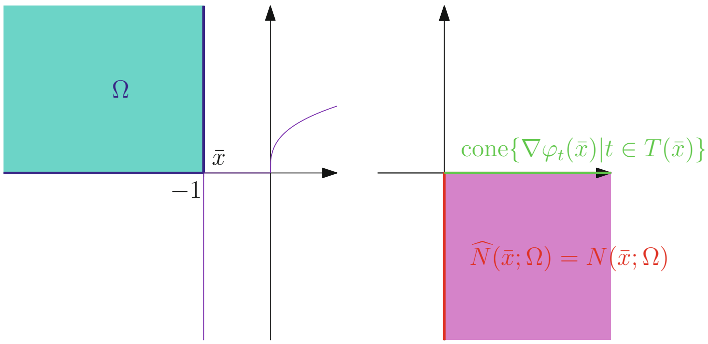


Fig. 8.2 Normal cones to countable inequality constraints.

The next consequence of Theorem 8.17 concerns infinite convex systems.

Corollary 8.19 (Normal Cone for Infinite Convex Systems). *Suppose that all φ_t, $t \in T$, are convex and uniformly Fréchet differentiable and that $h(x) := Ax$ is a surjective continuous linear operator, and that PMFCQ (equivalently SSC) holds at $\bar{x} \in \Omega$. Then the normal cone to the convex set Ω at the point $\bar{x}$ is calculated by*

$$N(\bar{x};\Omega)=\bigcap_{\varepsilon>0}\mathrm{cl}^*\mathrm{cone}\left\{\nabla\varphi_t(\bar{x})\big|\ t\in T_\varepsilon(\bar{x})\right\}+A^*Y^*.$$

If in addition NFMCQ holds at $\bar{x}$, then we have

$$N(\bar{x};\Omega)=\mathrm{cone}\left\{\nabla\varphi_t(\bar{x})\big|\ t\in T(\bar{x})\right\}+A^*Y^*. \tag{8.21}$$

Proof. It follows from Proposition 8.7 and Theorems 8.12 and 8.17. △

Finally in this subsection, we present specifications of the normal cone representation for the case of linear infinite systems.

Proposition 8.20 (Normal Cone Representations for Infinite Linear Constraint Systems). *Consider system (8.2) with $\varphi_t(x)=\langle a_t^*,x\rangle-b_t$, $t\in T$, and $h(x):=Ax$. Suppose that A is a surjective continuous linear operator and that the coefficient set $\{a_t^*|\ t\in T\}$ is bounded. If SSC holds at $\bar{x}$, then we have*

$$N(\bar{x};\Omega)=\bigcap_{\varepsilon>0}\mathrm{cl}^*\mathrm{cone}\left\{a_t^*\big|\ t\in T_\varepsilon(\bar{x})\right\}+A^*Y^*.$$

If in addition the convex conic hull of $\{(a_t^*,b_t)|\ t\in T\}$ is weak* closed in $X^*\times\mathbb{R}$ and $h\equiv 0$, then the simplified representation holds:

$$N(\bar{x};\Omega)=\mathrm{cone}\left\{a_t^*\big|\ t\in T(\bar{x})\right\}+A^*Y^*. \tag{8.22}$$

Proof. Both statements of the proposition follow from the results of Corollary 8.19, where the boundedness of the coefficient set $\{a_t^*|\ t\in T\}$ comes from the standing assumptions (SA) imposed in this section. △

8.1.3 Optimality Conditions for Nonlinear SIPs

Here we derive necessary optimality conditions for SIPs of type (8.1) in infinite-dimensional spaces by combining the above calculation formulas for the normal cones to the feasible solution set with subdifferential calculus. It is done by employing the standard scheme in nonsmooth optimization used in Section 6.1 for non-differentiable programs. The main point reflecting the nature of infinite constraint systems is the *normal cone calculations* given in Subsection 8.1.2 under appropriate constraint qualifications. For brevity we confine ourselves to deriving lower subdifferential optimality conditions while leaving those of the upper subdifferential type as exercises for the reader.

Let us begin with the following necessary optimality conditions in arbitrary Banach spaces under Fréchet differentiability assumptions.

Proposition 8.21 (Necessary Optimality Conditions for Differentiable SIPs in Banach Spaces). *Let $\bar{x}$ be a local minimizer of SIP (8.1) under the validity of PM-FCQ at $\bar{x}$. Suppose further that the inequality constraint functions φ_t, $t\in T$, are* *uniformly Fréchet differentiable at* $\bar{x}$ *and the cost function* φ *is Fréchet differentiable at this point. Then we have the inclusion*

$$0 \in \nabla\varphi(\bar{x}) + \bigcap_{\varepsilon>0} \mathrm{cl}^*\mathrm{cone}\,\big\{\nabla\varphi_t(\bar{x})\big|\; t \in T_\varepsilon(\bar{x})\big\} + \nabla h(\bar{x})^* Y^*. \tag{8.23}$$

If in addition the NFMCQ holds at $\bar{x}$*, then there are multipliers* $\lambda \in \mathbb{R}_+^{(T)}$ *and* $y^* \in Y^*$ *satisfying the differential KKT condition*

$$0 = \nabla\varphi(\bar{x}) + \sum_{t\in T(\bar{x})} \lambda_t \nabla\varphi_t(\bar{x}) + \nabla h(\bar{x})^* y^*. \tag{8.24}$$

Proof. We have that $\bar{x}$ is a local minimizer of unconstrained problem

$$\text{minimize} \quad f = \varphi(x) + \delta(x;\Omega)$$

with the infinite penalty on the constraints in (8.2). Applying the generalized Fermat rule from Proposition 1.30(i), which holds in any Banach space, yields

$$0 \in \widehat{\partial}\big(\varphi + \delta(\cdot;\Omega)\big)(\bar{x}).$$

Since φ is Fréchet differentiable at $\bar{x}$, it follows now from the elementary Banach space sum rule of Proposition 1.30(ii) that

$$0 \in \nabla\varphi(\bar{x}) + \widehat{\partial}\delta(\bar{x};\Omega)(\bar{x}) = \nabla\varphi(\bar{x}) + \widehat{N}(\bar{x};\Omega).$$

Now using representation (8.10) of Theorem 8.12, we arrive at (8.23). The second part (8.24) immediately follows from Theorem 8.17. △

The next result concerns SIPs with nonsmooth objectives on Asplund spaces being more involved in comparison with Proposition 8.21.

Theorem 8.22 (Necessary Optimality Conditions for Nonconvex SIPs in Asplund Spaces, I). *Let* $\bar{x}$ *be a local minimizer of* (8.1)*, where the space* X *is Asplund while* Y *is arbitrary Banach. Suppose that the constraint functions* φ_t*,* $t \in T$*, are uniformly strictly differentiable at* $\bar{x}$*, that the cost function* φ *is l.s.c. around* $\bar{x}$ *and SNEC at this point, and that the qualification condition*

$$\partial^\infty\varphi(\bar{x}) \cap \Big[-\bigcap_{\varepsilon>0} \mathrm{cl}^*\mathrm{cone}\big\{\nabla\varphi_t(\bar{x})\big| t \in T_\varepsilon(\bar{x})\big\} - \nabla h(\bar{x})^* Y^*\Big] = \{0\} \tag{8.25}$$

is fulfilled; the latter two assumptions are automatic when φ *is locally Lipschitzian around* $\bar{x}$*. If PMFCQ is satisfied at* $\bar{x}$*, then*

$$0 \in \partial\varphi(\bar{x}) + \bigcap_{\varepsilon>0} \mathrm{cl}^*\mathrm{cone}\big\{\nabla\varphi_t(\bar{x})\big| t \in T_\varepsilon(\bar{x})\big\} + \nabla h(\bar{x})^* Y^*. \tag{8.26}$$

If in addition we assume that NFMCQ holds at $\bar{x}$ and replace (8.25) by

$$\partial^\infty \varphi(\bar{x}) \cap \Big[-\operatorname{cone}\big\{\nabla\varphi_t(\bar{x}) \big|\ t \in T(\bar{x})\big\} - \nabla h(\bar{x})^* Y^* \Big] = \{0\}, \tag{8.27}$$

then there exist multipliers $\lambda \in \mathbb{R}_+^{(T)}$ and $y^* \in Y^*$ such that the following subdifferential KKT condition is satisfied:

$$0 \in \partial\varphi(\bar{x}) + \sum_{t \in T(\bar{x})} \lambda_t \nabla\varphi_t(\bar{x}) + \nabla h(\bar{x})^* y^*. \tag{8.28}$$

Proof. Observe first that the feasible set Ω is locally closed around $\bar{x}$. Indeed, it follows from (8.4) that there are $\gamma > 0$ and $\eta > 0$ sufficiently small such that for any sequence $\{x_k\} \subset \Omega \cap B_\eta(\bar{x})$ converging to some x_0, we have

$$\|h(x_0)\| \le (\|\nabla h(\bar{x})\| + \gamma)\|x_k - x_0\| \ \text{ and}$$
$$\varphi_t(x_0) \le \sup_{\tau \in T}(\|\nabla\varphi_\tau(\bar{x})\| + \gamma)\, \|x_k - x_0\| + \varphi_t(x_k)$$

for each $t \in T$ and $k \in \mathbb{N}$. By passing to the limit as $k \to \infty$, the latter yields that $h(x_0) = 0$ and $\varphi_t(x_0) \le 0$ for all $t \in T$, i.e., $x_0 \in \Omega \cap B_\eta(\bar{x})$, which justifies the claimed local closedness of Ω.

Employing now the generalized Fermat rule to the unconstrained form of (8.1) at $\bar{x}$ and using the subdifferential sum rule from Theorem 2.19 valid in any Asplund space under the SNEC assumption (see Exercise 2.54(i)) yield

$$0 \in \partial\big(\varphi + \delta(\cdot; \Omega)\big)(\bar{x}) \subset \partial\varphi(\bar{x}) + \partial\delta(\bar{x}; \Omega) = \partial\varphi(\bar{x}) + N(\bar{x}; \Omega)$$

provided that $\partial^\infty\varphi(\bar{x}) \cap \big(-N(\bar{x}; \Omega)\big) = \{0\}$. We apply further to both these conditions the basic normal cone representation of Theorem 8.12. It gives us (8.26) under the fulfillment of (8.25) and PMFCQ at $\bar{x}$. Employing finally Theorem 8.17 instead of Theorem 8.12 in the setting above, we arrive at the KKT condition (8.28) under the assumed NFMCQ at $\bar{x}$ and the qualification condition (8.27). This completes the proof of the theorem. △

Note that if φ is *strictly* differentiable at $\bar{x}$, necessary optimality conditions obtained in Theorem 8.22 and Proposition 8.21 look the same. However, the results of Proposition 8.21 require merely Fréchet differentiability of the cost and constraint functions in the general Banach space setting.

Let us present a consequence of Theorem 8.22 in the case where both spaces X and Y are finite-dimensional and the index set T is compact. This is a conventional situation in SIP theory, except the fact that now the cost function is far removed to be smooth.

Corollary 8.23 (Necessary Conditions for Finite-Dimensional SIPs with Compact Index Sets). *Let $\bar{x}$ be a local minimizer of (8.1), where $\dim Y < \dim X < \infty$, where T is a compact metric space, and where the mappings $(x,t) \mapsto \varphi_t(x)$ and $(x,t) \mapsto \nabla\varphi_t(x)$ are continuous while φ is l.s.c. around $\bar{x}$. If the qualification* *requirements in* (8.27) *and EMFCQ hold at* $\bar{x}$, *then there are multipliers* $\lambda \in \mathbb{R}_+^{(T)}$ *and* $y^* \in Y^*$ *satisfying the KKT condition* (8.28).

Proof. By Proposition 8.9 we have that NFMCQ holds at $\bar{x}$ under the assumptions made, which ensure also by Proposition 8.5 that PMFCQ reduces to EMFCQ. Then the formulated corollary follows from Theorem 8.22. $\triangle$

An important ingredient in the proof of Theorem 8.22 is applying the subdifferential sum rule in Asplund spaces given in Exercise 2.54(i) to the sum $\varphi + \delta(\cdot; \Omega)$. It requires that either φ is SNEC at $\bar{x}$ or Ω is SNC at this point. While the first possibility was used in the proof above, now we are going to explore the second one. The next proposition of its own interest presents verifiable conditions that ensure the SNC property of the infinite system (8.2) expressed entirely in terms of its initial data.

Proposition 8.24 (SNC Property of Infinite Systems). *Let* X *be an Asplund space, and let* $\dim Y < \infty$ *in the framework of* (8.1). *Assume that all the functions* φ_t, $t \in T$, *are Fréchet differentiable around some* $\bar{x} \in \Omega$ *and that the derivative family* $\{\nabla\varphi_t\}_{t\in T}$ *is* EQUICONTINUOUS *around this point in the sense that there exists* $\varepsilon > 0$ *such that for each* $x \in B_\varepsilon(\bar{x})$ *and each* $\gamma > 0$ *there is* $0 < \widetilde{\varepsilon} < \varepsilon$ *with the property:*

$$\|\nabla\varphi_t(x') - \nabla\varphi_t(x)\| \le \gamma \quad \text{whenever} \quad x' \in B_{\widetilde{\varepsilon}}(x) \cap \Omega \quad \text{and} \quad t \in T.$$

Then the feasible set Ω *in* (8.2) *is locally closed around* $\bar{x}$ *and SNC at this point provided that the validity of PMFCQ at* $\bar{x}$.

Proof. Consider first the set $\Omega_1 := \{x \in X|\ \varphi_t(x) \le 0,\ t \in T\}$. By using arguments similar to the proof of Theorem 8.22, we justify the local closedness of Ω_1 around $\bar{x}$. Now let us verify that Ω_1 is SNC at this point. To proceed, pick any sequence $(x_k, x_k^*) \in \Omega_1 \times X^*$, $k \in \mathbb{N}$, satisfying

$$x_k \xrightarrow{\Omega_1} \bar{x},\ \ x_k^* \in \widehat{N}(x_k; \Omega_1) \text{ and } x_k^* \xrightarrow{w^*} 0 \ \text{ as } \ k \to \infty$$

and, by using the imposed equicontinuity, observe that the functions φ_t satisfy the standing assumptions (SA) at x_k for all large $k \in \mathbb{N}$. Further, the reader can check via Exercise 8.91 that condition (8.4) holds at x_k. Since PMFCQ is satisfied at $\bar{x}$, there exist $\delta > 0$, $\varepsilon > 0$, and $\widetilde{x} \in X$ such that $\langle \nabla\varphi_t(\bar{x}), \widetilde{x}\rangle \le -2\delta$ for all $t \in T_{2\varepsilon}(\bar{x})$. Whenever $t \in T_\varepsilon(x_k)$ we have

$$0 \ge \varphi_t(\bar{x}) \ge \varphi_t(x_k) - \langle \nabla\varphi_t(\bar{x}), x_k - \bar{x}\rangle - \|x_k - \bar{x}\| s(\|x_k - \bar{x}\|) \ge -2\varepsilon$$

for large $k \in \mathbb{N}$, where the quantity $s(\cdot)$ is taken from (8.3). Hence we may suppose without loss of generality that

$$T_\varepsilon(x_k) \subset T_{2\varepsilon}(\bar{x}) \text{ and } \sup_{t\in T_\varepsilon(x_k)} \langle \nabla\varphi_t(x_k), \widetilde{x}\rangle \le -\delta \ \text{ for } \ k \in \mathbb{N}. \tag{8.29}$$

Applying now Theorem 8.12 in this setting, we have that whenever $k \in \mathbb{N}$ there exists a net $\{\lambda_{k_\nu}\}_{\nu\in\mathcal{N}} \subset \widetilde{\mathbb{R}}_+^{T_\varepsilon(x_k)}$ such that

$$x_k^* = w^* - \lim_{\nu\in\mathcal{N}} \sum_{t\in T_\varepsilon(x_k)} \lambda_{tk_\nu} \nabla\varphi_t(x_k).$$

Combining this with (8.29) gives us the estimate

$$\langle x_k^*, \widetilde{x}\rangle = \lim_{\nu\in\mathcal{N}} \sum_{t\in T_\varepsilon(x_k)} \lambda_{tk_\nu} \langle\nabla\varphi_t(x_k), \widetilde{x}\rangle \le -\delta \liminf_{\nu\in\mathcal{N}} \sum_{t\in T_\varepsilon(x_k)} \lambda_{tk_\nu}.$$

Furthermore, for each $x \in X$, we get the relationships

$$\begin{aligned}\|x_k^*\| &= \sup_{x\in\mathbb{B}} \left|\liminf_{\nu\in\mathcal{N}} \sum_{t\in T_\varepsilon(x_k)} \lambda_{tk_\nu} \langle\nabla\varphi_t(x_k), x\rangle\right| \\ &\le \liminf_{\nu\in\mathcal{N}} \sum_{t\in T_\varepsilon(x_k)} \lambda_{tk_\nu} \sup_{\tau\in T} \|\nabla\varphi_\tau(x_k)\| \le -\frac{\langle x_k^*, \widetilde{x}\rangle}{\delta} \sup_{\tau\in T} \|\nabla\varphi_\tau(x_k)\|.\end{aligned}$$

Since $x_k^* \xrightarrow{w^*} 0$, it shows that $\|x_k^*\| \to 0$ and thus Ω_1 is SNC at $\bar{x}$.

Consider next the set $\Omega_2 := \{x \in X|\ h(x) = 0\}$, which is obviously closed around $\bar{x}$. It follows from Exercise 2.30 and finite dimensionality of Y that Ω_2 is SNC at $\bar{x}$. Moreover, we get from Exercise 1.54(ii) that $N(\bar{x}; \Omega_2) = \nabla h(\bar{x})^* Y^*$. Thus for $x^* \in N(\bar{x}; \Omega_1) \cap (-N(\bar{x}; \Omega_2))$ there is $y^* \in Y^*$ satisfying $x^* + \nabla h(\bar{x})^* y^* = 0$. Since $x^* \in N(\bar{x}; \Omega_1)$, we find by Theorem 8.12 such a net $\{\lambda_\nu\}_{\nu\in\mathcal{N}} \in \mathbb{R}_+^{(T)}$ that $x^* = w^* - \lim_{\nu\in\mathcal{N}} \sum_{t\in T_\varepsilon(\bar{x})} \lambda_{t\nu} \nabla\varphi_t(\bar{x})$, which yields

$$0 = -\langle\nabla h(\bar{x})^* y^*, \widetilde{x}\rangle = \lim_{\nu\in\mathcal{N}} \sum_{t\in T_\varepsilon(\bar{x})} \lambda_{t\nu} \langle\nabla\varphi_t(\bar{x}), \widetilde{x}\rangle \le -2\delta \liminf_{\nu\in\mathcal{N}} \sum_{t\in T_\varepsilon(\bar{x})} \lambda_{t\nu}.$$

This in turn ensures the relationships

$$\begin{aligned}\langle x^*, x\rangle &= \liminf_{\nu\in\mathcal{N}} \sum_{t\in T_\varepsilon(\bar{x})} \lambda_{t\nu} \langle\nabla\varphi_t(\bar{x}), x\rangle \\ &\le \liminf_{\nu\in\mathcal{N}} \sum_{t\in T_\varepsilon(\bar{x})} \lambda_{t\nu} \sup_{\tau\in T} \|\nabla\varphi_\tau(\bar{x})\| \cdot \|x\| = 0,\ \ x \in X.\end{aligned}$$

Hence we have $x^* = 0$, and so $N(\bar{x}; \Omega_1) \cap (-N(\bar{x}; \Omega_2)) = \{0\}$. It finally follows from Exercise 2.45 that the intersection $\Omega = \Omega_1 \cap \Omega_2$ is SNC at $\bar{x}$. △

Observe that the assumption $\dim Y < \infty$ is essential in Proposition 8.24. To illustrate this, consider a particular case of (8.1) when $T = \emptyset$. It follows from Exercise 2.30 that the inverse image $\Omega = h^{-1}(0)$ is SNC at $\bar{x} \in \Omega$ if and only if the set $\{0\}$ is SNC at $0 \in Y$. Since $N(0; \{0\}) = Y^*$, the latter holds if and only if the weak* topology in Y^* agrees with the norm topology in Y^*, which is only the case if $\dim Y < \infty$ by the fundamental Josefson-Nissenzweig theorem from geometric theory of Banach spaces; see [207, Chapter 12].

The alternative SNC version of Theorem 8.22 is as follows.

Theorem 8.25 (Necessary Optimality Conditions for Nonconvex SIPs in Asplund Spaces, II). *Let $\bar{x}$ be a local minimizer of* (8.1) *under the assumptions of Proposition* 8.24*, and let the qualification condition* (8.25) *be satisfied. Then we have the asymptotic necessary optimality condition* (8.26)*. If in addition NFMCQ holds at $\bar{x}$ and* (8.25) *is replaced by* (8.27)*, then there exist multipliers $\lambda \in \mathbb{R}_+^{(T)}$ and $y^* \in Y^*$ such that the subdifferential KKT condition* (8.28) *is satisfied.*

Proof. It follows the lines in the proof of Theorem 8.22 with applying Proposition 8.24 that ensures the SNC and closedness properties of Ω in the subdifferential sum rule for $\varphi + \delta(\cdot; \Omega)$ used therein. △

The next result provides *necessary and sufficient* optimality conditions for convex SIPs in general Banach spaces.

Theorem 8.26 (Characterization of Optimal Solutions to Convex SIPs). *Let both spaces X and Y be Banach. Assume that the functions φ_t, $t \in T$, are convex and uniformly Fréchet differentiable and that $h(x) := Ax$ is a surjective continuous linear operator. Suppose further that the cost function φ is convex and continuous at some point in Ω. If PMFCQ (equivalently SSC) holds at $\bar{x}$, then $\bar{x}$ is a global minimizer of problem* (8.1) *if and only if*

$$0 \in \partial\varphi(\bar{x}) + \bigcap_{\varepsilon>0} \text{cl}^*\text{cone}\left\{\nabla\varphi_t(\bar{x}) \middle|\, t \in T_\varepsilon(\bar{x})\right\} + A^*Y^*.$$

If in addition NFMCQ is also satisfied at $\bar{x}$, then $\bar{x}$ is a global minimizer of problem (8.1) *if and only if there exist $\lambda \in \mathbb{R}_+^{(T)}$ and $y^* \in Y^*$ such that*

$$0 \in \partial\varphi(\bar{x}) + \sum_{t\in T(\bar{x})} \lambda_t \nabla\varphi_t(\bar{x}) + A^*y^*. \tag{8.30}$$

Proof. The convexity of (8.1) tells us that $\bar{x}$ is its global minimizer if and only of $0 \in \partial(\varphi + \delta(\cdot; \Omega))(\bar{x})$. Applying the convex subdifferential sum rule to the latter inclusion valid under the continuity assumption on φ, we get

$$0 \in \partial\varphi(\bar{x}) + \partial\delta(\bar{x}; \Omega) = \partial\varphi(\bar{x}) + N(\bar{x}; \Omega)$$

and then complete the proof by applying the results of Corollary 8.19. △

8.2 Lipschitzian Semi-infinite Programs

The next class of problem of our study consists of fully *nonsmooth SIPs* with inequality constraints given by locally *Lipschitzian* functions:

$$\begin{cases} \text{minimize } \ \varphi(x) \ \text{ subject to} \\ \varphi_t(x) \leq 0 \ \text{ with } \ t \in T, \ x \in X, \end{cases} \tag{8.31}$$

where T is an arbitrary index set. The generalized differential tools of our analysis revolve around basic subgradients of locally Lipschitzian functions, and we suppose that the decision space X is *Asplund* unless otherwise stated. The methods developed in this infinite-dimensional case are not essentially more complicated than in finite dimensions, and the reader can confine his/her main attention to the latter case for simplicity.

Our strategy to derive necessary optimality conditions for problem (8.31) is significantly different in comparison with the approach in the previous section. We consider now the above SIP in the equivalent *single-constrained* form

$$\text{minimize } \ \varphi(x) \ \text{ subject to } \ \psi(x) := \sup\left\{\varphi_t(x) \middle| \ t \in T\right\} \leq 0 \tag{8.32}$$

given by the intrinsically nonsmooth *supremum function* $\psi\colon X \to \overline{\mathbb{R}}$. Problem (8.32) allows us to apply the standard machinery of nondifferentiable programming to deriving necessary optimality conditions *provided* the possibility to evaluate appropriate subgradients of ψ, which of course is of its own significant interest. We proceed in this direction in what follows.

8.2.1 Some Technical Lemmas

In this subsection we present some technical lemmas, which are important to derive the main results in the two subsequent parts of this section. Our standing assumption for the rest of the section is that the constraint functions $\varphi_t\colon \overline{\mathbb{R}}$ are *uniformly locally Lipschitzian* around a given point $\bar{x} \in \operatorname{dom} \psi$ with some rank $K > 0$. This means the existence of a positive number δ such that

$$|\varphi_t(x) - \varphi_t(y)| \leq K\|x - y\| \ \text{ for all } \ x, y \in B_\delta(\bar{x}), \quad t \in T. \tag{8.33}$$

Note that (8.33) yields the local Lipschitz continuity of the supremum function (8.32) around $\bar{x}$ with rank K. Define the set of ε*-active indices* at $\bar{x}$ by

$$T_\varepsilon(\bar{x}) := \left\{t \in T \middle| \ \varphi_t(\bar{x}) \geq \psi(\bar{x}) - \varepsilon\right\}, \quad \varepsilon \geq 0, \tag{8.34}$$

with $T(\bar{x}) := T_0(\bar{x})$ and observe that $T_\varepsilon(\bar{x}) \neq \emptyset$ for $\varepsilon > 0$. We also denote

$$\begin{aligned} &\Delta(T) := \left\{\lambda \in \widetilde{\mathbb{R}}^T_+ \middle| \ \textstyle\sum_{t\in T} \lambda_t = 1\right\}, \\ &\Lambda_\varepsilon(\widehat{x}) := \left\{\lambda \in \Delta(T_\varepsilon(\bar{x})) \middle| \ \varphi_t(\widehat{x}) = \varphi_s(\widehat{x}) \ \text{ for all } t, s \in \operatorname{supp} \lambda\right\}. \end{aligned} \tag{8.35}$$

First we obtain certain "fuzzy estimates" for regular subgradients of supremum functions, which give us some preliminary information important for deriving the main results in what follows.

Lemma 8.27 (Fuzzy Estimates of Regular Subgradients for Supremum Functions). *Let V^* be a weak* neighborhood of the origin in X^*. Then the following assertions hold for the supremum function ψ in (8.32):*

(i) *For each regular subgradient $x^* \in \widehat{\partial}\psi(\bar{x})$ and each $\varepsilon > 0$, there are elements $\widehat{x} \in B_\varepsilon(\bar{x})$ and $\lambda \in \Lambda_\varepsilon(\widehat{x})$ from (8.35) such that*

$$x^* \in \sum_{t \in T_\varepsilon(\bar{x})} \lambda_t \partial\varphi_t(\widehat{x}) + V^*.$$

(ii) *For each regular subgradient $x^* \in \widehat{\partial}\psi(\bar{x})$ and each $\varepsilon > 0$, there are elements $\lambda \in \Delta(T_\varepsilon(\bar{x}))$ and $\widehat{x}_t \in B_\varepsilon(\bar{x})$ for all $t \in T_\varepsilon(\bar{x})$ such that*

$$x^* \in \sum_{t \in T_\varepsilon(\bar{x})} \lambda_t \widehat{\partial}\varphi_t(\widehat{x}_t) + V^*.$$

Proof. To justify (i), fix arbitrary $x^* \in \widehat{\partial}\psi(\bar{x})$ and then find $m \in \mathbb{N}$, $\gamma > 0$, and $x_k \in X$ as $k = 1, \ldots, m$ satisfying the inclusion

$$\gamma\mathbb{B}^* + \operatorname{span}\{x_1, \ldots, x_m\}^\perp \subset \frac{1}{2}V^*. \tag{8.36}$$

Without loss of generality, we assume that V^* is convex and that $2\varepsilon \leq \gamma$ and then define $L := \operatorname{span}\{x_1, \ldots, x_m\}$. Since $x^* \in \widehat{\partial}\psi(\bar{x})$, there is $\delta > 0$ with

$$\delta < \frac{1}{2}, \quad 2(K+1)\delta \leq \varepsilon, \quad \text{and} \quad \frac{\varepsilon + 2(K+1)\delta}{1 - 2\delta} \leq \gamma \tag{8.37}$$

such that φ_t are uniformly Lipschitzian with rank K in $B_{2\delta}(\bar{x})$ and that

$$\psi(x) - \psi(\bar{x}) - \langle x^*, x - \bar{x}\rangle \geq -\varepsilon\|x - \bar{x}\| \ \text{ for all } x \in B_\delta(\bar{x}). \tag{8.38}$$

Consider now the following constrained optimization problem:

$$\begin{cases} \text{minimize } \ y - \langle x^*, x - \bar{x}\rangle + \varepsilon\|x - \bar{x}\| - \psi(\bar{x}) \ \text{ subject to} \\ \varphi_t(x) - y \leq 0, \quad t \in T, \ (x, y) \in B_\delta(\bar{x}) \times \mathbb{R}. \end{cases} \tag{8.39}$$

It follows from (8.38) that $(\bar{x}, \psi(\bar{x}))$ is a local minimizer of (8.39). Define the l.s.c. function $g\colon X \times \mathbb{R} \to \overline{\mathbb{R}}$ by

$$g(x, y) := y - \langle x^*, x - \bar{x}\rangle + \varepsilon\|x - \bar{x}\| - \psi(\bar{x}) + \delta\big((x, y); \Omega\big)$$

with $\Omega := (L \cap B_\delta(\bar{x})) \times [\psi(\bar{x}) - 1, \psi(\bar{x}) + 1]$ and then a family of $g_t\colon X \times \mathbb{R} \to \overline{\mathbb{R}}$ by $g_t(x, y) := \varphi_t(x) - y$ for all $t \in T$. Due to (8.38) we have the inclusion

$$\big\{(x, y) \in X \times \mathbb{R} \big| g(x, y) + \delta^2 \leq 0\big\} \subset \bigcup_{t \in T} \big\{(x, y) \in \operatorname{int} B_{2\delta}(\bar{x}) \times \mathbb{R} \big| g_t(x, y) > 0\big\}.$$

The set on the left-hand side of the latter inclusion is closed and bounded and hence compact in the finite-dimensional space $L \times \mathbb{R}$. Furthermore, each set $\{(x, y) \in \text{int}B_{2\delta}(\bar{x}) \times \mathbb{R} |\ g_t(x, y) > 0\}$ is open due to the Lipschitz continuity of φ_t on $B_{2\delta}(\bar{x})$. Thus there exists a *finite subset* S of T such that

$$\big\{(x, y) \in X \times \mathbb{R} \big| g(x, y) + \delta^2 \le 0\big\} \subset \bigcup_{s \in S} \big\{(x, y) \in \text{int}B_{2\delta}(\bar{x}) \times \mathbb{R} \big| g_s(x, y) > 0\big\},$$

which implies that $(\bar{x}, \psi(\bar{x}))$ is a δ^2-*optimal solution* to the following optimization problem with *finitely many* inequality constraints:

$$\begin{cases} \text{minimize } \ g(x, y) \ \text{ subject to} \\ g_s(x, y) \le 0, \quad s \in S, \ (x, y) \in B_\delta(\bar{x}) \times \mathbb{R}. \end{cases} \tag{8.40}$$

Note further that $\partial g_s(x, y) \subset X^* \times \{-1\}$ for all $s \in S$ and that $N((x, y); B_\delta(\bar{x}) \times \mathbb{R}) \subset N(x; B_\delta(\bar{x})) \times \{0\}$. This ensures the implication

$$\Big[0 \in \sum_{s \in S(x,y)} \lambda_s \partial g_s(x, y) + N\big((x, y); B_\delta(\bar{x}) \times \mathbb{R}\big)\Big] \Longrightarrow \big[\lambda_s = 0, \ s \in S(x, y)\big]$$

whenever $\lambda_s \ge 0$ for $s \in S(x, y) := \{s \in S | g_s(x, y) = 0\} = \{s \in S | \varphi_s(x) = y\}$. Applying now to problem (8.40) the *suboptimality conditions* from Exercise 6.35(ii) held in Asplund spaces (see [522, Theorem 5.30]) gives us $(\widehat{x}, \widehat{y}) \in X \times \mathbb{R}$, $(\widehat{x}^*, 1) \in \partial g(\widehat{x}, \widehat{y})$, $(x_s^*, -1) \in \partial g_s(\widehat{x}, \widehat{y})$ as $s \in S$, $(u^*, 0) \in N((\widehat{x}, \widehat{y}); B_\delta(\bar{x}) \times \mathbb{R}))$, and $\lambda \in \mathbb{R}^S_+$ such that $\|\widehat{x} - \bar{x}\| + |\widehat{y} - \psi(\bar{x})| \le \delta/2$ and

$$\Big\|(\widehat{x}^*, 1) + \sum_{s \in S(\widehat{x}, \widehat{y})} \lambda_s (x_s^*, -1) + (u^*, 0)\Big\| \le 2\delta. \tag{8.41}$$

Since $\widehat{x} \in B_{\frac{\delta}{2}}(\bar{x}) \subset \text{int}B_\delta(\bar{x})$, we have $u^* \in N(\widehat{x}; B_\delta(\bar{x})) = \{0\}$. Moreover, it follows from the convexity of the function g that

$$\begin{aligned} \widehat{x}^* &\in -x^* + \varepsilon \mathbb{B}^* + N\big(\widehat{x}; L \cap B_\delta(\bar{x})\big) \subset -x^* + \varepsilon \mathbb{B}^* + N(\widehat{x}; L) + N\big(\widehat{x}; B_\delta(\bar{x})\big) \\ &\subset -x^* + \varepsilon \mathbb{B}^* + L^\perp. \end{aligned}$$

Thus estimate (8.41) yields $\| \sum_{s \in S(\widehat{x}, \widehat{y})} \lambda_s - 1\| \le 2\delta$ and

$$x^* \in \sum_{s \in S(\widehat{x}, \widehat{y})} \lambda_s x_s^* + (\varepsilon + 2\delta)\mathbb{B}^* + L^\perp. \tag{8.42}$$

By $\delta < 1/2$ we have $\sum_{s \in S(\widehat{x}, \widehat{y})} \lambda_s > 0$. Let us further define

$$\lambda_s' := \lambda_s \Big[\sum_{t \in S(\widehat{x}, \widehat{y})} \lambda_t\Big]^{-1} \ \text{ for all } s \in S(\widehat{x}, \widehat{y}),$$

which gives us $\sum_{s\in S(\widehat{x},\widehat{y})}\lambda'_s=1$. Since $\|x^*\|\le K$, we derive from (8.36), (8.37), and (8.42) the following chain of inclusions:

$$\begin{aligned}
x^* &\in \frac{1}{\sum_{s\in S(\widehat{x},\widehat{y})}\lambda_s}x^*+\left|1-\frac{1}{\sum_{s\in S(\widehat{x},\widehat{y})}\lambda_s}\right|\|x^*\|\mathbb{B}^*\\
&\subset \sum_{s\in S(\widehat{x},\widehat{y})}\lambda'_s x^*_s+\frac{\varepsilon+2\delta}{\sum_{s\in S(\widehat{x},\widehat{y})}\lambda_s}\mathbb{B}^*+L^\perp+\frac{\left|\sum_{s\in S(\widehat{x},\widehat{y})}\lambda_s-1\right|}{\sum_{s\in S(\widehat{x},\widehat{y})}\lambda_s}K\mathbb{B}^*\\
&\subset \sum_{s\in S(\widehat{x},\widehat{y})}\lambda'_s x^*_s+\frac{\varepsilon+2\delta}{1-2\delta}\mathbb{B}^*+L^\perp+\frac{2\delta}{1-2\delta}K\mathbb{B}^* \qquad (8.43)\\
&\subset \sum_{s\in S(\widehat{x},\widehat{y})}\lambda'_s x^*_s+\frac{\varepsilon+2(K+1)\delta}{1-2\delta}\mathbb{B}^*+L^\perp\\
&\subset \sum_{s\in S(\widehat{x},\widehat{y})}\lambda'_s x^*_s+\gamma\mathbb{B}^*+L^\perp\subset \sum_{s\in S(\widehat{x},\widehat{y})}\lambda'_s x^*_s+V^*.
\end{aligned}$$

Now we claim that $S(\widehat{x},\widehat{y})\subset T_\varepsilon(\bar{x})$. Indeed, it follows from (8.33) that

$$\varphi_s(\bar{x})\ge\varphi_s(\widehat{x})-K\|\bar{x}-\widehat{x}\|\ge\widehat{y}-K\frac{\delta}{2}\ge\psi(\bar{x})-\frac{\delta}{2}-K\frac{\delta}{2}\ge\psi(\bar{x})-\varepsilon$$

for each $s\in S(\widehat{x},\widehat{y})$, which implies that $s\in T_\varepsilon(\bar{x})$. It yields $S(\widehat{x},\widehat{y})\subset T_\varepsilon(\bar{x})$ and together with (8.43) verifies assertion (i) of the lemma.

To justify assertion (ii), we get $x^*_s\in\partial\varphi_s(\widehat{x})$ for $s\in S(\widehat{x},\widehat{y})$ from the proof of (i) and then by the first subdifferential representation in (1.37), which valid in Asplund spaces, find $x_s\in X$ and $\widehat{x}^*_s\in\widehat{\partial}\varphi_s(x_s)$ such that $\|x_s-\widehat{x}\|\le\delta$ and $x^*_s\in\widehat{x}^*_s+V^*$. This gives us the inclusions

$$\begin{aligned}
x^* &\in \sum_{s\in S(\widehat{x},\widehat{y})}\lambda'_s x^*_s+V^*\subset\sum_{s\in S(\widehat{x},\widehat{y})}\lambda'_s\widehat{x}^*_s+\sum_{s\in S(\widehat{x},\widehat{y})}\lambda'_s V^*+V^*\\
&\subset \sum_{s\in S(\widehat{x},\widehat{y})}\lambda'_s\widehat{x}^*_s+V^*+V^*\subset\sum_{s\in S(\widehat{x},\widehat{y})}\lambda'_s\widehat{x}^*_s+2V^*.
\end{aligned}$$

Since $\|x_s-\bar{x}\|\le\|x_s-\widehat{x}\|+\|\widehat{x}-\bar{x}\|\le 2\delta\le\varepsilon$ and $S(\widehat{x},\widehat{y})\subset T_\varepsilon(\bar{x})$, we arrive at the claimed inclusion in (ii) and complete the proof of the lemma. △

The next two lemmas don't directly relate to either SIPs or subgradients of supremum functions while being of their own interest and important for the proofs of the main results of this section.

Lemma 8.28 (Weak* Closed Conic Hulls). *Let $A\subset X^*$ be weak* compact with $0\notin A$ for a Banach space X. Then the conic hull $\mathbb{R}_+A$ is weak* closed.*

Proof. To show that the cone $\mathbb{R}_+A$ is weak* closed in X^*, take any net $\{x^*_\nu\}_{\nu\in\mathcal{N}}\subset\mathbb{R}_+A$ weak* converging to some $x^*\in X^*$. Hence there exist nets $\{\lambda_\nu\}_{\nu\in\mathcal{N}}\subset\mathbb{R}_+$ and $\{u^*_\nu\}_{\nu\in\mathcal{N}}\subset A$ such that $\lambda_\nu u^*_\nu=x^*_\nu\xrightarrow{w^*}x^*$. Define $\lambda:=\limsup_{\nu\in\mathcal{N}}\lambda_\nu$. If $\lambda=\infty$, then we find a subnet $\{\lambda_\nu\}$ (without relabeling) converging to ∞. Since A

is weak* compact, assume without loss of generality that $u_\nu^* \xrightarrow{w^*} u^*$. Furthermore, the relationships $\langle \lambda_\nu u_\nu^*, x\rangle \to \langle x^*, x\rangle$ and $\langle u_\nu^*, x\rangle \to \langle u^*, x\rangle$ for all $x \in X$ imply that $\langle u_\nu^*, x\rangle \to 0$ for all $x \in X$ due to $\lambda_\nu \to \infty$. This gives us $0 \in A$, which contradicts the assumption made. Thus $\lambda < \infty$. By similar arguments we show that $\lambda_\nu \to \lambda \in \mathbb{R}_+$ and $u_\nu^* \xrightarrow{w^*} u^* \in A$. It follows then that $x^* = \lambda u^* \in \mathbb{R}_+ A$, which tells us that $\mathbb{R}_+ A$ is weak* closed and thus completes the proof of the lemma. $\triangle$

The last lemma establishes some relationships for the Painlevé-Kuratowski sequential outer limits of increasing set-valued mappings in Asplund spaces.

Lemma 8.29 (Outer Limits of Increasing Mappings). *Let X be an Asplund space, and let $F : \mathbb{R}_+ \rightrightarrows X^*$ be a set-valued mapping. Suppose that there is $\varepsilon > 0$ such that $F(\varepsilon)$ is bounded in X^* and that F is increasing, i.e., $F(\varepsilon_1) \subset F(\varepsilon_2)$ whenever $0 \le \varepsilon_1 \le \varepsilon_2$. Then the following assertions hold:*

(i) $\operatorname{cl}^*[\operatorname*{Lim\,sup}_{\varepsilon\downarrow 0} F(\varepsilon)] = \bigcap_{\varepsilon>0} \operatorname{cl}^* F(\varepsilon)$.

(ii) $\operatorname{cl}^*\operatorname{co}\big[\operatorname*{Lim\,sup}_{\varepsilon\downarrow 0} F(\varepsilon)\big] = \operatorname{cl}^*\big[\operatorname*{Lim\,sup}_{\varepsilon\downarrow 0} \operatorname{co} F(\varepsilon)\big]$.

(iii) *If* $0 \notin \operatorname{cl}^*\big[\operatorname*{Lim\,sup}_{\varepsilon\downarrow 0} F(\varepsilon)\big]$*, then we have*

$$\mathbb{R}_+\operatorname{cl}^*\big[\operatorname*{Lim\,sup}_{\varepsilon\downarrow 0} F(\varepsilon)\big] = \operatorname{cl}^*\operatorname*{Lim\,sup}_{\varepsilon\downarrow 0}\big[\mathbb{R}_+ F(\varepsilon)\big],$$

Proof. The inclusion "$\subset$" in (i) immediately follows from the definitions. To justify the converse inclusion, pick any x^* belonging to the set on right-hand side set of (i) and take an arbitrary convex weak* neighborhood V^* of the origin in X^*. This gives us sequences $\varepsilon_k \downarrow 0$ and $x_k^* \in F(\varepsilon_k)$ such that $x^* \in x_k^* + \frac{V^*}{2}$. Since $F(\varepsilon) \subset K\mathbb{B}^*$ for some $\varepsilon, K > 0$, there is a subsequence of $\{x_k^*\}$ (without relabeling) weak* converging to some $u^* \in \operatorname{Lim\,sup}_{\varepsilon\downarrow 0} F(\varepsilon)$. For large $k \in \mathbb{N}$, we have $x_k^* \in u^* + \frac{V^*}{2}$, which yields

$$x^* \in u^* + \frac{V^*}{2} + \frac{V^*}{2} = u^* + V^* \subset \operatorname*{Lim\,sup}_{\varepsilon\downarrow 0} F(\varepsilon) + V^*$$

The arbitrary choice of V^* allows us to complete the proof of (i).

To proceed with the proof of (ii), observe from (i) that

$$\operatorname*{Lim\,sup}_{\varepsilon\downarrow 0} F(\varepsilon) \subset \bigcap_{\varepsilon>0} \operatorname{cl}^* F(\varepsilon) \subset \bigcap_{\varepsilon>0} \operatorname{cl}^*\operatorname{co} F(\varepsilon) = \operatorname{cl}^*\big[\operatorname*{Lim\,sup}_{\varepsilon\downarrow 0} \operatorname{co} F(\varepsilon)\big],$$

which ensures the inclusion "$\subset$" in (ii). To prove the converse inclusion therein, it suffices to verify that

$$\operatorname*{Lim\,sup}_{\varepsilon\downarrow 0} \operatorname{co} F(\varepsilon) \subset \operatorname{cl}^*\operatorname{co}\big[\operatorname*{Lim\,sup}_{\varepsilon\downarrow 0} F(\varepsilon)\big].$$

Assuming the contrary, find sequences $\varepsilon_k \downarrow 0$ and $x_k^* \xrightarrow{w^*} x^*$ with $x_k^* \in \text{co}\,F(\varepsilon_k)$ such that x^* is not in the set on the right-hand side of (ii). The classical convex separation theorem gives us $0 \neq v \in X$ and $\alpha, \beta \in \mathbb{R}$ satisfying

$$\langle x^*, v\rangle > \alpha > \beta > \langle u^*, v\rangle \quad \text{for all} \quad u^* \in \mathop{\text{Lim}\,\text{sup}}_{\varepsilon\downarrow 0} F(\varepsilon). \tag{8.44}$$

Since $x_k^* \xrightarrow{w^*} x^*$, this allows us to suppose without loss of generality that $\langle x_k^*, v\rangle > \frac{\alpha+\beta}{2}$ as $k \in \mathbb{N}$. By $x_k^* \in \text{co}\,F(x_k)$ there exist a finite index set S_k, $\lambda_k \in \Delta(S_k)$ from (8.35), and $x_{k_s}^* \in F(\varepsilon_k)$ as $s \in S_k$ such that

$$x_k^* = \sum_{s\in S_k} \lambda_{k_s} x_{k_s}^*, \qquad k \in \mathbb{N}.$$

Among elements of the set $\{x_{k_s}^* \mid s \in S_k\}$ for each $k \in \mathbb{N}$, we select $\widehat{x}_k^* \in F(\varepsilon_k)$ such that $\langle \widehat{x}_k^*, v\rangle = \max\{\langle x_{k_s}^*, v\rangle \mid s \in S_k\}$. It follows therefore that

$$\langle \widehat{x}_k^*, v\rangle \geq \sum_{s\in S_k} \lambda_{k_s}\langle x_{k_s}^*, v\rangle = \langle x_k^*, v\rangle > \frac{\alpha+\beta}{2}.$$

Since $\{\widehat{x}_k^*\}$ is bounded in X^* and X is Asplund, we can assume that $\{\widehat{x}_k^*\}$ weak* converges to some $u^* \in X^*$. This yields $u^* \in \text{Lim}\,\text{sup}_{\varepsilon\downarrow 0} F(\varepsilon)$ and $\langle u^*, v\rangle \geq \frac{\alpha+\beta}{2}$, which contradicts to (8.44) and thus justifies (ii).

It remains to prove (iii) under the extra assumption made. The inclusion "$\subset$" therein is obvious. To justify the converse inclusion, we show that

$$\mathop{\text{Lim}\,\text{sup}}_{\varepsilon\downarrow 0}[\mathbb{R}_+ F(\varepsilon)] \subset \mathbb{R}_+ \text{cl}^*\Big[\mathop{\text{Lim}\,\text{sup}}_{\varepsilon\downarrow 0} F(\varepsilon)\Big], \tag{8.45}$$

where the set on the the right-hand side is weak* closed by Lemma 8.28. To prove (8.45), pick any element $x^* \neq 0$ from the set on the left-hand side of (8.45) and find $\varepsilon_k \downarrow 0$, $\lambda_k \in \mathbb{R}_+$, and $u_k^* \in F(\varepsilon_k)$ as $k \in \mathbb{N}$ such that $\lambda_k u_k^* \xrightarrow{w^*} x^*$. Following the proof of Lemma 8.28, suppose without loss of generality that $\lambda_k \to \lambda \in \mathbb{R}_+$ and $u_k^* \xrightarrow{w^*} u^* \in \text{Lim}\,\text{sup}_{\varepsilon\downarrow 0} F(\varepsilon)$. It yields $\lambda_k u_k^* \xrightarrow{w^*} \lambda u^* = x^*$, and so x^* belongs to the set on the right-hand side of (8.45). This justifies (iii) and completes the proof of the lemma. △

8.2.2 Basic Subgradients of Supremum Functions

By using the preliminary results of Subsection 8.2.1, we now proceed with deriving *pointbased* upper estimates for the basic subdifferential of the supremum function ψ in (8.32) under various assumptions in Asplund spaces.

Our first theorem here employs the notion of the *weak* outer stability* of a mapping $F\colon Z \rightrightarrows X^*$ at $\bar{z}$ meaning the validity of the inclusion $\operatorname{Lim\,sup}_{z\to\bar{z}} F(z) \subset \operatorname{cl}^* F(\bar{z})$. Note that it relates to the standard notion of weak* outer semicontinuity of F at $\bar{z}$ (not used below), where the operation cl^* is omitted on the right-hand side of the latter inclusion.

Theorem 8.30 (Pointbased Estimates of Basic Subgradients for Supremum Functions). *Given φ_t in (8.31), define $C : \mathbb{R}_+ \rightrightarrows X^*$ by*

$$C(\varepsilon) := \bigcup\Big\{ \sum_{t\in T_\varepsilon(\bar{x})} \lambda_t \partial\varphi_t(x) \Big|\, x \in B_\varepsilon(\bar{x}),\ \lambda \in \Lambda_\varepsilon(x)\Big\},\ \varepsilon \ge 0, \tag{8.46}$$

where $\Lambda_\varepsilon(x)$ is taken from (8.35) for $\varepsilon > 0$, and where $\Lambda_0(x) := \Delta(T(x))$. Then the following assertions hold:

(i) *The basic subdifferential of ψ in (8.32) at $\bar{x}$ is estimated by*

$$\partial\psi(\bar{x}) \subset \bigcap_{\varepsilon>0} \operatorname{cl}^* C(\varepsilon). \tag{8.47}$$

(ii) *The weak* outer stability of the mapping (8.46) at zero ensures that*

$$\partial\psi(\bar{x}) \subset \operatorname{cl}^*\Big[\bigcup\Big\{ \sum_{t\in T(\bar{x})} \lambda_t \partial\varphi_t(\bar{x}) \Big|\, \lambda \in \Delta\big(T(\bar{x})\big)\Big\}\Big]. \tag{8.48}$$

(iii) *If in addition X is reflexive and φ_t, $t \in T(\bar{x})$, are lower regular at $\bar{x}$, then ψ is also lower regular at $\bar{x}$ and (8.48) holds as equality.*

Proof. To justify (i), pick any $x^* \in \partial\psi(\bar{x})$ and fix an arbitrary weak* neighborhood V^* of the origin in X^*. By definition of x^* there are sequences $x_k \to \bar{x}$ and $x_k^* \in \widehat{\partial}\psi(x_k)$ satisfying $x_k^* \xrightarrow{w^*} x^*$. Select a neighborhood U^* of $0 \in X^*$ with $\operatorname{cl}^* U^* \subset V^*$ and find a sequence $\delta_k \downarrow 0$ satisfying $\delta_k > \|x_k - \bar{x}\|$ for all $k \in \mathbb{N}$. It follows from Lemma 8.27 that there exist $\widehat{x}_k \in B_{\delta_k}(x_k)$ and $\lambda_k \in \Delta(T_{\delta_k}(x_k))$ with $\varphi_t(\widehat{x}_k) = \varphi_s(\widehat{x}_k)$ for all $t, s \in \operatorname{supp}\lambda_k$ and such that

$$x_k^* \in \sum_{t\in T_{\delta_k}(x_k)} \lambda_{k_t} \partial\varphi_t(\widehat{x}_k) + U^*. \tag{8.49}$$

Note further that for all $k \in \mathbb{N}$ sufficiently large, we have

$$\begin{aligned} \varphi_t(\bar{x}) &\ge \varphi_t(x_k) - K\|x_k - \bar{x}\| \ge \psi(x_k) - \delta_k - K\|x_k - \bar{x}\| \\ &\ge \psi(\bar{x}) - 2K\|x_k - \bar{x}\| - \delta_k \ge \psi(\bar{x}) - (2K+1)\delta_k \end{aligned}$$

whenever $t \in T_{\delta_k}(x_k)$. Defining $\varepsilon_k := \max\{2\delta_k, (2K+1)\delta_k\}$ and using the above inequalities give us the inclusions $\widehat{x}_k \in B_{\varepsilon_k}(\bar{x})$ and $T_{\delta_k}(x_k) \subset T_{\varepsilon_k}(\bar{x})$. This implies that $\lambda_k \in \Lambda_{\varepsilon_k}(\widehat{x}_k)$ and $x_k^* \in C(\varepsilon_k) + U^*$ by (8.49). It follows that there are $\widehat{x}_k^* \in C(\varepsilon_k)$ and $u_k^* \in U^*$ satisfying $x_k^* = \widehat{x}_k^* + u_k^*$. Observe further that $C(\varepsilon_k)$ is contained in $K\mathbb{B}^*$ for all large $k \in \mathbb{N}$. The sequence $\{\widehat{x}_k^*\}$ is bounded in X^* and hence contains a weak* convergent subsequence by the Asplund property of X. Assuming without loss of generality that it itself converges to some $x^* \in X^*$, we get $u_k^* \xrightarrow{w^*} x^* - \widehat{x}^* \in \operatorname{cl}^* U^*$ and therefore

$$x^* = \widehat{x}^* + (x^* - \widehat{x}^*) \in \Big[\operatorname*{Lim\,sup}_{\varepsilon \downarrow 0} C(\varepsilon)\Big] + \operatorname{cl}^* U^* \subset \Big[\operatorname*{Lim\,sup}_{\varepsilon \downarrow 0} C(\varepsilon)\Big] + V^*$$

for any V^*, which implies that $x^* \in \operatorname{cl}^*[\operatorname{Lim\,sup}_{\varepsilon \downarrow 0} C(\varepsilon)]$. Applying now Lemma 8.29 yields (8.47) and thus justifies (i). Assertion (ii) follows from (i) by the assumed weak* outer stability of the mapping $C(\cdot)$ in (8.46).

It remains to prove (iii) under the additional assumptions made. Take any $x^* \in C(0)$ and find $\lambda \in \Delta(T(\bar{x}))$ such that

$$x^* \in \sum_{t \in T(\bar{x})} \lambda_t \partial \varphi_t(\bar{x}) = \sum_{t \in T(\bar{x})} \lambda_t \widehat{\partial} \varphi_t(\bar{x}).$$

We can easily check the inclusions

$$\sum_{t \in T(\bar{x})} \lambda_t \widehat{\partial} \varphi_t(\bar{x}) \subset \widehat{\partial}\Big(\sum_{t \in T(\bar{x})} \lambda_t \varphi_t\Big)(\bar{x}) \ \text{ and } \ \widehat{\partial}\Big(\sum_{t \in T(\bar{x})} \lambda_t \varphi_t\Big)(\bar{x}) \subset \widehat{\partial}\psi(\bar{x})$$

implied by $\sum_{t \in T(\bar{x})} \lambda_t \varphi_t(\bar{x}) = \psi(\bar{x})$ and $\sum_{t \in T(\bar{x})} \lambda_t \varphi_t(x) \le \psi(x)$ for all $x \in X$. Thus $C(0) \subset \widehat{\partial}\psi(\bar{x})$ and it follows from (8.48) that

$$\widehat{\partial}\psi(\bar{x}) \subset \partial\psi(\bar{x}) \subset \operatorname{cl}^* C(0) \subset \operatorname{cl}^* \widehat{\partial}\psi(\bar{x}) = \widehat{\partial}\psi(\bar{x}),$$

where the last equality holds due to the reflexivity of the space X. This justifies the equality in (8.48) and completes the proof of the theorem. △

Let us construct an example in $\mathbb{R}^2$ showing that the set on the right-hand side of (8.47) is generally *nonconvex*. In this example the *equality* holds in (8.47) and the usage of the perturbed set $\Lambda_\varepsilon(x)$ in (8.46) is essential.

Example 8.31 (Nonconvex Estimate for Basic Subgradients of Supremum Functions). Let $X = \mathbb{R}^2$ and $T = (0, 1) \subset \mathbb{R}$, and let the supremum function $\psi : \mathbb{R}^2 \to \mathbb{R}$ be defined by

$$\psi(x) := \sup\Big\{t x_1^3 - \frac{1}{(t+1)^2}|x_2| + t^3 - 1 \,\Big|\, t \in T\Big\}.$$

Denote $\varphi_t(x) := t x_1^3 - \frac{1}{(t+1)^2}|x_2| + t^3 - 1$, $t \in T$, and let $\bar{x} = (0, 0)$. It is easy to check that the functions φ_t are uniformly Lipschitz continuous around $\bar{x}$, that $T(\bar{x}) = \emptyset$, and that $T_\varepsilon(\bar{x}) = \{t \in T \,|\, t \ge \sqrt[3]{1-\varepsilon}\}$ for all $\varepsilon > 0$. Pick any $x^* \in C(\varepsilon)$ and by (8.46) find $x \in B_\varepsilon(\bar{x})$ and $\lambda \in \Delta(T_\varepsilon(\bar{x}))$ such that

$$x^* \in \sum_{t \in T_\varepsilon(\bar{x})} \lambda_t \partial \varphi_t(x) \ \text{ and } \ \varphi_t(x) = \varphi_s(x) \ \text{ for all } \ t, s \in \operatorname{supp} \lambda.$$

If $\varphi_t(x) = \varphi_s(x)$ and $t \neq s$, we get from the above that

$$tx_1^3 - \frac{1}{(t+1)^2}|x_2| + t^3 - 1 = sx_1^3 - \frac{1}{(s+1)^2}|x_2| + s^3 - 1,$$

which is equivalent to the equation

$$x_1^3 - \frac{t+s+2}{(t+1)^2(s+1)^2}|x_2| = -t^2 + ts - s^2.$$

When $\varepsilon > 0$ is sufficiently small, this equation has no solution since its left-hand side is close to 0 while the other side is close to -1 for $x \in B_\varepsilon(\bar{x})$ and $t \in T_\varepsilon(\bar{x})$. It follows therefore that $C(\varepsilon) = \bigcup\{\partial\varphi_t(x)|\ t \in T_\varepsilon(\bar{x}),\ x \in B_\varepsilon(\bar{x})\}$ for small ε. Note further that $\partial\varphi_t(x) \subset \{(3tx_1^2, \frac{1}{(t+1)^2}), (3tx_1^2, -\frac{1}{(t+1)^2})\}$, where the equality holds for $x_2 = 0$. Applying Lemma 8.29 yields the representation

$$\bigcap_{\varepsilon>0} \mathrm{cl}\, C(\varepsilon) = \mathrm{cl}\big[\mathop{\mathrm{Lim\,sup}}_{\varepsilon\downarrow 0} C(\varepsilon)\big] = \Big\{\Big(0, \frac{1}{4}\Big), \Big(0, -\frac{1}{4}\Big)\Big\},$$

resulting in a nonconvex set. We also have $\psi(x) = x_1^3 - \frac{1}{4}|x_2|$ for all x around $\bar{x}$. Hence the equality holds in (8.47), and the set $\partial\psi(\bar{x})$ is nonconvex as well.

Now we introduce a subdifferential property for infinite families of functions, which can be viewed as a nonsmooth extension of the uniform strict differentiability exploited in Section 8.1 and, on the other hand, makes them behave similarly to collections of finitely many Lipschitzian functions.

Definition 8.32 (Equicontinuous Subdifferentiability). *We say that the functions $\varphi_t\colon X \to \overline{\mathbb{R}}$, $t \in T$, are* EQUICONTINUOUSLY SUBDIFFERENTIABLE *at $\bar{x}$ if for any weak* neighborhood V^* of $0 \in X^*$ there is $\varepsilon > 0$ such that*

$$\partial\varphi_t(x) \subset \partial\varphi_t(\bar{x}) + V^* \ \textit{for all}\ \ t \in T_\varepsilon(\bar{x}),\ x \in B_\varepsilon(\bar{x}). \tag{8.50}$$

The next proposition shows that property (8.50) holds automatically for the aforementioned classes of functions.

Proposition 8.33 (Sufficient Conditions for Equicontinuous Subdifferentiability). *The functions $\varphi_t(x)$, $t \in T$, are equicontinuously subdifferentiable at $\bar{x}$ if one of the following conditions is satisfied:*

(i) *either the index set T is finite and the functions φ_t are locally Lipschitzian around $\bar{x}$ for all $t \in T$,*

(ii) *or the functions φ_t indexed by an arbitrary set T are uniformly strictly differentiable at $\bar{x}$ in the sense described in* (8.4).

Proof. To justify (i), consider finitely many functions φ_t locally Lipschitzian around $\bar{x}$; they are obviously uniformly Lipschitzian around $\bar{x}$ with some rank K. Take any weak* neighborhood V^* of $0 \in X^*$ and suppose that V^* is convex. If φ_t are not equicontinuously subdifferentiable at $\bar{x}$, then we find by the basic subgradient representation some sequences $\varepsilon_k \downarrow 0$, $x_k \in B_{\varepsilon_k}(\bar{x})$, $u_k \in B_{\varepsilon_k}(x_k)$, $t_k \in T_{\varepsilon_k}(\bar{x})$, $x_k^* \in \partial\varphi_{t_k}(x_k)$, and $u_k^* \in \widehat{\partial}\varphi_{t_k}(u_k)$ such that $x_k^* \notin \partial\varphi_{t_k}(\bar{x}) + V^*$ and $x_k^* \in u_k^* + \frac{V^*}{2}$. Since T is finite, there is a subsequence $\{t_{k_m}\}$ of $\{t_k\}$ whose elements are constant, say $\widehat{t}$. When m is sufficiently large, the norms $\|u_{k_m}^*\|$ are bounded by K. Hence there is a subsequence (no relabeling) of $\{u_{k_m}^*\}$ weak* converging to $u^* \in \partial\varphi_{\widehat{t}}(\bar{x})$. It yields

$$x_{k_m}^* \in u_{k_m}^* + \frac{V^*}{2} \subset u^* + \frac{V^*}{2} + \frac{V^*}{2} \subset \partial\varphi_{\widehat{t}}(\bar{x}) + V^* = \partial\varphi_{t_{k_m}}(\bar{x}) + V^*$$

for large m, which leads us to a contradiction and thus justifies (i).

To prove (ii), fix any $\delta > 0$ such that $\delta\mathbb{B}^* \subset V^*$, where V^* is supposed to be convex. It follows from (8.4) that each function φ_t is strictly differentiable at $\bar{x}$. Thus $\partial\varphi_t(\bar{x}) = \{\nabla\varphi_t(\bar{x})\}$ for each $t \in T$. Moreover, (8.4) allows us to find $\eta > 0$ such that $r(\eta) < \frac{\delta}{2}$. Define $\varepsilon := \frac{\eta}{2}$ and take any $x \in B_\varepsilon(\bar{x})$ and $x_t^* \in \partial\varphi_t(x)$ for some $t \in T_\varepsilon(\bar{x})$. Then there are $x_t \in B_\varepsilon(x)$, $\widehat{x}_t^* \in \widehat{\partial}\varphi_t(x_t)$, and $\varepsilon_t \in (0, \varepsilon)$ such that $x_t^* \in \widehat{x}_t^* + V^*$ and that

$$\varphi_t(u) - \varphi_t(x_t) \geq \langle \widehat{x}_t^*, u - x_t\rangle - \frac{\delta}{2}\|u - x_t\| \;\text{ for all }\; u \in B_{\varepsilon_t}(x_t).$$

Employing (8.4) again, we get the relationship

$$\varphi_t(u) - \varphi_t(x_t) \leq \langle \nabla\varphi_t(\bar{x}), u - x_t\rangle + r(\eta)\|u - x_t\|$$

for all $u \in B_{\varepsilon_t}(x_t) \subset B_{\varepsilon+\varepsilon_t}(\bar{x}) \subset B_\eta(\bar{x})$. Combining the above yields

$$\langle \widehat{x}_t^* - \nabla\varphi_t(\bar{x}), u - x_t\rangle \leq \Big(r(\eta) + \frac{\delta}{2}\Big)\|u - x_t\| \leq \delta\|u - x_t\|$$

for all $u \in B_{\varepsilon_t}(x_t)$, which shows in turn that $\|\widehat{x}_t^* - \nabla\varphi_t(\bar{x})\| \leq \delta$. Hence

$$x_t^* \in \widehat{x}_t^* + V^* \subset \nabla\varphi_t(\bar{x}) + \delta\mathbb{B}^* + V^* \subset \partial\varphi_t(\bar{x}) + V^* + V^*.$$

By the convexity of V^*, we conclude that $\partial\varphi_t(x) \subset \partial\varphi_t(\bar{x}) + 2V^*$ for all $t \in T_\varepsilon(\bar{x})$ and $x \in B_\varepsilon(\bar{x})$ and thus complete the proof of the proposition. △

The equicontinuous subdifferentiability of φ_t allows us to improve the pointbased evaluations of basic subgradients obtained in Theorem 8.30.

Corollary 8.34 (Enhanced Estimates of Basic Subgradients for Supremum Functions Under Equicontinuous Subdifferentiability). *Assuming the equicontinuous subdifferentiability of the functions φ_t at $\bar{x}$ in the setting of Theorem 8.30, we have the inclusion*

$$\partial\psi(\bar{x}) \subset \bigcap_{\varepsilon>0} \operatorname{cl}^* D(\varepsilon), \tag{8.51}$$

where the mapping $D : \mathbb{R}_+ \rightrightarrows X^*$ is defined by

$$D(\varepsilon) := \bigcup\Big\{ \sum_{t\in T_\varepsilon(\bar{x})} \lambda_t \partial\varphi_t(\bar{x}) \Big| \lambda \in \Delta(T_\varepsilon(\bar{x}))\Big\} \text{ for all } \varepsilon \geq 0. \tag{8.52}$$

If in addition the mapping D in (8.52) is weak outer stable at zero, then*

$$\partial\psi(\bar{x}) \subset \mathrm{cl}^*\Big[\bigcup\Big\{ \sum_{t\in T(\bar{x})} \lambda_t \partial\varphi_t(\bar{x}) \Big| \lambda \in \Delta(T(\bar{x}))\Big\}\Big]. \tag{8.53}$$

Proof. Let V^* be an arbitrary convex weak* neighborhood of $0 \in X^*$. Since φ_t are equicontinuously subdifferentiable at $\bar{x}$, there is a number $\bar{\varepsilon} > 0$ such that inclusion (8.50) holds for all positive numbers $\varepsilon < \bar{\varepsilon}$. Employing Theorem 8.30 and Lemma 8.29, we can derive inclusion (8.51) by showing that

$$\operatorname*{Lim\,sup}_{\varepsilon\downarrow 0} C(\varepsilon) \subset \mathrm{cl}^* \operatorname*{Lim\,sup}_{\varepsilon\downarrow 0} D(\varepsilon). \tag{8.54}$$

To proceed with (8.54), pick any x^* from the set on the left-hand side of (8.54) and find sequences $\varepsilon_k \downarrow 0$, $x_k \in B_{\varepsilon_k}(\bar{x})$, $\lambda_k \in \Lambda_{\varepsilon_k}(x_k)$, and $x_k^* \xrightarrow{w^*} x^*$ such that

$$x_k^* \in \sum_{t\in T_{\varepsilon_k}(\bar{x})} \lambda_{k_t} \partial\varphi_t(x_k).$$

Since φ_t are equicontinuously subdifferentiable at $\bar{x}$, the latter yields

$$x_k^* \in \sum_{t\in T_{\varepsilon_k}(\bar{x})} \lambda_{k_t}(\partial\varphi_t(\bar{x}) + V^*) \subset \sum_{t\in T_{\varepsilon_k}(\bar{x})} \lambda_{k_t}\partial\varphi_t(\bar{x}) + \lambda_{k_t} V^* \subset D(\varepsilon_k) + V^*$$

for all large $k \in \mathbb{N}$, and thus there is $u_k^* \in D(\varepsilon_k)$ such that $x_k^* \in u_k^* + V^*$. The uniform boundedness of $D(\varepsilon_k)$ allows us to conclude that $u_k^* \xrightarrow{w^*} u^* \in \operatorname{Lim\,sup}_{\varepsilon\downarrow 0} D(\varepsilon)$ along a subsequence and then get the inclusions

$$x^* \in x_k^* + V^* \subset u_k^* + V^* + V^* \subset u^* + V^* + V^* + V^* \subset \operatorname*{Lim\,sup}_{\varepsilon\downarrow 0} D(\varepsilon) + 3V^*.$$

This means that x^* belongs to the right-hand side of (8.54), and so (8.51) holds. The rest of the proof is similar to Theorem 8.30. △

The next corollary provides a verifiable sufficient condition, which ensures the weak* outer stability of mapping (8.52) at zero and allows us to eliminate the weak* closure in the subdifferential upper estimate (8.53).

Corollary 8.35 (Subdifferential Estimate Without Weak* Closure). *Let φ_t in (8.32) be equicontinuously subdifferentiable at $\bar{x}$, and let the set*

$$\bigcup\Big\{ \sum_{t\in T} \lambda_t \big(\partial\varphi_t(\bar{x}), \varphi_t(\bar{x})\big) \Big| \lambda \in \Delta(T)\Big\}$$

be weak closed in $X^* \times \mathbb{R}$. Then we have the estimate*

$$\partial\psi(\bar{x}) \subset \bigcup\Big\{\sum_{t\in T(\bar{x})}\lambda_t\partial\varphi_t(\bar{x})\Big|\ \lambda\in\Delta(T(\bar{x}))\Big\}.$$

Proof. To justify the claimed inclusion, it is sufficient to prove by (8.53) that mapping (8.52) is weak* outer stable at zero and that the set $D(0)$ is weak* closed under the assumption made. Pick any $x^* \in \text{cl}^*[\text{Lim}\sup_{\varepsilon\downarrow 0} D(\varepsilon)]$ and employ Lemma 8.29. Given $\varepsilon > 0$, this allows us to find a net $(\lambda_\nu)_{\nu\in\mathcal{N}} \subset \Delta(T_\varepsilon(\bar{x}))$ and subgradients $x^*_{\nu_t} \in \partial\varphi_t(\bar{x})$ for each $\nu\in\mathcal{N}$ and $t\in T$ such that

$$x^* = w^* - \lim_{\nu\in\mathcal{N}}\sum_{t\in T_\varepsilon(\bar{x})}\lambda_{\nu_t}x^*_{\nu_t}. \tag{8.55}$$

Since $\text{supp}\,\lambda_\nu \subset T_\varepsilon(\bar{x})$, we observe that

$$\psi(\bar{x}) \ge \sum_{t\in T_\varepsilon(\bar{x})}\lambda_{\nu_t}\varphi_t(\bar{x}) \ge \sum_{t\in T_\varepsilon(\bar{x})}\lambda_{\nu_t}\big(\psi(\bar{x})-\varepsilon\big) = \psi(\bar{x})-\varepsilon.$$

The latter implies by (8.55) that

$$\big(x^*,\psi(\bar{x})\big) \in \text{cl}^*\Big[\bigcup\Big\{\sum_{t\in T}\lambda_t\big(\partial\varphi_t(\bar{x}),\varphi_t(\bar{x})\big)\Big|\ \lambda\in\Delta(T)\Big\}\Big] + \{0\}\times[0,\varepsilon].$$

Letting $\varepsilon\downarrow 0$ in the above, we obtain the relationships

$$\begin{aligned}\big(x^*,\psi(\bar{x})\big) &\in \text{cl}^*\Big[\bigcup\Big\{\sum_{t\in T}\lambda_t\big(\partial\varphi_t(\bar{x}),\varphi_t(\bar{x})\big)\Big|\ \lambda\in\Delta(T)\Big\}\Big]\\ &= \bigcup\Big\{\sum_{t\in T}\lambda_t\big(\partial\varphi_t(\bar{x}),\varphi_t(\bar{x})\big)\Big|\ \lambda\in\Delta(T)\Big\}.\end{aligned}$$

This gives us $\lambda\in\Delta(T)$ with $\big(x^*,\psi(\bar{x})\big)\in\sum_{t\in T}\lambda\big(\partial\varphi_t(\bar{x}),\varphi_t(\bar{x})\big)$ and obviously implies that $0=\sum_{t\in T}\lambda_t\big(\varphi_t(\bar{x})-\psi(\bar{x})\big)$. It shows that $\text{supp}\,\lambda\subset T(\bar{x})$ and that $x^*\in\sum_{t\in T(\bar{x})}\lambda_t\partial\varphi_t(\bar{x})\subset D(0)$, which justify the weak* outer stability of mapping (8.52) at zero. To prove finally that the set $D(0)$ is weak* closed in X^*, take any $u^*\in\text{cl}^*D(0)$ and observe similarly to the above that

$$\big(u^*,\psi(\bar{x})\big)\in\bigcup\Big\{\sum_{t\in T}\lambda_t\big(\partial\varphi_t(\bar{x}),\phi_t(\bar{x})\big)\Big|\ \lambda\in\Delta(T)\Big\}.$$

Therefore we get $u^*\in D(0)$, which verifies that the set $D(0)$ is weak* closed in X^* and thus completes the proof of the corollary. △

We conclude this subsection with yet another consequence of Theorem 8.30 that provides a *precise calculation* of the basic subdifferential of the supremum function (8.32) generated by uniformly strictly differentiable functions.

Corollary 8.36 (Calculating Basic Subgradients for Suprema of Uniformly Strictly Differentiable Functions). *Let the functions φ_t in (8.31) be uniformly strictly differentiable at $\bar{x}$, and let their gradient set $\{\nabla\varphi_t(\bar{x})\}$ be bounded in X^*.*

Then the supremum function (8.32) *is lower regular at* $\bar{x}$ *and its basic subdifferential* $\partial\psi(\bar{x})$ *at this point is calculated by*

$$\partial\psi(\bar{x}) = \bigcap_{\varepsilon>0} \mathrm{cl}^*\mathrm{co}\left\{\nabla\varphi_t(\bar{x})\Big|\ t \in T_\varepsilon(\bar{x})\right\}. \tag{8.56}$$

If in addition the set $\mathrm{co}\big\{\big(\nabla\varphi_t(\bar{x}),\ f_t(\bar{x})\big)\big|\ t \in T\big\}$ *is weak* closed, then*

$$\partial\psi(\bar{x}) = \mathrm{co}\left\{\nabla\varphi_t(\bar{x})\big|\ t \in T(\bar{x})\right\}. \tag{8.57}$$

Proof. The inclusion "$\subset$" in (8.56) follows from Proposition 8.33 and Corollary 8.35. To justify the converse inclusion, take any $\delta > 0$ and pick x^* from the set on the right-hand side of (8.56). Then for each $\varepsilon > 0$, we find a net $(\lambda_\nu)_{\nu\in\mathcal{N}} \in \Delta(T_\varepsilon(\bar{x}))$ ensuring the representation

$$x^* = w^* - \lim_{\nu\in\mathcal{N}} \sum_{t\in T_\varepsilon(\bar{x})} \lambda_{\nu_t}\nabla\varphi_t(\bar{x}).$$

It follows from (8.4) that there is $\eta > 0$ such that

$$\varphi_t(x) - \varphi_t(\bar{x}) \geq \langle\nabla\varphi_t(\bar{x}), x - \bar{x}\rangle - \delta\|x - \bar{x}\| \text{ for all } x \in B_\eta(\bar{x}),\ t \in T.$$

Then we get by the above representation of x^* that

$$\begin{aligned}&\psi(x) - \psi(\bar{x}) + \varepsilon \geq \limsup_{\nu\in\mathcal{N}} \sum_{t\in T_\varepsilon(\bar{x})} \lambda_{\nu_t}\big(\varphi_t(x) - \varphi_t(\bar{x})\big)\\ &\geq \limsup_{\nu\in\mathcal{N}} \sum_{t\in T_\varepsilon(\bar{x})} \lambda_{\nu_t}\big(\langle\nabla\varphi_t(\bar{x}), x - \bar{x}\rangle - \delta\|x - \bar{x}\|\big) \geq \langle x^*, x - \bar{x}\rangle - \delta\|x - \bar{x}\|\end{aligned}$$

whenever $x \in B_\eta(\bar{x})$. Letting now $\varepsilon \downarrow 0$ gives us

$$\psi(x) - \psi(\bar{x}) \geq \langle x^*, x - \bar{x}\rangle - \delta\|x - \bar{x}\| \text{ for all } x \in B_\eta(\bar{x}),$$

which means that $x^* \in \widehat{\partial}\psi(\bar{x})$ and thus yields, by taking into account the inclusion "$\subset$" in (8.56), the validity of the inclusions

$$\partial\psi(\bar{x}) \subset \bigcap_{\varepsilon>0} \mathrm{cl}^*\mathrm{co}\left\{\nabla\varphi_t(\bar{x})\Big|\ t \in T_\varepsilon(\bar{x})\right\} \subset \widehat{\partial}\psi(\bar{x}).$$

Since $\widehat{\partial}\psi(\bar{x}) \subset \partial\psi(\bar{x})$, this shows that ψ is lower regular at $\bar{x}$ and (8.56) holds. Finally, (8.57) follows from Corollary 8.35 by the lower regularity of ψ. △

8.2.3 Optimality Conditions for Lipschitzian SIPs

Here we apply the subdifferential evaluations of the supremum functions (8.32) and subdifferential calculus rules to derive necessary optimality conditions for Lipschitzian SIPs of type (8.31). In this way we obtain *qualified* (nonzero multipliers associated with the cost function) optimality conditions of both *asymptotic* (i.e., with weak* closure) and *KKT* (without it) forms. Define

$$\Omega := \left\{x \in X \middle| \varphi_t(x) \le 0,\ t \in T\right\}$$

and recall our standing assumption that the functions φ_t are uniformly locally Lipschitzian with rank $K > 0$ around the reference point $\bar{x}$. For simplicity we suppose that the *cost* function φ in (8.31) is locally Lipschitzian around $\bar{x}$ too. In what follows we also suppose that $\psi(\bar{x}) = 0$, since the case of $\psi(\bar{x}) < 0$ is trivial. Then we have the expressions

$$T_\varepsilon(\bar{x}) = \left\{t \in T \middle| \varphi_t(\bar{x}) \ge -\varepsilon\right\} \text{ and } T(\bar{x}) = \left\{t \in T \middle| \varphi_t(\bar{x}) = 0\right\}.$$

The first theorem provides several versions of necessary optimality conditions of the KKT type in terms of basic subgradients.

Theorem 8.37 (Qualified Necessary Optimality Conditions via Basic Subgradients). *Let $\bar{x}$ be a local minimizer for (8.31) under the constraint qualification $0 \notin \bigcap_{\varepsilon>0} \operatorname{cl}^* C(\varepsilon)$ with $C(\cdot)$ defined in (8.46). Then we have*

$$0 \in \partial\varphi(\bar{x}) + \mathbb{R}_+ \bigcap_{\varepsilon>0} \operatorname{cl} C(\varepsilon). \tag{8.58}$$

If the mapping $\mathbb{R}_+ C : \mathbb{R}_+ \rightrightarrows X^*$ is weak outer stable at zero, then*

$$0 \in \partial\varphi(\bar{x}) + \operatorname{cl}^*\left\{ \sum_{t \in T(\bar{x})} \lambda_t \partial\varphi_t(\bar{x}) \middle| t \in T(\bar{x})\right\}. \tag{8.59}$$

If furthermore the set $\mathbb{R}_+ C(0)$ is weak closed, then there is a multiplier $\lambda \in \mathbb{R}_+^{(T)}$ such that we have the KKT form*

$$0 \in \partial\varphi(\bar{x}) + \sum_{t \in T(\bar{x})} \lambda_t \partial\varphi_t(\bar{x}). \tag{8.60}$$

Proof. To justify (8.58) under $0 \notin \bigcap_{\varepsilon>0} \operatorname{cl}^* C(\varepsilon)$, consider the function

$$\vartheta(x) := \max\left\{\psi(x) - \psi(\bar{x}), \varphi(x)\right\}, \quad x \in X,$$

defined via (8.32) and observe that $\bar{x}$ is a local minimizer of $\vartheta(\cdot)$ over X. Thus $0 \in \partial\vartheta(\bar{x})$. Applying the maximum rule from Theorem 4.10(ii) and the sum rule for Lipschitzian functions from Corollary 2.20, which both hold without any change in Asplund spaces (see Exercises 4.32 and 2.54(i), respectively), and remembering that $\psi(\bar{x}) = 0$, we find $\mu \in [0, 1]$ such that

$$0 \in \partial\big(\mu\psi + (1-\mu)\varphi\big)(\bar{x}) \subset \mu\partial\psi(\bar{x}) + (1-\mu)\partial\varphi(\bar{x}).$$

Then Theorem 8.30(i) excludes the case of $\mu = 0$ in the above inclusion due to the imposed constraint qualification, and thus we arrive at (8.58). If the mapping $\mathbb{R}_+C : \mathbb{R}_+ \rightrightarrows X^*$ is weak* outer stable at zero, then Lemma 8.29 allows us to deduce from (8.58) that

$$0 \in \partial\psi(\bar{x}) + \text{cl}^*\Big[\mathop{\text{Lim}\,\text{sup}}_{\varepsilon\downarrow 0} \mathbb{R}_+C(\varepsilon)\Big] \subset \partial\psi(\bar{x}) + \text{cl}^*[\mathbb{R}_+C(0)],$$

which justifies (8.59). The KKT condition (8.60) clearly follows from (8.59) provided that the cone $\mathbb{R}_+C(0)$ is weak* closed. △

When the constraint functions in (8.31) are equicontinuously subdifferentiable at the reference point, the results of Theorem 8.37 can be simplified by replacing the set-valued mapping $C(\cdot)$ with that of $D(\cdot)$ defined in (8.52).

Corollary 8.38 (Simplified Necessary Conditions for Equicontinuously Subdifferentiable Functions). *Let φ_t be equicontinuously subdifferentiable at the local minimizer $\bar{x}$ for* (8.31)*, and let $D\colon \mathbb{R}_+ \rightrightarrows X^*$ be defined in* (8.52)*. Then the qualification condition $0 \notin \bigcap_{\varepsilon>0} \text{cl}^* D(\varepsilon)$ implies that*

$$0 \in \partial\varphi(\bar{x}) + \mathbb{R}_+ \bigcap_{\varepsilon>0} \text{cl}^* D(\varepsilon).$$

If we assume in addition that the set

$$Q := \bigcup\Big\{\sum_{t\in T} \lambda_t\big(\partial\varphi_t(\bar{x}), \varphi_t(\bar{x})\big)\Big|\ \lambda \in \mathbb{R}_+^{(T)}\Big\} \ \textit{is weak* closed in } X^* \times \mathbb{R},$$

then there is a multiplier $\lambda \in \mathbb{R}_+^{(T)}$ such that the KKT condition (8.60) *holds.*

Proof. Following the lines in the proof of Corollary 8.35, we get that the mapping $\mathbb{R}_+C : \mathbb{R}_+ \rightrightarrows X^*$ is weak* outer stable at zero and that $\mathbb{R}_+C(0)$ is weak* closed in X^* provided the weak* closedness of Q. This together with Theorem 8.37 justifies the results in this corollary. △

Observe that in the case of linear functions φ_t, the weak* closedness of Q reduces to the Farkas-Minkowski property (7.49). More generally, for uniformly strictly differentiable functions φ_t, the imposed condition on Q is equivalent to NFMCQ introduced in Definition 8.8.

Next we define and employ an extension to the Lipschitzian case of another constraint qualification for SIPs developed in Section 8.1 for smooth functions, namely, PMFCQ from Definition 8.4. The following condition is formulated in terms of the generalized directional derivative (1.77).

Definition 8.39 (Generalized PMFCQ). *We say that SIP* (8.31) *satisfies the* GENERALIZED PERTURBED MANGASARIAN-FROMOVITZ CONSTRAINT QUALIFICATION (GENERALIZED PMFCQ) *at* $\bar{x}$ *if there is* $d \in X$ *such that*

$$\inf_{\varepsilon>0} \sup_{t\in T_\varepsilon(\bar{x})} \varphi_t^\circ(\bar{x}; d) < 0. \tag{8.61}$$

If φ_t are uniformly strictly differentiable at $\bar{x}$, then (8.61) reduces to

$$\inf_{\varepsilon>0} \sup_{t\in T_\varepsilon(\bar{x})} \langle \nabla\varphi_t(\bar{x}), d\rangle < 0 \ \text{ for some } \ d \in \mathbb{R},$$

which is exactly PMFCQ from Definition 8.4 employed in Section 8.1.

The final result of this subsection employs the generalized PMFCQ (8.61) to derive necessary optimality conditions for (8.31) in both asymptotic and KKT forms expressed via the generalized gradient (1.78) in the case of equicontinuously subdifferentiable constraint functions.

Theorem 8.40 (Necessary Optimality Conditions Under the Generalized PMFCQ). *Let* $\bar{x}$ *be a local minimizer of* (8.31)*, and let* φ_t *be equicontinuously subdifferentiable at* $\bar{x}$*. If the generalized PMFCQ holds at* $\bar{x}$*, then*

$$0 \in \overline{\partial}\varphi(\bar{x}) + \mathbb{R}_+ \bigcap_{\varepsilon>0} \mathrm{cl}^*\mathrm{co}\Big[\bigcup \Big\{\overline{\partial}\varphi_t(\bar{x})\Big|\ t \in T_\varepsilon(\bar{x})\Big\}\Big]. \tag{8.62}$$

If furthermore the convex conic hull $\mathrm{cone}\{(\overline{\partial}\varphi_t(\bar{x}), \varphi_t(\bar{x}))|\ t \in T\}$ *is weak** *closed in* $X^* \times \mathbb{R}$*, then there is a multiplier* $\lambda \in \mathbb{R}_+^{(T)}$ *such that*

$$0 \in \overline{\partial}\varphi(\bar{x}) + \sum_{t\in T(\bar{x})} \lambda_t \overline{\partial}\varphi_t(\bar{x}).$$

Proof. Show first that (8.61) can be equivalently written in the dual form

$$0 \notin \bigcap_{\varepsilon>0} \mathrm{cl}^*\mathrm{co}\Big[\bigcup \Big\{\overline{\partial}\varphi_t(\bar{x})\Big|\ t \in T_\varepsilon(\bar{x})\Big\}\Big]. \tag{8.63}$$

Indeed, (8.61) gives us numbers $\varepsilon, \delta > 0$ such that $\sup_{t\in T_\varepsilon(\bar{x})} f^\circ(x; d) < -\delta$. Supposing that condition (8.63) is not satisfied tells us that $0 \in \mathrm{cl}^*\mathrm{co}\big[\cup \{\overline{\partial}\varphi_t(\bar{x})|\ t \in T_{\varepsilon_1}(\bar{x})\}\big]$ for some $\varepsilon_1 \in (0, \varepsilon)$. Then there are nets $(\lambda_\nu)_{\nu\in\mathcal{N}} \in \Delta(T_{\varepsilon_1}(\bar{x}))$ and $x^*_{\nu_t} \in \overline{\partial}\varphi_t(\bar{x})$ as $t \in T_{\varepsilon_1}(\bar{x})$ and $\nu \in \mathcal{N}$ for which

$$0 = w^* - \lim_{\nu\in\mathcal{N}} \sum_{t\in T_{\varepsilon_1}(\bar{x})} \lambda_{\nu_t} x^*_t.$$

This implies by using (1.78) that for the selected direction d in (8.61), we have

$$0 = \lim_{\nu \in \mathcal{N}} \sum_{t \in T_{\varepsilon_1}(\bar{x})} \lambda_{\nu_t} \langle x_t^*, d \rangle \le \limsup_{\nu \in \mathcal{N}} \sum_{t \in T_{\varepsilon_1}(\bar{x})} \lambda_{\nu_t} \varphi_t^\circ(x; d) \le -\delta < 0,$$

a contradiction that justifies the validity of (8.61)⇒(8.63). Assuming now that (8.63) is satisfied, find $\varepsilon > 0$ with $0 \notin \text{cl}^*\text{co}\big[\cup \{\overline{\partial}\varphi_t(\bar{x})|\ t \in T_\varepsilon(\bar{x})\}\big]$. Then the convex separation theorem gives us an element $d \in X$ such that

$$\sup\big\{\langle x^*, d\rangle\big|\ x^* \in \overline{\partial}\varphi_t(\bar{x}),\ t \in T_\varepsilon(\bar{x})\big\} < 0.$$

Since $\sup\{\langle x^*, d\rangle|\ x^* \in \overline{\partial}\varphi_t(\bar{x})\} = \varphi_t^\circ(x; d)$ for any $t \in T_\varepsilon(\bar{x})$, the last inequality ensures that (8.61) is satisfied and thus verifies the claimed equivalence.

Since (8.63) holds, it follows from Exercise 8.97 that $0 \notin \overline{\partial}\psi(\bar{x})$, and so $0 \notin \partial\psi(\bar{x})$. Similarly to the proof of Theorem 8.37, we find $\mu \in \mathbb{R}_+$ with

$$0 \in \partial\varphi(\bar{x}) + \mu\partial\psi(\bar{x}) \subset \overline{\partial}\varphi(\bar{x}) + \mu\overline{\partial}\psi(\bar{x}).$$

Then using again the result of Exercise 8.97 verifies (8.62). Furthermore, the weak* closedness of the convex conic hull cone $\{(\overline{\partial}\varphi_t(\bar{x}), \varphi_t(\bar{x}))|\ t \in T\}$ allows us reducing (8.62) to the KKT optimality condition stated in the theorem by arguments similar to those in the proof of Corollary 8.38. △

8.3 Nonsmooth Cone-Constrained Optimization

In this section we explore a different approach in comparison with that in Section 8.2 to the study of Lipschitzian SIPs. It involves reducing SIPs to problems of *cone-constrained* (or *conic*) *programming* in *infinite-dimensional* spaces, even if the decision space in finite-dimensional. The latter class of optimization problems of their strong independent interest is formalized as

$$\begin{cases} \text{minimize } \ \varphi(x) \ \text{ subject to} \\ f(x) \in -\Theta \subset Y, \ x \in \Omega \subset X, \end{cases} \tag{8.64}$$

where Θ is *closed* and *convex cone*, which is the standing assumption in this section. We show below that Lipschitzian SIPs considered in Section 8.2, with the additional presence of the geometric constraint $x \in \Omega$, can be written in form (8.64), where the cone Θ is given by either $\Theta = \mathcal{C}_+(T)$ or $\Theta = l_+^\infty(T)$, i.e., it is a collection of either positive continuous or essentially bounded functions over a compact or noncompact index set T, respectively. We have already dealt with these spaces in Chapter 7 while investigating SIPs with linear and convex data by using approaches that are completely different from those which we are going to employ in what follows.

8.3.1 Subgradients of Scalarized Supremum Functions

Besides the assumption on Θ made above and our standing l.s.c. assumption on $\varphi\colon X \to \overline{\mathbb{R}}$, we suppose in what follows that X is *Asplund*, Y is arbitrary *Banach*, and the mapping $f\colon X \to Y$ is *locally Lipschitzian* around the reference point $\bar{x}$, i.e., there are constants $K, \rho > 0$ such that

$$\|f(x) - f(u)\| \le K\|x - u\| \ \text{ for all } \ x, u \in B_\rho(\bar{x}). \tag{8.65}$$

We now show that the *conic constraint* $f(x) \in -\Theta$ in (8.64) can be rewritten in the inequality form via a certain *scalarized supremum function*.

Proposition 8.41 (Conic Constraints via Supremum Functions). *We have the following conic constraint representation:*

$$\big\{x \in X \big|\ f(x) \in -\Theta\big\} = \big\{x \in X \big|\ \vartheta(x) \le 0\big\},$$

where the scalarized supremum function $\vartheta\colon X \to \overline{\mathbb{R}}$ *is defined by*

$$\begin{aligned}&\vartheta(x) := \sup_{y^* \in \Xi} \langle y^*, f(x)\rangle \ \textit{ with}\\ &\Xi := \big\{y^* \in Y^* \big|\ \|y^*\| = 1,\ \langle y^*, y\rangle \ge 0,\ y \in \Theta\big\}.\end{aligned} \tag{8.66}$$

Proof. We obviously get $\langle y^*, f(x)\rangle \le 0$ for all $y^* \in \Xi$ if the inclusion $f(x) \in -\Theta$ holds. To verify the converse implication, suppose that it fails and then find by convex separation such nonzero elements $\bar{y}^* \in Y^* \setminus \{0\}$ and $\gamma > 0$ that

$$\langle \bar{y}^*, f(x)\rangle > \gamma > 0 \ge \langle \bar{y}^*, y\rangle \ \text{ for all } \ y \in -\Theta.$$

This yields $\bar{y}^*\|\bar{y}^*\|^{-1} \in \Xi$, and hence we arrive at the contradiction

$$0 \ge \big\langle \bar{y}^*\|\bar{y}^*\|^{-1}, f(x)\big\rangle > \gamma\|\bar{y}^*\|^{-1} > 0,$$

which thus completes the proof of the proposition. △

Note that the scalarized supremum function ϑ is significantly different from the supremum function ψ considered in the SIP framework (8.32).

The main goal of this subsection is to evaluate *basic subgradients* for a more general class of scalarized supremum functions defined by

$$\psi(x) := \sup_{y^* \in \Lambda} \big\langle y^*, f(x)\big\rangle, \tag{8.67}$$

where Λ is an arbitrary nonempty subset of the *positive polar cone*

$$\Theta^+ := \big\{y^* \in Y^* \big|\ \big\langle y^*, y\big\rangle \ge 0 \ \text{ for all } \ y \in \Theta\big\}. \tag{8.68}$$

Since $\Xi \subset \Theta^+$ for the set Ξ in Proposition 8.41, the results obtained below for (8.67) immediately apply to the function ϑ therein.

Our first result provides a "fuzzy" upper estimate for *basic* subgradients of the scalarized supremum function (8.67) at the reference point $\bar{x}$ via *regular* subgradients of the scalarization $x \mapsto \langle y^*, f\rangle(x)$ at some neighboring points.

Theorem 8.42 (Fuzzy Estimate of Basic Subgradients for Scalarized Supremum Functions). *Let* $\bar{x} \in \operatorname{dom}\psi$ *for function* (8.67)*, and let* V^* *be a weak* neighborhood of* $0 \in X^*$. *Then for any* $x^* \in \partial\psi(\bar{x})$ *and any* $\varepsilon > 0$, *there exist* $x_\varepsilon \in B_\varepsilon(\bar{x})$ *and* $y_\varepsilon^* \in \operatorname{co}\Lambda$ *with* $|\langle y_\varepsilon^*, f(x_\varepsilon)\rangle - \psi(\bar{x})| < \varepsilon$ *such that*

$$x^* \in \widehat{\partial}\langle y_\varepsilon^*, f\rangle(x_\varepsilon) + V^*. \tag{8.69}$$

Proof. Fix arbitrary $x^* \in \partial\psi(\bar{x})$ and $\varepsilon > 0$. It is easy to check that each function $\langle y^*, f(x)\rangle$ is locally Lipschitzian around $\bar{x}$ with the same constants K and ρ as in (8.65) for all $y^* \in \Lambda$ and so is the scalarized supremum function ψ. Without loss of generality, assume that V^* is convex and that $\varepsilon \le \rho$. Then find $k \in \mathbb{N}$, $\varepsilon_k > 0$, and $x_j \in X$ for $j = 1, \ldots, k$ such that

$$\bigcap_{j=1}^{k} \left\{ v^* \in X^* \,\middle|\, \langle v^*, x_j\rangle < \varepsilon_k \right\} \subset \frac{1}{4}V^*.$$

Consider further a finite-dimensional subspace $L \subset X$ by $L := \operatorname{span}\{x_1, \ldots, x_k\}$ and observe that $L^\perp := \{v^* \in X^* |\ \langle v^*, x\rangle = 0,\ x \in L\} \subset \frac{1}{4}V^*$. By the choice of x^*, we find $\widehat{x} \in \operatorname{dom}\psi \cap B_{\frac{\varepsilon}{2}}(\bar{x})$ and $u^* \in X^*$ such that $|\psi(\widehat{x}) - \psi(\bar{x})| \le \frac{\varepsilon}{2}$, $u^* \in \widehat{\partial}\psi(\widehat{x})$ and that $x^* \in u^* + \frac{V^*}{4}$. Fix $\delta > 0$ satisfying

$$4\delta \le \varepsilon, \quad \frac{12\delta}{1-2\delta}\mathbb{B}^* \subset V^*, \quad \text{and} \quad \frac{16\delta}{1-2\delta}\|u^*\|\mathbb{B}^* \subset V^*. \tag{8.70}$$

Since $u^* \in \widehat{\partial}\psi(\widehat{x})$, there is some number $\eta \in (0, \delta)$ with

$$\psi(x) - \psi(\widehat{x}) + \delta\|x - \widehat{x}\| \ge \langle u^*, x - \widehat{x}\rangle \ \text{ for all } \ x \in B_\eta(\widehat{x}) \subset B_\rho(\bar{x}).$$

This implies that $(\widehat{x}, \psi(\widehat{x}))$ is a *local minimizer* of the following problem:

$$\begin{cases} \text{minimize } \ r + \delta\|x - \widehat{x}\| - \langle u^*, x - \widehat{x}\rangle - \psi(\widehat{x}) \ \text{ subject to} \\ \langle y^*, f(x)\rangle - r \le 0 \ \text{ for } \ y^* \in \Lambda \ \text{ and } (x, r) \in B_\eta(\widehat{x}) \times \mathbb{R}. \end{cases}$$

Define $A := (L \cap B_\eta(\widehat{x})) \times [\psi(\widehat{x}) - 1, \psi(\widehat{x}) + 1]$, $\Psi(x, r) := r + \delta\|x - \widehat{x}\| - \langle u^*, x - \widehat{x}\rangle - \psi(\widehat{x})$, and $\varphi_{y^*} : X \times \mathbb{R} \to \overline{\mathbb{R}}$ by

$$\varphi_{y^*}(x, r) := \langle y^*, f(x)\rangle - r \ \text{ for all } \ y^* \in \Lambda \ \text{ and } \ (x, r) \in X \times \mathbb{R}.$$

It readily follows from these constructions that

$$\big\{(x,r)\in A\big|\ \Psi(x,r)+\eta^2\le 0\big\}\subset\bigcup_{y^*\in\Lambda}\big\{(x,r)\in A\big|\ \varphi_{y^*}(x,r)>0\big\}.$$

The set on the left-hand side above is clearly compact in the finite-dimensional space $L\times\mathbb{R}$, and each subset $\{(x,r)\in A|\ \varphi_{y^*}(x,r)>0\}$ is open in A due to the Lipschitz continuity of φ_{y^*} on the set $B_\rho(\bar{x})\times\mathbb{R}$, which is larger than A. Thus we find a *finite* subset $\Upsilon\subset\Lambda$ containing $s\in\mathbb{N}$ elements so that

$$\big\{(x,r)\in A\big|\ \Psi(x,r)+\eta^2\le 0\big\}\subset\bigcup_{y^*\in\Upsilon}\Big\{(x,r)\in A\Big|\ \varphi_{y^*}(x,r)>0\Big\}.$$

Based on this, we arrive at the relationships

$$\Psi(x,r)+\eta^2\ge 0=\Psi\big(\widehat{x},\psi(\widehat{x})\big)\ \text{if}\ (x,r)\in\widetilde{A}:=\big\{(x,r)\in A\big|\varphi_{y^*}(x,r)\le 0,\ y^*\in\Upsilon\big\},$$

where $\widetilde{A}$ is a closed subset of $B_\rho(\bar{x})\times\mathbb{R}$. Using now Ekeland's variational principle gives us $(\widetilde{x},\widetilde{r})\in\widetilde{A}$ with $\|\widetilde{x}-\widehat{x}\|+|\widetilde{r}-\varphi(\widehat{x})|\le\frac{\eta}{2}$ and

$$\Psi(x,r)+2\eta(\|x-\widetilde{x}\|+|r-\widetilde{r}|)\ge\Psi(\widetilde{x},\widetilde{r})\ \ \text{whenever}\ \ (x,r)\in\widetilde{A}.$$

The latter means that $(\widetilde{x},\widetilde{r})$ is a local optimal solution to the problem

$$\begin{cases}\text{minimize}\ \widetilde{\Psi}(x,r):=\Psi(x,r)+2\eta\big(\|x-\widetilde{x}\|+|r-\widetilde{r}|\big)\ \ \text{subject to}\\ \varphi_{y^*}(x,r)\le 0\ \ \text{for}\ \ y^*\in\Upsilon\ \ \text{and}\ \ (x,r)\in A.\end{cases}$$

It is obvious that the functions $\widetilde{\Psi}(\cdot,\cdot)$ and $\varphi_{y^*}(\cdot,\cdot)$ are Lipschitz continuous around $(\widetilde{x},\widetilde{r})$ for all $y^*\in\Upsilon$. Applying now to this problem with $s\in\mathbb{N}$ elements in the set Υ the necessary optimality conditions from Theorem 6.5(ii) that hold in any Asplund space, we find multipliers $\lambda_0,\ldots,\lambda_s\ge 0$, not equal to zero simultaneously, and dual elements $y_1^*,\ldots,y_s^*\in\Upsilon(\widetilde{x},\widetilde{r}):=\{y^*\in\Upsilon|\ \varphi_{y^*}(\widetilde{x},\widetilde{r})=0\}$ satisfying the inclusion

$$(0,0)\in\partial\Big(\lambda_0\widetilde{\Psi}+\sum_{m=1}^{s}\lambda_m\varphi_{y_m^*}\Big)(\widetilde{x},\widetilde{r})+N\big((\widetilde{x},\widetilde{r});A\big).$$

Since $(\widetilde{x},\widetilde{r})\in\text{int}(B_\eta(\widehat{x})\times[\psi(\widehat{x})-1,\psi(\widehat{x})+1])$, it follows that

$$\begin{aligned}&(0,0)\in\partial\Big(\lambda_0\widetilde{\Psi}+\sum_{m=1}^{s}\lambda_m\varphi_{y_m^*}\Big)(\widetilde{x},\widetilde{r})+N\big((\widetilde{x},\widetilde{r});(L\cap B_\eta(\widehat{x}))\times[\psi(\widehat{x})-1,\psi(\widehat{x})+1]\big)\\ &=\partial\Big(\lambda_0\widetilde{\Psi}+\sum_{m=1}^{s}\lambda_m\varphi_{y_m^*}\Big)(\widetilde{x},\widetilde{r})+N(\widetilde{x};L)\times\{0\}\\ &\subset\partial\Big(\lambda_0\widetilde{\Psi}+\sum_{m=1}^{s}\lambda_m\varphi_{y_m^*}\Big)(\widetilde{x},\widetilde{r})+L^\perp\times\{0\}.\end{aligned}$$

If $\lambda_0 = 0$ therein, we have the inclusion

$$(0,0) \in \partial\Big(\sum_{m=1}^{s} \lambda_m \langle y_m^*, f\rangle\Big)(\widetilde{x}) \times \Big\{-\sum_{m=1}^{s} \lambda_m\Big\} + L^\perp \times \{0\},$$

which implies in turn that $\sum_{m=1}^{s} \lambda_m = 0$, i.e., $\lambda_m = 0$ for all $m = 0, \ldots, s$. This contradiction shows that $\lambda_0 \neq 0$. Thus we put $\lambda_0 = 1$ and then obtain

$$\begin{aligned}(u^*, 0) \in \partial\Big(\sum_{m=1}^{s} \lambda_m \langle y_m^*, f\rangle\Big)(\widetilde{x}) \times \Big\{1 - \sum_{m=1}^{s} \lambda_m\Big\}\\ +(\delta + 2\eta)\mathbb{B}^* \times 2[-\eta, \eta] + L^\perp \times \{0\}.\end{aligned}$$

Define $\widetilde{\lambda} := \sum_{m=1}^{s} \lambda_m$, $\widetilde{\lambda}_m := \widetilde{\lambda}^{-1}\lambda_m$, and $\widetilde{u}^* := \widetilde{\lambda}^{-1}u^*$. Since $|1 - \widetilde{\lambda}| \le 2\eta < 2\delta$ by the last inclusion, we divide its both sides by $\widetilde{\lambda}$ and get

$$\begin{aligned}\widetilde{u}^* \in \partial\Big(\sum_{m=1}^{s} \widetilde{\lambda}_m \langle y_m^*, f\rangle\Big)(\widetilde{x}) + \frac{\delta + 2\eta}{\widetilde{\lambda}}\mathbb{B}^* + \frac{L^\perp}{\widetilde{\lambda}} \subset \partial\Big(\sum_{m=1}^{s} \langle \widetilde{\lambda}_m y_m^*, f\rangle\Big)(\widetilde{x})\\ + \frac{3\delta}{1 - 2\delta}\mathbb{B}^* + L^\perp \subset \partial\Big(\sum_{m=1}^{s} \langle \widetilde{\lambda}_m y_m^*, f\rangle\Big)(\widetilde{x}) + \frac{V^*}{4} + \frac{V^*}{4} \subset \partial\langle y_\varepsilon^*, f\rangle(\widetilde{x}) + \frac{V^*}{2}\end{aligned}$$

with $y_\varepsilon^* := \sum_{m=1}^{s} \widetilde{\lambda}_m y_m^* \in \mathrm{co}\,\Upsilon \subset \mathrm{co}\,\Lambda$ by taking into account (8.70), the above constructions of L and Ψ, and the estimate of $L^\perp$. Thus there is a basic subgradient $v^* \in \partial\langle y_\varepsilon^*, f\rangle(\widetilde{x})$ satisfying $\widetilde{u}^* \in v^* + \frac{V^*}{2}$. The basic subdifferential representation in Asplund spaces from Exercise 1.65(ii) allows us to find elements $x_\varepsilon \in B_\delta(\widetilde{x})$ and $w^* \in \widehat{\partial}\langle y_\varepsilon^*, f\rangle(x_\varepsilon)$ such that $|\langle y_\varepsilon^*, f(x_\varepsilon)\rangle - \langle y_\varepsilon^*, f(\widetilde{x})\rangle| \le \delta$ and $v^* \in w^* + \frac{V^*}{8}$. We clearly have the inequalities

$$\|x_\varepsilon - \bar{x}\| \le \|x_\varepsilon - \widetilde{x}\| + \|\widetilde{x} - \widehat{x}\| + \|\widehat{x} - \bar{x}\| \le \delta + \delta + \frac{\varepsilon}{2} \le \varepsilon.$$

Moreover, the following estimates hold:

$$\begin{aligned}|\langle y_\varepsilon^*, f(x_\varepsilon)\rangle - \psi(\bar{x})| \le |\langle y_\varepsilon^*, f(x_\varepsilon) - f(\widetilde{x})\rangle| + |\langle y_\varepsilon^*, f(\widetilde{x})\rangle - \widetilde{r}| + |\widetilde{r} - \psi(\widehat{x})|\\ +|\psi(\widehat{x}) - \psi(\bar{x})| \le \delta + \Big|\sum_{m=1}^{s} \widetilde{\lambda}_m \langle y_m^*, f(\widetilde{x})\rangle - \widetilde{r}\Big| + \frac{\eta}{2} + \frac{\varepsilon}{2} = \delta + \frac{\eta}{2} + \frac{\varepsilon}{2} \le \varepsilon\end{aligned}$$

by taking into account that $\langle y_m^*, f(\widetilde{x})\rangle = \widetilde{r}$ as $m = 1, \ldots, s$. Note further that

$$\|u^* - \widetilde{u}^*\| = \frac{1 - \widetilde{\lambda}}{\widetilde{\lambda}}\|u^*\| \le \frac{2\eta}{1 - 2\eta}\|u^*\| \le \frac{2\delta}{1 - 2\delta}\|u^*\|,$$

which implies the chain of inclusions:

$$x^* \in u^* + \frac{V^*}{4} \subset \widetilde{u}^* + \frac{2\delta}{1-2\delta}\|u^*\|\mathbb{B}^* + \frac{V^*}{4} \subset v^* + \frac{V^*}{2} + \frac{V^*}{8} + \frac{V^*}{4}$$

$$\subset w^* + \frac{V^*}{8} + \frac{V^*}{2} + \frac{V^*}{8} + \frac{V^*}{4} \subset \widehat{\partial}\langle y_\varepsilon^*, f\rangle(x_\varepsilon) + V^*.$$

This verifies (8.69) and thus completes the proof of the theorem. △

Our next result provides *pointbased* upper estimates of the basic subdifferential of the scalarized supremum function (8.67) involving only the reference point $\bar{x}$. To proceed, define the *partial order* $\leq_\Theta$ on Y generated by a closed and convex ordering cone $\Theta \subset Y$ as follows:

$$y_1 \leq_\Theta y_2 \quad \text{if and only if} \quad y_2 - y_1 \in \Theta \quad \text{for } \ y_1, y_2 \in Y. \tag{8.71}$$

The Θ-*epigraph* of $f\colon X \to Y$ with respect to the order $\leq_\Theta$ is given by

$$\text{epi}_\Theta f := \big\{(x,y) \in X \times Y \big|\ f(x) \leq_\Theta y\big\}.$$

Recall also that f is Θ-*convex* if for any $x_1, x_2 \in X$ and $t \in [0,1]$ we have

$$f\big(tx_1 + (1-t)x_2\big) \leq_\Theta tf(x_1) + (1-t)f(x_2),$$

which is equivalent to the fact that the set $\text{epi}_\Theta f$ is convex in $X \times Y$.

Definition 8.43 **(Θ-Coderivatives).** *Under the standing assumption on the mapping $f\colon X \to Y$ and the cone $\Theta \subset Y$, we define the following:*

(i) *The* REGULAR Θ-CODERIVATIVE *of f at $\bar{x}$ is*

$$\widehat{D}_\Theta^* f(\bar{x})(y^*) := \Big\{x^* \in X^* \Big|\ \limsup_{(x,y)\overset{\text{epi}_\Theta f}{\to}(\bar{x}, f(\bar{x}))} \frac{\langle x^*, x-\bar{x}\rangle - \langle y^*, y - f(\bar{x})\rangle}{\|x-\bar{x}\| + \|y - f(\bar{x})\|} \leq 0\Big\}.$$

(ii) *The (sequential)* NORMAL Θ-CODERIVATIVE *of f at $\bar{x}$ is*

$$D_{N,\Theta}^* f(\bar{x})(y^*) := \Big\{x^* \in X^* \Big|\ \exists \text{ sequences } \ x_k \to \bar{x},\ x_k^* \in \widehat{D}_\Theta^* f(x_k)(y_k^*) \text{ such that } \ (x_k^*, y_k^*) \overset{w^*}{\to} (x^*, y^*)\Big\}.$$

(iii) *The* TOPOLOGICAL NORMAL Θ-CODERIVATIVE *of f at $\bar{x}$ is*

$$\widetilde{D}_{N,\Theta}^* f(\bar{x})(y^*) := \Big\{x^* \in X^* \Big|\ \exists \text{ nets } \ x_\alpha \to \bar{x},\ x_\alpha^* \in \widehat{D}_\Theta^* f(x_\alpha)(y_\alpha^*) \text{ such that } \ (x_\alpha^*, y_\alpha^*) \overset{w^*}{\to} (x^*, y^*)\Big\}.$$

(iv) *The* CLUSTER NORMAL Θ-CODERIVATIVE *of* f *at* $\bar{x}$ *is*

$$\breve{D}^*_{N,\Theta} f(\bar{x})(y^*) := \Big\{x^* \in X^* \Big| \ \exists \text{ sequences } x_k \to \bar{x},\ x^*_k \in \widehat{D}^*_\Theta f(x_k)(y^*_k) \text{ such that } (x^*, y^*) \text{ is a weak}^* \text{ cluster point of } (x^*_k, y^*_k)\Big\}.$$

Note that the limiting procedures employed in Definition 8.43(ii,iii) are similar to those used for mappings without any ordering structure (cf. Chapter 1) while we don't consider here the "mixed" coderivative counterparts. However, the one suggested in (iv) seems to be new even in the nonordering setting while being important for our results on cone-constrained problems with general Banach image spaces Y and their applications to SIPs.

Observe that $\operatorname{dom} \widehat{D}^*_\Theta f(x) \subset \Theta^+$ for any $x \in X$, where Θ^+ stands for the positive polar cone (8.68) to Θ. Since Θ^+ is a weak* closed subset of Y^*, it follows from the inclusion above that the domains $\operatorname{dom} D^*_N f(\bar{x})$, $\operatorname{dom} \widetilde{D}^*_{N,\Theta} f(\bar{x})$, and $\operatorname{dom} \breve{D}^*_{N,\Theta} f(\bar{x})$ are also subsets of Θ^+. It is easy to check that for mappings $f : X \to Y$ locally Lipschitzian around $\bar{x}$ we have the *scalarization formula*

$$\widehat{D}^*_\Theta f(\bar{x})(y^*) := \widehat{\partial}\langle y^*, f\rangle(\bar{x}) \ \text{ if and only if } \ y^* \in \Theta^+, \tag{8.72}$$

where $\langle y^*, f\rangle(x) := \langle y^*, f(x)\rangle$. However, such a scalarization for the limiting coderivatives $D^*_{N,\Theta}$, $\widetilde{D}^*_{N,\Theta}$, and $\breve{D}^*_{N,\Theta}$ requires stronger Lipschitzian assumptions; see Exercise 8.103(ii) for mappings with values in spaces without ordering. The following limiting counterparts of scalarization, which can be proved similarly to [522, Theorem 1.90], are needed below: for all $y^* \in \Theta^+$ we have

$$D^*_{N,\Theta} f(\bar{x})(y^*) = \widetilde{D}^*_{N,\Theta} f(\bar{x})(y^*) = \breve{D}^*_{N,\Theta} f(\bar{x})(y^*) = \big\{\nabla f(\bar{x})^* y^*\big\} \tag{8.73}$$

provided that f is strictly differentiable at $\bar{x}$. Furthermore, it can be derived directly from the constructions above that

$$D^*_\Theta f(\bar{x})(y^*) = \partial\langle y^*, f\rangle(\bar{x}), \quad y^* \in \Theta^+, \tag{8.74}$$

for all the Θ-coderivatives in Definition 8.43 if f is Θ-convex.

Now we are ready to establish the aforementioned pointbased estimates for the basic subdifferential $\partial\psi(\bar{x})$ for the scalarized supremum function (8.67).

Theorem 8.44 (Pointbased Estimates of Basic Subgradients of Scalarized Supremum Functions via Coderivatives). *In the setting of Theorem 8.42, assume that Λ is bounded in Y^*. Then the basic subdifferential of ψ at $\bar{x}$ is upper estimated by*

$$\partial\psi(\bar{x}) \subset \big\{x^* \in \widetilde{D}^*_{N,\Theta} f(\bar{x})(y^*) \big| \ y^* \in \operatorname{cl}^*\operatorname{co}\Lambda,\ \langle y^*, f(\bar{x})\rangle = \psi(\bar{x})\big\} \tag{8.75}$$

via the topological Θ-coderivative of f at $\bar{x}$. If $\dim X < \infty$*, we have*

$$\partial\psi(\bar{x}) \subset \big\{x^* \in \breve{D}^*_{N,\Theta} f(\bar{x})(y^*) \big| \ y^* \in \operatorname{cl}^*\operatorname{co}\Lambda,\ \langle y^*, f(\bar{x})\rangle = \psi(\bar{x})\big\} \tag{8.76}$$

via the cluster Θ*-coderivative counterpart. If in addition the closed unit ball* $\mathbb{B}^*$ *is weak** *sequentially compact in* Y^**, then the cluster* Θ*-coderivative can be replaced in* (8.76) *by its sequential counterpart* $D^*_{N,\Theta}f(\bar{x})(y^*)$.

Proof. To justify (8.75), we first construct a filter $\{V^*_\alpha\}_{\alpha\in A}$ of neighborhoods of $0 \in X^*$ and a net $\{\varepsilon_\alpha\}_{\alpha\in A} \subset \mathbb{R}_+$ such that $\varepsilon_\alpha \downarrow 0$. Let $\mathcal{N}_{X^*}$ be the set of weak* neighborhoods of $0 \in X^*$, and let the A be bijective with $\mathcal{N}_{X^*}$. Denote this bijective correspondence by $\mathcal{N}_{X^*} = \{V^*_\alpha|\ \alpha \in A\}$ and observe that A is a *directed set*, where the direction is given by $\alpha \succeq \beta$ if and only if V^*_α is contained in V^*_β. Fix any $v^* \in X^*$ with $\|v^*\| = 1$ and define

$$\varepsilon_\alpha := \sup\big\{r \in [0,\rho)\big|\ rv^* \in V^*_\alpha\big\}\ \text{ for all }\ \alpha \in A,$$

where ρ is taken from (8.65). Note that $\varepsilon_\alpha > 0$ for all $\alpha \in A$ and that $\varepsilon_\alpha \downarrow 0$. Indeed, for any $\alpha \in A$, there is $\delta \in (0,\rho)$ sufficiently small to get $\delta\mathbb{B}^* \subset V^*_\alpha$. It is obvious that $\varepsilon_\alpha > \delta$. Furthermore, for any $\varepsilon > 0$, the existence of $\alpha_0 \in A$ with $\varepsilon_{\alpha_0} < \varepsilon$ implies that $\varepsilon_\alpha < \varepsilon$ for all $\alpha \succeq \alpha_0$ by the definition of A. Hence if the net $\{\varepsilon_\alpha\}$ doesn't converge to 0, there is $\varepsilon > 0$ with $\varepsilon_\alpha > \varepsilon$ for all $\alpha \in A$, which yields $\varepsilon v^* \in V^*_\alpha$ as $\alpha \in A$. This contradiction justifies $\varepsilon_\alpha \downarrow 0$.

Now pick an arbitrary basic subgradient $x^* \in \partial\psi(\bar{x})$. Employing Theorem 8.42 for any $\alpha \in A$ allows us to find $x_\alpha \in B_{\varepsilon_\alpha}(\bar{x})$ and $y^*_\alpha \in \text{co}\,\Lambda$ with

$$x^* \in \widehat{\partial}\langle y^*_\alpha, f\rangle(x_\alpha) + V^*_\alpha\ \text{ and }\ |\langle y^*_\alpha, f(x_\alpha)\rangle - \psi(\bar{x})| \le \varepsilon_\alpha.$$

By the scalarization formula (8.72), we get $u^*_\alpha \in \widehat{\partial}\langle y^*_\alpha, f\rangle(x_\alpha) = \widehat{D}^*_\Theta f(x_\alpha)(y^*_\alpha)$ and $v^*_\alpha \in V^*_\alpha$ with $x^* = u^*_\alpha + v^*_\alpha$. Since the filter $\{V^*_\alpha\}_{\alpha\in A}$ weak* converges to zero, the directed net $\{v^*_\alpha\}_{\alpha\in A}$ weak* converges to zero as well. This implies that $u^*_\alpha \xrightarrow{w^*} x^*$. Due to the boundedness of $\text{co}\,\Lambda \subset Y^*$, the classical Alaoglu-Bourbaki theorem allows us to find a subnet of $\{y^*_\alpha\}_{\alpha\in A}$ (no relabeling) that weak* converges to some $y^* \in \text{cl}^*\text{co}\,\Lambda$. This yields $x^* \in \widetilde{D}^*_{N,\Theta}f(\bar{x})(y^*)$. Moreover, by $\varepsilon_\alpha \downarrow 0$, $x_\alpha \to \bar{x}$, and $y^*_\alpha \xrightarrow{w^*} y^*$, we have

$$0 = \lim \varepsilon_\alpha = \lim\langle y^*_\alpha, f(x_\alpha)\rangle - \psi(\bar{x}) = \langle y^*, f(\bar{x})\rangle - \psi(\bar{x}),$$

which thus justifies the validity of (8.75) via the topological coderivative.

When $\dim X < \infty$, we can choose $\widetilde{\mathcal{N}}_{X^*} := \{\frac{1}{k}\mathbb{B}^*|\ k \in \mathbb{N}\}$ instead of $\mathcal{N}_{X^*}$ in the proof above, then find $A = \mathbb{N}$ and a *sequence* $\varepsilon_k \in (0,\rho)$ such that $\varepsilon_k \downarrow 0$. Following similar arguments gives us estimate (8.76) via the cluster coderivative $\breve{D}^*f(\bar{x})(y^*)$. Finally, assuming the weak* sequential compactness of the unit ball $\mathbb{B}^* \subset Y^*$ ensures that all the limiting elements of $\breve{D}^*_{N,\Theta}f(\bar{x})(y^*)$ belong to $D^*_{N,\Theta}f(\bar{x})(y^*)$ and thus completes the proof. △

8.3.2 Pointbased Optimality and Qualification Conditions

It follows from Proposition 8.41 that the original cone-constrained optimization problem (8.64) can be equivalently represented as

$$\text{minimize } \varphi(x) \text{ subject to } \vartheta(x) \le 0,\ x \in \Omega$$

via the scalarized supremum function ϑ taken from (8.66). We now use this representation together with the subdifferential estimates of Theorem 8.44 and generalized differential calculus of variational analysis to derive *pointbased* necessary optimality conditions for (8.64) expressed in terms of the above limiting constructions under appropriate constraint qualifications.

The following theorem presents the main results of this subsection.

Theorem 8.45 (Necessary Optimality Conditions for Cone-Constrained Programs). *Let $\bar{x}$ be a local optimal solution to problem (8.64) under our standing assumptions. Suppose also that either the function φ is SNEC at $\bar{x}$ or the set Ω is SNC at this point and that the qualification condition*

$$\partial^\infty \varphi(\bar{x}) \cap \big(-N(\bar{x};\Omega)\big) = \{0\} \tag{8.77}$$

is satisfied; both SNEC and (8.77) are automatic when φ is locally Lipschitzian around $\bar{x}$. Then we have the following assertions:

(a) *either there exists $y^* \in \Theta^+$ such that*

$$0 \in \partial\varphi(\bar{x}) + \widetilde{D}^*_{N,\Theta} f(\bar{x})(y^*) + N(\bar{x};\Omega) \ \text{ and } \ \langle y^*, f(\bar{x})\rangle = 0, \tag{8.78}$$

(b) *or there exists $y^* \in \mathrm{cl}^*\mathrm{co}\,\Xi$ such that*

$$0 \in \partial^\infty\varphi(\bar{x}) + \widetilde{D}^*_{N,\Theta} f(\bar{x})(y^*) + N(\bar{x};\Omega) \ \text{ and } \ \langle y^*, f(\bar{x})\rangle = 0. \tag{8.79}$$

If $\dim X < \infty$, *the above holds with replacing $\widetilde{D}^*_N f(\bar{x})$ by $\breve{D}^*_N f(\bar{x})$. If furthermore $\mathbb{B}^* \subset Y^*$ is weak* sequentially compact, then the topological coderivative $\widetilde{D}^*_{N,\Theta} f(\bar{x})$ can be replaced in (8.78) and (8.79) by the sequential one $D^*_N f(\bar{x})$.*

Proof. Observe first that the validity of both the SNEC property of φ at $\bar{x}$ and the qualification condition (8.77) for local Lipschitzian cost functions φ on Asplund spaces follows from Exercises 2.49 and 4.34, respectively. Further, it is easy to see that $\bar{x}$ is a local optimal solution to the unconstrained problem of minimizing the maximum function

$$\Psi(x) := \max\Big\{\big(\varphi + \delta(\cdot;\Omega)\big)(x) - \varphi(\bar{x}), \vartheta(x)\Big\}, \quad x \in X, \tag{8.80}$$

where $\vartheta(x)$ is given in (8.66), and where Ψ is obviously l.s.c. around $\bar{x}$. If $\vartheta(\bar{x}) < 0$, then there is a neighborhood U of $\bar{x}$ such that $\Psi(x) - \vartheta(x) > 0$ for $x \in U$, which implies that $\Psi(x) = (\varphi + \delta(\cdot;\Omega))(x)$ for $x \in U$. Since $\bar{x}$ is a local optimal solution to (8.80), we have by the generalized Fermat rule that

$$0 \in \partial\Psi(\bar{x}) = \partial\big(\varphi + \delta(\cdot;\Omega)\big)(\bar{x}).$$

It follows from the assumptions imposed on φ and Ω and the sum rules for the basic and singular subdifferentials from Exercise 2.54(i) that

$$\begin{array}{l}\partial\big(\varphi + \delta(\cdot;\Omega)\big)(\bar{x}) \subset \partial\varphi(\bar{x}) + N(\bar{x};\Omega)\\ \partial^\infty\big(\varphi + \delta(\cdot;\Omega)\big)(\bar{x}) \subset \partial^\infty\varphi(\bar{x}) + N(\bar{x};\Omega).\end{array} \tag{8.81}$$

Thus we have $0 \in \partial\varphi(\bar{x}) + N(\bar{x};\Omega)$, which ensures the validity of the necessary optimality conditions in (8.78) with $y^* = 0$ in this case.

Next we consider the case of $\vartheta(\bar{x}) = 0$. Since the function ϑ is locally Lipschitzian around $\bar{x}$, it follows from Theorem 4.10, which holds in Asplund spaces with no change in the Lipschitz case under consideration, that

$$\begin{array}{l}\partial^\infty\Psi(\bar{x}) \subset \partial^\infty\big(\vartheta + \delta(\cdot;\Omega)\big)(\bar{x}) + \partial^\infty\varphi(\bar{x}) = \partial^\infty\big(\vartheta + \delta(\cdot;\Omega)\big)(\bar{x}),\\ \partial\Psi(\bar{x}) \subset \bigcup\Big\{\lambda_1 \circ \partial\big(\vartheta + \delta(\cdot;\Omega)\big)(\bar{x}) + \lambda_2\partial\varphi(\bar{x})\Big|\ (\lambda_1,\lambda_2) \in \mathbb{R}^2_+,\ \lambda_1 + \lambda_2 = 1\Big\},\end{array}$$

where $\lambda \circ \partial\vartheta(\bar{x})$ denotes $\lambda\partial\vartheta(\bar{x})$ when $\lambda > 0$ and $\partial^\infty\vartheta(\bar{x})$ when $\lambda = 0$. Since $0 \in \partial\Psi(\bar{x})$, we get from (8.81) and the latter inclusions that there exist $x^* \in N(\bar{x};\Omega)$ and $(\lambda_1,\lambda_2) \in \mathbb{R}^2_+$ such that $\lambda_1 + \lambda_2 = 1$ and that

$$0 \in \lambda_1 \circ \partial\vartheta(\bar{x}) + \lambda_2\partial\varphi(\bar{x}) + x^*. \tag{8.82}$$

If $\lambda_1 \neq 0$ in (8.82), then there is $u^* \in \partial\vartheta(\bar{x})$ with $-x^* - \lambda_1 u^* \in \lambda_2\partial\varphi(\bar{x})$. If $\lambda_2 = 0$ and thus $\lambda_1 = 1$ in (8.82), we obtain (8.78) with $y^* = 0$ by

$$0 = u^* + x^* \in \partial\varphi(\bar{x}) + \widetilde{D}^*_{N,\Theta} f(\bar{x})(0) + N(\bar{x};\Omega).$$

Otherwise Theorem 8.44 with $\Lambda = \Xi$ allows us to find $y^* \in \text{cl}^*\text{co}\,\Xi$ satisfying

$$\frac{-x^* - \lambda_1 u^*}{\lambda_2} \in \widetilde{D}^*_{N,\Theta} f(\bar{x})(y^*) \ \text{ and } \ \langle y^*, f(\bar{x})\rangle = \varphi(\bar{x}) = 0.$$

Hence we arrive at the inclusions

$$0 \in u^* + \frac{\lambda_2}{\lambda_1}\widetilde{D}^*_{N,\Theta} f(\bar{x})(y^*) + \frac{x^*}{\lambda_1} \subset \partial\varphi(\bar{x}) + \widetilde{D}^*_{N,\Theta} f(\bar{x})\Big(\frac{\lambda_2 y^*}{\lambda_1}\Big) + N(\bar{x};\Omega),$$

which justify the conditions of (8.78) in this case.

Supposing then that $\lambda_1 = 0$, we deduce from (8.82) the existence of $v^* \in \partial^\infty\varphi(\bar{x})$ such that $-v^* - x^* \in \partial\vartheta(\bar{x})$. Applying Theorem 8.44 again gives us $z^* \in \text{cl}^*\text{co}\,\Xi$ satisfying the conditions $-v^* - x^* \in \widetilde{D}^*_{N,\Theta} f(\bar{x})(z^*)$ and $\langle z^*, f(\bar{x})\rangle = 0$, which readily yield (8.79). The rest of the theorem, which deals with the particular structures of the spaces X and Y, follows by the above arguments from the corresponding results of Theorem 8.44. △

Note that the (qualification) condition (b) of Theorem 8.45 holds trivially if $0 \in \text{cl}^*\text{co}\,\Xi$. Indeed, in this case we always have $0 \in \widetilde{D}^*_{N,\Theta} f(\bar{x})(0) \cap \partial^\infty\varphi(\bar{x}) \cap N(\bar{x};\Omega)$.

The next proposition shows that the origin is never an element of $\text{cl}^*\text{co}\,\Xi$ if, in particular, the interior of the cone Θ is nonempty.

Proposition 8.46 (Solid Cone Constraints). *The following are equivalent:*

(i) $0 \notin \text{cl}^*\text{co}\,\Xi$.

(ii) *There are* $r > 0$ *and* $y_0 \in Y$ *such that* $\langle y^*, y_0\rangle > r$ *for all* $y^* \in \Xi$.

(iii) $\text{int}\,\Theta \neq \emptyset$.

Proof. Implication (i)⇒(ii) follows directly from the convex separation theorem. To prove (ii)⇒(iii), we get from (ii) for any $y \in B_r(y_0)$ that

$$\langle y^*, y\rangle = \langle y^*, y_0\rangle + \langle y^*, y - y_0\rangle \geq r - \|y^*\| \cdot \|y - y_0\| > r - r = 0$$

whenever $y^* \in \Xi$. This implies that $y \in \Theta$ and so ensures (iii). Finally, suppose that (iii) is satisfied and then find $y_1 \in \Theta$ and $s > 0$ such that $B_s(y_1) \subset \Theta$. For any $y^* \in \Xi$, we clearly have

$$\langle y^*, y_1\rangle = \langle y^*, y_1\rangle - s\|y^*\| + s \geq \inf_{y \in B_s(y_1)} \langle y^*, y\rangle + s \geq s > 0,$$

which yields $\langle y^*, y_1\rangle > s$ if $y^* \in \text{co}\,\Xi$, and thus (i) holds. △

Next we present several remarkable consequences of Theorem 8.45. The first one shows that in the case of solid cone constraints, the necessary optimality conditions in (8.78) hold under a certain *enhanced* constraint qualification.

Corollary 8.47 (Optimality Conditions Under Enhanced Qualifications for Solid Cone Constraints). *Suppose in the setting of Theorem 8.44 that* $\text{int}\,\Theta \neq \emptyset$ *and that the qualification condition*

$$\big(\partial^\infty\varphi(\bar{x}) + N(\bar{x}; \Omega)\big) \cap \big(-\widetilde{D}^*_{N,\Theta} f(\bar{x})(\Xi_0)\big) = \emptyset \tag{8.83}$$

holds with $\Xi_0 := \{y^* \in \Xi |\ \langle y^*, f(\bar{x})\rangle = 0\}$. *Then there is* $y^* \in \Theta^+$ *so that the conditions in (8.78) are satisfied. If* $\dim X < \infty$, *then* $\widetilde{D}^*_{N,\Theta} f(\bar{x})$ *can be replaced by* $\breve{D}^*_{N,\Theta} f(\bar{x})$ *in (8.78). Furthermore,* $\widetilde{D}^*_{N,\Theta} f(\bar{x})$ *can be replaced by* $D^*_{N,\Theta} f(\bar{x})$ *if in addition the unit ball* $\mathbb{B}^* \subset Y^*$ *is weak* sequentially compact.*

Proof. Following the proof of Theorem 8.44, it is sufficient to show that $\lambda_1 \neq 0$ under the assumptions made. Arguing by contradiction, suppose that $\lambda_1 = 0$ and then find $x^* \in N(\bar{x}; \Omega)$, $v^* \in \partial^\infty\varphi(\bar{x})$, and $z^* \in \text{cl}^*\text{co}\,\Xi$ such that

$$-v^* - x^* \in \widetilde{D}^*_{N,\Theta} f(\bar{x})(z^*) \;\text{ and }\; \langle z^*, f(\bar{x})\rangle = 0.$$

It follows from Proposition 8.46 that $z^* \neq 0$. Hence we have

$$\partial^\infty\varphi(\bar{x}) + N(\bar{x}; \Omega) \ni \frac{v^*}{\|z^*\|} + \frac{x^*}{\|z^*\|} = -\frac{-v^* - x^*}{\|z^*\|} \in -\widetilde{D}^*_{N,\Theta} f(\bar{x})\Big(\frac{z^*}{\|z^*\|}\Big),$$

which contradicts the qualification condition (8.83) due to $\frac{z^*}{\|z^*\|} \in \Xi_0$. △.

Our last consequence in this subsection concerns the settings of (8.64) where the cost function φ is locally Lipschitzian at $\bar{x}$ and where the constraint mapping f is either strictly differentiable at $\bar{x}$ or Θ-convex. We can see that in such settings the qualification condition (8.83) is equivalent to the classical Robinson and Slater constraint qualifications, respectively.

Corollary 8.48 (Cone-Constrained Problems in Special Settings). *Assume in the framework of Corollary 8.47 that φ is locally Lipschitzian around $\bar{x}$ and the constraint set $\Omega \subset X$ is convex. The following assertions hold:*

(i) *If f is strictly differentiable at $\bar{x}$, then the qualification condition (8.83) is equivalent to the Robinson constraint qualification:*

$$0 \in \operatorname{int}\big\{f(\bar{x}) + \nabla f(\bar{x})(\Omega - \bar{x}) + \Theta\big\} \tag{8.84}$$

and the optimality condition (5.20) reduces to the existence of $y^* \in \Theta^+$ with $\langle y^*, f(\bar{x})\rangle = 0$ and $x^* \in \partial\varphi(\bar{x})$ satisfying

$$\langle x^* + \nabla f(\bar{x})^* y^*, x - \bar{x}\rangle \geq 0 \ \text{ for all } \ x \in \Omega. \tag{8.85}$$

(ii) *If f is Θ-convex, then the qualification condition (8.83) is equivalent to the Slater constraint qualification:*

$$\text{there is } \ x_0 \in \Omega \ \text{ with } \ f(x_0) \in -\operatorname{int}\Theta \tag{8.86}$$

while the optimality condition (8.78) reduces to the existence of $y^* \in \Theta^+$ with $\langle y^*, f(\bar{x})\rangle = 0$, $u^* \in \partial\langle y^*, f\rangle(\bar{x})$, and $x^* \in \partial\varphi(\bar{x})$ satisfying

$$\langle x^* + u^*, x - \bar{x}\rangle \geq 0 \ \text{ for all } \ x \in \Omega. \tag{8.87}$$

Proof. Since $\partial^\infty\varphi(\bar{x}) = \{0\}$ for locally Lipschitzian functions and due to the convexity of Ω, the qualification condition (8.83) has the form

$$\nexists x^* \in -\widetilde{D}^*_{N,\Theta} f(\bar{x})(\Xi_0) \ \text{ with } \ \langle x^*, x - \bar{x}\rangle \leq 0 \ \text{ for all } \ x \in \Omega.$$

To justify (i), assume that f is strictly differentiable at $\bar{x}$ and observe by applying the supporting hyperplane theorem that (8.84) is equivalent to

$$N\big(0; f(\bar{x}) + \nabla f(\bar{x})(\Omega - \bar{x}) + \Theta\big) = \{0\}. \tag{8.88}$$

Suppose that (8.83) holds and show that (8.88) is satisfied. Indeed, if on the contrary there is $y^* \in N\big(0; f(\bar{x}) + \nabla f(\bar{x})(\Omega - \bar{x}) + \Theta\big)$ with $\|y^*\| = 1$, then

$$\langle y^*, f(\bar{x}) + \nabla f(\bar{x})(x - \bar{x}) + z\rangle \leq 0 \ \text{ for all } \ x \in \Omega \ \text{ and } \ z \in \Theta,$$

which yields $y^* \in -\Theta^+$ with $\langle y^*, f(\bar{x})\rangle \leq 0$. Moreover, $\langle y^*, f(\bar{x})\rangle \geq 0$ by $y^* \in -\Theta^+$ and $f(\bar{x}) \in -\Theta$. It follows also that $y^* \in -\Xi_0$ and $\nabla f(\bar{x})^* y^* \in N(\bar{x}; \Omega)$. By scalarization (8.73) we arrive at a contradiction with (8.83).

Conversely, suppose that the Robinson constraint qualification (8.84) is satisfied. If there exists $z^* \in \Xi_0$ with $N(\bar{x}; \Omega) \cap \big(-\widetilde{D}^*_{N,\Theta} f(\bar{x})(z^*)\big) \neq \emptyset$, we easily get from (8.73) that $-z^* \in N\big(0; f(\bar{x}) + \nabla f(\bar{x})(\Omega - \bar{x}) + \Theta\big)$ while yielding $z^* = 0$. This is a contradiction that justifies the equivalence between (8.83) and (8.84) in assertion (i). The equivalence between the necessary optimality conditions (8.78) and (8.85) in this case follows from the structure of the normal cone to convex sets and the coderivative scalarization (8.73), which therefore complete the proof of assertion (i).

Next we verify assertion (ii), where the constraint mapping f is Θ-convex in (8.64). Suppose first that the Slater condition (8.86) doesn't hold, i.e., $f(\Omega) \cap (-\text{int } \Theta) = \emptyset$. Then it is easy to check that $A \cap (-\text{int } \Theta) = \emptyset$, where $A := \{f(x) + \Theta |\ x \in \Omega\}$ is a convex set in Y. Applying the separation theorem to these two sets gives us $w^* \in Y^*$ with $\|w^*\| = 1$ such that

$$\langle w^*, f(x)\rangle \geq \langle w^*, -z\rangle \ \text{ for all } \ x \in \Omega,\ z \in \Theta.$$

It follows that $w^* \in \Theta^+$ and $\langle w^*, f(x)\rangle \geq 0$ as $x \in \Omega$. Since $f(\bar{x}) \in -\Theta$, we get that $\langle w^*, f(\bar{x})\rangle = 0$ and $\langle w^*, f(x)\rangle - \langle w^*, f(\bar{x})\rangle \geq 0$ for $x \in \Omega$. This yields

$$0 \in \partial(\langle w^*, f\rangle + \delta(\cdot; \Omega))(\bar{x}) \subset \partial\langle w^*, f\rangle(\bar{x}) + N(\bar{x}; \Omega).$$

Thus we arrive at $N(\bar{x}; \Omega) \cap \big(-\widetilde{D}^*_{N,\Theta} f(\bar{x})(w^*)\big) \neq \emptyset$ due to the scalarization formula in (8.74), which shows that condition (8.83) is violated.

Conversely, assume that the Slater condition (8.86) holds and then find $x_0 \in \Omega$ with $f(x_0) \in -\text{int } \Theta$. Supposing that there exists

$$u^* \in \Xi_0 \ \text{ with } \ N(\bar{x}; \Omega) \cap \big(-\widetilde{D}^*_{N,\Theta} f(\bar{x})(u^*)\big) \neq \emptyset,$$

we get from the coderivative scalarization (8.74) that $0 \in \partial\langle u^*, f\rangle(\bar{x}) + N(\bar{x}; \Omega)$. This implies that $0 \leq \langle u^*, f(x_0)\rangle - \langle u^*, f(\bar{x})\rangle = \langle u^*, f(x_0)\rangle$. Since $-f(x_0) \in \text{int}\Theta$, it follows from the proof of the implication (iii)$\Rightarrow$(i) in Proposition 8.46 that $\langle u^*, -f(x_0)\rangle > 0$, which is a contradiction. Thus we justify the equivalence between the qualification conditions (8.83) and (8.86) in the convex setting under consideration. Finally, the necessary optimality conditions in (8.78) reduce to those in (8.87) in this setting due to the convexity of Ω and the scalarization formula (8.74). $\triangle$

8.3.3 Qualified Optimality Conditions Without CQs

In this subsection we present necessary optimality conditions of a new type for cone-constrained programs (8.64). These results are essentially different from those obtained in Subsection 8.3.2 in the following major aspects:

(i) The results below are obtained in a *qualified* form (i.e., with nonzero multipliers corresponding to cost functions), while they are established *without any constrained qualification* (CQs).

(ii) The obtained results are given in an *approximate/fuzzy form*, i.e., they involve neighborhoods of the reference local optimal solution.

Note that some necessary optimality conditions of the fuzzy type have been derived in the literature for nonlinear programs *under* certain qualification conditions; see more discussions and references in Section 8.6.

We start with the following simple lemma.

Lemma 8.49 (Fuzzy Estimates of Basic Normals to Inverse Images). *Let* $\bar{x} \in f^{-1}(-\Theta)$ *for* $f\colon X \to Y$ *under the standing assumptions, and let* V^* *be a weak* neighborhood of* $0 \in X^*$*. Then for any basic normal* $x^* \in N(\bar{x}; f^{-1}(-\Theta))$ *and any* $\varepsilon > 0$ *there exist* $x_\varepsilon \in B_\varepsilon(\bar{x})$ *and* $y^*_\varepsilon \in \Theta^+$ *such that*

$$x^* \in \widehat{\partial}\langle y^*_\varepsilon, f\rangle(x_\varepsilon) + V^* \ \ \textit{with} \ \ |\langle y^*_\varepsilon, f(x_\varepsilon)\rangle| \le \varepsilon. \tag{8.89}$$

Proof. It follows from the convex separation theorem that

$$\delta\big(x; f^{-1}(-\Theta)\big) = \sup_{y^*\in\Theta^+} \langle y^*, f(x)\rangle \ \text{ for all } \ x \in X.$$

Applying Theorem 8.42 to the case of $\Lambda := \Theta^+$ ensures the existence of $y^*_\varepsilon \in \operatorname{co}\Lambda = \Theta^+$ and $x_\varepsilon \in B_\varepsilon(\bar{x})$ satisfying the relationships

$$|\langle y^*_\varepsilon, f(x_\varepsilon)\rangle| = |\langle y^*_\varepsilon, f(x_\varepsilon)\rangle - \delta\big(\bar{x}; f^{-1}(-\Theta)\big)| \le \varepsilon \ \text{ and } \ x^* \in \widehat{\partial}\langle y^*_\varepsilon, f\rangle(x_\varepsilon) + V^*,$$

which verify (8.89) and thus complete the proof of the lemma. △

The main result of this subsection is as follows.

Theorem 8.50 (Fuzzy Optimality Conditions for Cone-Constrained Programs). *Let* $\bar{x}$ *be a local optimal solution to problem* (8.64)*. Then for any weak* neighborhood* V^* *of* $0 \in X^*$ *and any* $\varepsilon > 0$*, there exist* $x_0, x_1, x_\varepsilon \in B_\varepsilon(\bar{x})$ *and* $y^*_\varepsilon \in \Theta^+$ *such that* $|\vartheta(x_0) - \vartheta(\bar{x})| \le \varepsilon$*,* $x_1 \in \Omega$*, and*

$$0 \in \widehat{\partial}\vartheta(x_0) + \widehat{\partial}\langle y^*_\varepsilon, f\rangle(x_\varepsilon) + \widehat{N}(x_1; \Omega) + V^* \ \textit{with} \ |\langle y^*_\varepsilon, f(x_\varepsilon)\rangle| \le \varepsilon. \tag{8.90}$$

Proof. Suppose without loss of generality that V^* is convex in X^*. Since $\bar{x}$ is a local solution to (8.64), we have by the generalized Fermat rule that

$$0 \in \widehat{\partial}\big(\vartheta + \delta(\cdot; \Omega) + \delta(\cdot; f^{-1}(-\Theta))\big)(\bar{x}).$$

Using the weak fuzzy sum rule from Exercise 2.27 gives us $x_0 \in B_\varepsilon(\bar{x})$ with $|\varphi(x_0) - \varphi(\bar{x})| \le \varepsilon$, $x_1 \in \Omega \cap B_\varepsilon(\bar{x})$, and $x_2 \in f^{-1}(-\Theta) \cap B_{\frac{\varepsilon}{2}}(\bar{x})$ such that

$$0 \in \widehat{\partial}\varphi(x_0) + \widehat{N}(x_1; \Omega) + \widehat{N}\big(x_2; f^{-1}(-\Theta)\big) + \frac{V^*}{2}.$$

Thus there is $x^* \in \widehat{N}(x_2; f^{-1}(-\Theta)) \subset N(x_2; f^{-1}(-\Theta))$ satisfying

$$0 \in x^* + \widehat{\partial}\varphi(x_0) + \widehat{N}(x_1; \Omega) + \frac{V^*}{2}.$$

By Proposition 8.49 we find $x_\varepsilon \in B_{\frac{\varepsilon}{2}}(x_2)$ and $y_\varepsilon^* \in \Theta^+$ such that

$$x^* \in \widehat{\partial}\langle y_\varepsilon^*, f\rangle(x_\varepsilon) + \frac{V^*}{2} \;\text{ with }\; |\langle y_\varepsilon^*, f(x_\varepsilon)\rangle| \le \varepsilon.$$

This immediately yields the inclusions

$$\begin{aligned} 0 &\in \widehat{\partial}\varphi(x_0) + \widehat{N}(x_1; \Omega) + \frac{V^*}{2} + \widehat{\partial}\langle y_\varepsilon^*, f\rangle(x_\varepsilon) + \frac{V^*}{2} \\ &\subset \widehat{\partial}\varphi(x_0) + \widehat{\partial}\langle y_\varepsilon^*, f\rangle(x_\varepsilon) + \widehat{N}(x_1; \Omega) + V^*, \end{aligned}$$

which imply in turn the optimality conditions in (8.90) by taking into account the obvious estimates $\|x_\varepsilon - \bar{x}\| \le \|x_\varepsilon - x_2\| + \|x_2 - \bar{x}\| \le \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$. △

As a consequence of the fuzzy optimality conditions of Theorem 8.50, we derive the following *sequential* qualified optimality conditions for a particular setting of cone-constrained programs (8.64) *without constraint qualifications*.

Corollary 8.51 (Sequential Optimality Conditions for Cone-Constrained Programs). *Suppose in the framework of Theorem 8.50 that* $\dim X < \infty$, $\Omega = X$, *and the cost function* φ *is Lipschitz continuous around* $\bar{x}$. *Then there exist a basic subgradient* $x^* \in \partial\varphi(\bar{x})$ *and sequences* $\{x_k\} \subset X$, $\{x_k^*\} \subset X^*$, *and* $\{y_k^*\} \subset \Theta^+$ *with* $x_k^* \in \widehat{\partial}\langle y_k^*, f\rangle(x_k)$ *for all* $k \in \mathbb{N}$ *such that*

$$x_k \to \bar{x}, \;\; x_k^* \to -x^*, \;\; \textit{and} \;\; \langle y_k^*, f(x_k)\rangle \to 0 \;\; \textit{as} \;\; k \to \infty. \tag{8.91}$$

Proof. Since $\dim X < \infty$, we can select $V^* = \frac{1}{k}\mathbb{B}^*$, $\varepsilon = \frac{1}{k}$ and then find from (8.90) vectors $u_k, x_k \to \bar{x}$ together with dual elements $u_k^* \in \widehat{\partial}\varphi(u_k)$, $y_k^* \in \Theta^+$, and $x_k^* \in \widehat{\partial}\langle y_k^*, f\rangle(x_k)$ satisfying

$$-u_k^* \in x_k^* + \frac{1}{k}\mathbb{B}^* \;\text{ and }\; |\langle y_k^*, f(x_k)\rangle| \le \frac{1}{k} \;\text{ as }\; k \to \infty. \tag{8.92}$$

It follows from the local Lipschitz continuity of φ around $\bar{x}$ that the sequence $\{u_k^*\}$ is bounded, and hence it converges (without relabeling) to some basic subgradient $x^* \in \partial\varphi(\bar{x})$. This implies due to the inclusion in (8.92) that $x_k^* \to -x^*$, which justifies (8.91) and thus completes the proof. △

Yet another remarkable consequence of Theorem 8.50 is its following enhanced version for the case of nondifferentiable programming.

Corollary 8.52 (Fuzzy Optimality Conditions in Nondifferentiable Programming). *Let $\bar{x}$ be a local optimal solution to the program:*

$$\begin{cases} \text{minimize } \varphi(x) \text{ subject to } x \in \Omega \subset X, \\ \varphi_i(x) \le 0, \ i = 1, \ldots, m, \ \text{ and } \ \varphi_i(x) = 0, \ \ i = m+1, \ldots, m+r, \end{cases}$$

where in addition to the standing assumptions, we suppose that all the functions $\varphi_i\colon X \to \mathbb{R}$ are Lipschitz continuous around $\bar{x}$. Then for any weak* neighborhood V^* of $0 \in X^*$ and any $\varepsilon > 0$, there are vectors $x_0, x_1, \ldots, x_{m+r}, \widehat{x} \in B_\varepsilon(\bar{x})$ and multipliers $(\lambda_1, \ldots, \lambda_{m+r}) \in \mathbb{R}^m_+ \times \mathbb{R}^r$ such that

$$0 \in \widehat{\partial}\varphi(x_0) + \sum_{i=1}^{m} \lambda_i \widehat{\partial}\varphi_i(x_i) + \sum_{i=m+1}^{m+r} \widehat{\partial}(\lambda_i\varphi_i)(x_i) + \widehat{N}(\widehat{x}; \Omega) + V^* \tag{8.93}$$

with $\widehat{x} \in \Omega$, $\big|\sum_{i=1}^{m+r} \lambda_i\varphi_i(x_i)\big| \le \varepsilon$, and $|\varphi(x_0) - \varphi(\bar{x})| \le \varepsilon$.

Proof. Employing Theorem 8.50 in the case of $Y := \mathbb{R}^{m+r}$, $f := (\varphi_1, \ldots, \varphi_{m+r})$, and $\Theta := \mathbb{R}^m_+ \times 0_r \subset Y$ gives us $x_0, x_\varepsilon, \widehat{x} \in B_{\frac{\varepsilon}{2}}(\bar{x})$, and $(\lambda_1, \ldots, \lambda_{m+r}) \in \Theta^+ = \mathbb{R}^m_+ \times \mathbb{R}^r$ such that $|\vartheta(x_0) - \vartheta(\bar{x})| \le \varepsilon$, $\widehat{x} \in \Omega$, and

$$\begin{aligned} &0 \in \widehat{\partial}\varphi(x_0) + \widehat{\partial}\Big(\sum_{i=1}^{m+r} \lambda_i\varphi_i\Big)(x_\varepsilon) + \widehat{N}(\widehat{x}; \Omega) + \frac{V^*}{2} \\ &\text{with } \Big|\sum_{i=1}^{m+r} \lambda_i\varphi_i(x_\varepsilon)\Big| \le \frac{\varepsilon}{2}. \end{aligned} \tag{8.94}$$

Thus there is $x^* \in \widehat{\partial}\big(\sum_{i=1}^{m+r} \lambda_i\varphi_i\big)(x_\varepsilon)$ satisfying

$$0 \in x^* + \widehat{\partial}\vartheta(x_0) + \widehat{N}(\widehat{x}; \Omega) + \frac{V^*}{2}.$$

Then we apply to x^* the weak fuzzy sum rule from Exercise 2.27 and find $x_1^*, \ldots, x_{m+r}^*$ together with $x_1, \ldots, x_{m+r} \in B_{\frac{\varepsilon}{2}}(x_\varepsilon)$ such that

$$x_i^* \in \widehat{\partial}(\lambda_i\varphi_i)(x_i), \ |\lambda_i\varphi_i(x_i) - \lambda_i\varphi_i(x_\varepsilon)| \le \frac{\varepsilon}{2(m+r)} \ \text{ for } \ i = 1, \ldots, m+r,$$
$$x^* \in \sum_{i=1}^{m+r} x_i^* + \frac{V^*}{2}.$$

It follows from the above that $\|x_i - \bar{x}\| \le \|x_i - x_\varepsilon\| + \|x_\varepsilon - \bar{x}\| \le \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$ for all $i = 1, \ldots, m+r$ and that the inclusions

$$\begin{aligned} 0 &\in \widehat{\partial}\varphi(x_0) + \widehat{N}(\widehat{x};\Omega) + \frac{V^*}{2} + \sum_{i=1}^{m+r}\widehat{\partial}(\lambda_i\varphi_i)(x_i) + \frac{V^*}{2} \\ &\in \widehat{\partial}\varphi(x_0) + \sum_{i=1}^{m}\lambda_i\widehat{\partial}\varphi_i(x_i) + \sum_{i=m+1}^{m+r}\widehat{\partial}(\lambda_i\varphi_i)(x_i) + \widehat{N}(\widehat{x};\Omega) + V^* \end{aligned} \tag{8.95}$$

hold. Moreover, we get from (8.94) that

$$\Big|\sum_{i=1}^{m+r}\lambda_i\varphi_i(x_i)\Big| \le \sum_{i=1}^{m+r}\Big|\lambda_i\varphi_i(x_i) - \lambda_i\varphi_i(x_\varepsilon)\Big| + \Big|\sum_{i=1}^{m+r}\lambda_i\varphi_i(x_\varepsilon)\Big| \le \frac{\varepsilon(m+r)}{2(m+r)} + \frac{\varepsilon}{2}.$$

This together with (8.95) yields (8.93) and thus completes the proof. $\triangle$

8.3.4 Well-Posedness of Cone-Constrained Systems

In this subsection we return to investigating *well-posedness* properties of parametric systems, which have been studied in Chapter 3 in the general/abstract framework of finite-dimensional spaces (while most of the results are valid in Asplund spaces as discussed in Sections 3.4 and 3.5) and also in Chapter 7 for infinite linear and convex SIP systems in Banach spaces. Here we consider nonconvex and nonsmooth *cone-constrained systems* given by

$$F(x) := f(x) + \Theta = \big\{y \in Y \big|\ f(x) - y \in -\Theta\big\} \tag{8.96}$$

under the standing assumptions of this section. In contrast to Chapter 7, where the results revolve around the Lipschitz-like property of the particular systems in question, we focus in this section on the equivalent (up to considering the inverse mappings) property of *metric regularity* for (8.96) the study of which has some difference from Lipschitzian stability. Observe, in particular, that the image (Banach) space Y and domain (Asplund) space X in (8.96) have the opposite meaning for F^{-1}. We also refer the reader to Section 3.3 for some challenges in the study of metric regularity concerning parametric variational systems in both finite and infinite dimensions. Our approach to metric regularity of (8.96) is based on variational techniques and using the results developed in the previous subsections above.

In what follows we assume for simplicity that the domain space X is *finite-dimensional* while Y is arbitrary *Banach*. This is sufficient, in particular, for applications to SIPs considered in Subsection 8.3.5 with $Y = \mathcal{C}(T), l^\infty(T)$. The proofs below can be extended to the case of general Asplund spaces X.

First we derive an upper estimate with the case of equality for the *exact regularity bound* $\operatorname{reg} F(\bar{x}, 0)$ of F at $(\bar{x}, 0)$ via the regular coderivative of f at neighboring points. The obtained estimate and equality clearly imply a sufficient as well as a necessary and sufficient condition for metric regularity, respectively. Note that $\widehat{D}^*_\Theta f(x)(y^*)$ can be replaced by the regular subdifferential $\widehat{\partial}\langle y^*, f\rangle(x)$ with $y^* \in \Theta^+$ due to the scalarization formula (8.72).

Theorem 8.53 (Neighborhood Evaluation of the Regularity Bound for Cone-Constrained Systems). *Let $\bar{x}$ be such that $f(\bar{x}) \in -\Theta$ for system (8.96), and let Ξ be defined in (8.66). Then we have the upper estimate*

$$\operatorname{reg} F(\bar{x}, 0) \le \inf_{\eta>0} \sup \left\{ \frac{1}{\|x^*\|} \,\middle|\, x^* \in \widehat{D}^*_\Theta f(x)(y^*),\ x \in B_\eta(\bar{x}),\right. \left. y^* \in \Xi,\ |\langle y^*, f(\bar{x})\rangle| < \eta \right\}, \tag{8.97}$$

which holds as equality provided that $f(\bar{x}) = 0$.

Proof. Denote by $a(\bar{x})$ the right-hand side of (8.97) and consider the nontrivial case in (8.97) when $a(\bar{x}) < \infty$. Arguing by contradiction, suppose that $\operatorname{reg} F(\bar{x}, 0) > a(\bar{x})$ and thus $x^* \ne 0$ in (8.97). Hence there are sequences $(x_k, y_k) \to (\bar{x}, 0)$ and $\nu < \alpha_k < \nu + 1$ for some $\nu > a(\bar{x})$ such that we have

$$\operatorname{dist}\big(x_k; F^{-1}(y_k)\big) > \alpha_k \operatorname{dist}\big(y_k; F(x_k)\big) > 0, \quad k \in \mathbb{N}. \tag{8.98}$$

Define $\psi_k(x) := \operatorname{dist}(y_k; F(x))$ and then get $\varepsilon_k := \psi_k(x_k) > 0$. Since the set $F(x) = f(x) + \Theta$ is convex for all $x \in X$, we apply the classical Fenchel duality theorem to obtain the representations

$$\begin{aligned}
\psi_k(x) &= \inf_{y\in Y} \Big\{ \|y - y_k\| + \delta\big(y; F(x)\big) \Big\} \\
&= \max_{y^*\in Y^*} \Big\{ -\sup_{y\in Y} \big(\langle y^*, y\rangle - \|y - y_k\|\big) - \sup_{v\in Y} \big(\langle -y^*, v\rangle - \delta\big(v; f(x) + \Theta\big)\big) \Big\} \\
&= \max_{y^*\in Y^*} \Big\{ -\sup_{y\in Y} \big(\langle y^*, y + y_k\rangle - \|y\|\big) - \sup_{v\in\Theta} \langle -y^*, f(x) + v\rangle \Big\} \\
&= \max_{y^*\in Y^*} \Big\{ -\langle y^*, y_k\rangle - \delta(y^*; \mathbb{B}^*) + \langle y^*, f(x)\rangle - \delta(y^*; \Theta^+) \Big\} \\
&= \max_{y^*\in\widetilde{\Xi}} \langle y^*, f(x) - y_k\rangle \ \text{ for each } \ k \in \mathbb{N},
\end{aligned} \tag{8.99}$$

where $\widetilde{\Xi} := \Theta^+ \cap \mathbb{B}^* \subset Y^*$. Thus the distance function $\psi_k(x)$ defined above can be represented as the *supremum* of Lipschitzian functions as in Theorem 8.42. This function is Lipschitz continuous on $B_\rho(\bar{x})$ with modulus K, where K and ρ are taken from (8.65). Suppose without loss of generality that $x_k \in B_\rho(\bar{x})$ for all $k \in \mathbb{N}$ and therefore arrive at the estimates

$$\begin{aligned}
\varepsilon_k = \psi_k(x_k) \le \psi_k(\bar{x}) + K\|x_k - \bar{x}\| &= \max_{y^*\in\widetilde{\Xi}} \langle y^*, f(\bar{x}) - y_k\rangle + K\|x_k - \bar{x}\| \\
&\le \max_{y^*\in\widetilde{\Xi}} \langle y^*, -y_k\rangle + K\|x_k - \bar{x}\| \le \|y_k\| + K\|x_k - \bar{x}\|,
\end{aligned}$$

which ensures that $\varepsilon_k \downarrow 0$ as $k \to \infty$. Since $\psi_k(x)$ is nonnegative for all $x \in X$, we have by the definition of ε_k that

$$\psi_k(x) + \varepsilon_k \geq \psi_k(x_k) \ \text{ whenever } \ x \in B_\rho(\bar{x}), \quad k \in \mathbb{N}.$$

Applying now Ekeland's variational principle gives us $\widehat{x}_k \in B_\rho(\bar{x})$ satisfying

$$\|\widehat{x}_k - x_k\| \leq \alpha_k \varepsilon_k < (\nu + 1)\varepsilon_k, \ \ \psi_k(x) + \alpha_k^{-1}\|x - \widehat{x}_k\| \geq \psi(\widehat{x}_k) \tag{8.100}$$

on $B_\rho(\bar{x})$. It follows from (8.98) and (8.100) that $\|\widehat{x}_k - x_k\| < \text{dist}(x_k; F^{-1}(y_k))$, which yields $\widehat{x}_k \notin F^{-1}(y_k)$, i.e., $y_k \notin F(\widehat{x}_k)$. Thus $\psi_k(\widehat{x}_k) = \text{dist}(y_k; F(\widehat{x}_k)) > 0$. Moreover, we deduce from (8.100) that

$$0 \in \partial\big(\psi_k + \alpha_k^{-1}\| \cdot - \widehat{x}_k\|\big)(\widehat{x}_k) \subset \partial\psi_k(\widehat{x}_k) + \alpha_k^{-1}\mathbb{B}^*.$$

Hence there exists $x_k^* \in \alpha_k^{-1}\mathbb{B}^*$ with $x_k^* \in \partial\psi_k(\widehat{x}_k)$. By the representation of ψ_k in (8.99) with the usage of Theorem 8.42 in the setting under consideration ($V^* = \gamma_k\mathbb{B}^*$) for any $\delta_k \in (0, \psi_k(\widehat{x}_k))$ sufficiently small we find $\widetilde{x}_k \in B_{\gamma_k}(\widehat{x}_k)$ and $y_k^* \in \text{co}\,\widetilde{\Xi} = \widetilde{\Xi}$ such that

$$x_k^* \in \widehat{\partial}\langle y_k^*, f\rangle(\widetilde{x}_k) + \gamma_k\mathbb{B}^* \text{ and } |\langle y_k^*, f(\widetilde{x}_k) - y_k\rangle - \psi_k(\widehat{x}_k)| < \gamma_k. \tag{8.101}$$

Due to the obvious upper estimates

$$\|\bar{x} - \widetilde{x}_k\| \leq \|\bar{x} - x_k\| + \|x_k - \widehat{x}_k\| + \|\widehat{x}_k - \widetilde{x}_k\| \leq \|\bar{x} - x_k\| + (\nu + 1)\varepsilon_k + \gamma_k$$

it follows from (8.99) and (8.101) that $y_k^* \neq 0$ and that

$$\begin{aligned}
\psi_k(\widehat{x}_k) &\leq \langle y_k^*, f(\widetilde{x}_k) - y_k\rangle + \gamma_k = \langle y_k^*, f(\widetilde{x}_k) - f(\widehat{x}_k)\rangle + \langle y_k^*, f(\widehat{x}_k) - y_k\rangle + \gamma_k \\
&\leq \|y_k^*\| \cdot \|f(\widetilde{x}_k) - f(\widehat{x}_k)\| + \|y_k^*\|\Big\langle \frac{y_k^*}{\|y_k^*\|}, f(\widehat{x}_k) - y_k\Big\rangle + \gamma_k \\
&\leq \|y_k^*\|K\|\widetilde{x}_k - \widehat{x}_k\| + \|y_k^*\|\psi_k(\widehat{x}_k) + \gamma_k \leq K\gamma_k + \|y_k^*\|\psi_k(\widehat{x}_k) + \gamma_k,
\end{aligned}$$

which implies in turn the inequalities

$$1 \geq \|y_k^*\| \geq 1 - \frac{(K+1)\gamma_k}{\psi_k(\widehat{x}_k)}. \tag{8.102}$$

Deduce further from (8.101) the estimates

$$\begin{aligned}
|\langle y_k^*, f(\bar{x})\rangle| &\leq \|y_k^*\|K\|\bar{x} - \widetilde{x}_k\| + \gamma_k + \psi_k(x_k) + K\|\widehat{x}_k - x_k\| + \|y_k^*\| \cdot \|y_k\| \\
&\leq K\|\bar{x} - \widetilde{x}_k\| + \gamma_k + \varepsilon_k + K(k+1)\varepsilon_k + \|y_k\|.
\end{aligned}$$

This ensures together with (8.102) that

$$|\langle \widehat{y}_k^*, f(\bar{x})\rangle| \leq \Big(K\|\bar{x} - \widetilde{x}_k\| + \gamma_k + \varepsilon_k + K(\nu+1)\varepsilon_k + \|y_k\|\Big)\Big(1 - \frac{(K+1)\gamma_k}{\psi_k(\widehat{x}_k)}\Big)^{-1},$$

where $\widehat{y}_k^* := \|y_k^*\|^{-1} y_k^* \in \Xi$. Furthermore, it follows from (8.101) and the coderivative scalarization formula (8.72) that there is $u_k^* \in \widehat{\partial}\langle y_k^*, f\rangle(\widetilde{x}_k) = \widehat{D}_\Theta^* f(\widetilde{x}_k)(y_k^*)$ satisfying $\|x_k^* - u_k^*\| \le \gamma_k$. Combining this with (8.102) and $x_k^* \in \alpha_k^{-1}\mathbb{B}^*$ yields the conditions

$$\widehat{u}_k^* := \|y_k^*\|^{-1} u_k^* \in \widehat{D}_\Theta^* f(\widetilde{x}_k)(\widehat{y}_k^*),$$

$$\|\widehat{u}_k^*\| \le \frac{\|x_k^*\| + \gamma_k}{\|y_k^*\|} \le (\alpha_k^{-1} + \gamma_k)\Big(1 - \frac{(K+1)\gamma_k}{\psi_k(\widehat{x}_k)}\Big)^{-1}.$$

Since $\alpha_k > \nu > a(\bar{x})$, we may choose γ_k so small that the right-hand side of the last estimate is strictly smaller than $\nu^{-1} < a(\bar{x})^{-1}$ and that $\max\{\|\widetilde{x}_k - \bar{x}\|, |\langle \widehat{y}_k^*, f(\bar{x})\rangle|\} \to 0$ as $k \to \infty$ due to the above estimates of $\|\bar{x} - \widetilde{x}_k\|$ and $|\langle \widehat{y}_k^*, f(\bar{x})\rangle|$. Hence for small $\eta > 0$ we get $\widetilde{x}_k \in B_\eta(\bar{x})$ and $\langle \widehat{y}_k^*, f(\bar{x})\rangle < \eta$ with $\widehat{y}_k^* \in \Xi$ and $\|\widehat{u}_k^*\| < \nu^{-1} < a(\bar{x})^{-1}$ if k is sufficiently large. This contradicts the definition of $a(\bar{x})$ and thus justifies the regularity estimate (8.97).

To complete the proof of the theorem, it remains to show that the equality holds in (8.97) if $f(\bar{x}) = 0$. It follows from the definition of $\operatorname{reg} F(\bar{x}, 0)$ that for any $\varepsilon > 0$ there are neighborhoods U of $\bar{x}$ and V of $f(\bar{x}) = 0$ with

$$\operatorname{dist}\big(x; F^{-1}(y)\big) \le \big(\operatorname{reg} F(\bar{x}, 0) + \varepsilon\big)\|y - f(x)\| \tag{8.103}$$

for $x \in U$ and $y \in V$. Picking $y^* \in \Xi$ and $x^* \in \widehat{D}_\Theta^* f(x)(y^*)$ for some x with $x \in U$ and $f(x) \in V$ allows us to find by Definition 8.43(i) such $\gamma > 0$ that

$$\langle x^*, u - x\rangle - \langle y^*, f(u) - f(x)\rangle \le \varepsilon(\|u - x\| + \|f(u) - f(x)\|) \tag{8.104}$$

for $u \in B_\gamma(x)$. It follows from (8.103) that for any $y \in Y$ close to $f(x)$ there is $u \in F^{-1}(y)$ near x such that

$$\|x - u\| \le (\operatorname{reg} F(\bar{x}, 0) + 2\varepsilon)\|y - f(x)\| \;\text{ with }\; y - f(u) \in \Theta.$$

Combining this with (8.104) gives us the estimates

$$\begin{aligned}
\langle -y^*, y - f(x)\rangle \le \langle -y^*, f(u) - f(x)\rangle &\le \varepsilon(\|u - x\| \\
&\quad + \|f(u) - f(x)\|) - \langle x^*, u - x\rangle \\
&\le \big(\varepsilon(1 + K) + \|x^*\|\big)\|u - x\| \\
&\le \big(\varepsilon(1 + K) + \|x^*\|\big)\big(\operatorname{reg} F(\bar{x}, 0) + 2\varepsilon\big)\|y - f(x)\|
\end{aligned}$$

for y near $f(x)$. Thus we get $\eta > 0$ with $B_\eta(f(x)) \subset V$ and hence

$$\begin{aligned}
1 = \|y^*\| &= \sup_{y \in B_\eta(f(x)) \setminus f(x)} \frac{\langle -y^*, y - f(x)\rangle}{\|y - f(x)\|} \\
&\le \big(\varepsilon(1 + K) + \|x^*\|\big)\big(\operatorname{reg} F(\bar{x}, 0) + 2\varepsilon\big),
\end{aligned}$$

which implies in turn the inequality

$$\|x^*\|^{-1} \le \left[\left(\operatorname{reg} F(\bar{x},0)+2\varepsilon\right)^{-1}-\varepsilon(1+K)\right]^{-1}.$$

Letting finally $\varepsilon \downarrow 0$ gives us $a(\bar{x}) \le \operatorname{reg} F(\bar{x},0)$ and thus justifies the equality in (8.97) while completing the proof of the theorem. △

Next we consider the *pointbased* condition

$$\left(\ker \breve{D}^*_{N,\Theta} f(\bar{x})\right) \cap \Xi_0 = \emptyset \ \text{ with } \ \Xi_0 = \left\{y^* \in \Xi \,\middle|\, \langle y^*, f(\bar{x})\rangle = 0\right\} \quad (8.105)$$

under our standing assumptions while remembering from Corollary 8.48 that (8.105) is equivalent to the Robinson constraint qualification for smooth mappings f. Now we show that (8.105) is sufficient for metric regularity of the conic systems (8.96) around $(\bar{x},0)$, provides a verifiable upper estimate of the exact regularity bound $\operatorname{reg} F(\bar{x},0)$ calculated at $\bar{x}$, and justifies the equality therein when f is either Θ-convex or strictly differentiable at $\bar{x}$.

Theorem 8.54 (Pointbased Conditions for Metric Regularity of Conic Systems). *Let $f(\bar{x}) \in -\Theta$ and $\operatorname{int}\Theta \neq \emptyset$ be as in Theorem 8.53. Then the constrained qualification (8.105) is sufficient for the metric regularity of the conic system F from (8.96) around $(\bar{x},0)$ with the exact regularity bound of F at $(\bar{x},0)$ estimated from the above by*

$$\operatorname{reg} F(\bar{x},0) \le b(\bar{x}) := \sup\left\{\frac{1}{\|x^*\|} \,\middle|\, x^* \in \breve{D}^*_{N,\Theta} f(\bar{x})(y^*),\ y^* \in \operatorname{cl}^*\Xi,\ \langle y^*, f(\bar{x})\rangle = 0\right\}, \quad (8.106)$$

where $x^* \neq 0$ due to (8.105). If furthermore Ξ is weak* closed in Y^* and if f is either Θ-convex or strictly differentiable at $\bar{x}$, then we have the equality $\operatorname{reg} F(\bar{x},0) = b(\bar{x})$ in (8.106), where $b(\bar{x})$ is calculated by

$$b(\bar{x}) = \sup\left\{\|y^*\| \,\middle|\, (y^*,-x^*) \in N\left((0,\bar{x}); \operatorname{gph} F^{-1}\right),\ \|x^*\| = 1\right\}. \quad (8.107)$$

The latter reduces to the the explicit formulas

$$b(\bar{x}) = \left\{\|y^*\| \,\middle|\, \langle y^*, y\rangle \le \langle x^*, x-\bar{x}\rangle \ \textit{for all} \ y \in F(x),\ \|x^*\| = 1\right\}$$

in the case of Θ-convex mappings f and

$$b(\bar{x}) = \sup\left\{\frac{1}{\|\nabla f(\bar{x})^* y^*\|} \,\middle|\, y^* \in \Xi \ \textit{with} \ \langle y^*, f(\bar{x})\rangle = 0\right\}$$

when the mapping f is strictly differentiable at $\bar{x}$.

Proof. First we check that the qualification condition (8.105) guarantees that the number $a(\bar{x})$, which is the right-hand side of (8.97), is finite. Indeed, the contrary means the existence of a sequence $(x_k, x_k^*, y_k^*) \in X \times X^* \times Y^*$ with

$$x_k \to \bar{x},\ \|x_k^*\| \to 0,\ y_k^* \in \Xi,\ x_k^* \in \widehat{D}_\Theta^* f(\bar{x})(y_k^*),\ \text{ and } \langle y_k^*, f(x_k)\rangle \to 0$$

as $k \to \infty$. By $\|y_k^*\| = 1$ for all $k \in \mathbb{N}$ we find a subnet of $\{y_k^*\}$ weak* converging to some $y^* \in \text{cl}^*\Xi$. Then it follows from the convergence above and the cluster coderivative construction from Definition 8.43 that $0 \in \breve{D}_{N,\Theta}^* f(\bar{x})(y^*)$ with $\langle y^*, f(\bar{x})\rangle = 0$. Proposition 8.46 ensures that $y^* \neq 0$ and therefore

$$\frac{y^*}{\|y^*\|} \in \big(\ker \breve{D}_{N,\Theta}^* f(\bar{x})\big) \cap \Xi_0.$$

This contradicts (8.105) and thus justifies that the number $a(\bar{x})$ is finite. By Theorem 8.53 we have that F is metrically regular around $(\bar{x}, 0)$.

Since $a(\bar{x})$ is finite, it follows from the regularity bound estimate in (8.97) that there is a sequence $(x_k, x_k^*, y_k^*) \in X \times X^* \times Y^*$ such that

$$x_k \to \bar{x},\ \frac{1}{\|x_k^*\|} \to a(\bar{x}),\ y_k^* \in \Xi,\ x_k^* \in \widehat{D}_\Theta^* f(\bar{x})(y_k^*),\ \text{ and } \langle y_k^*, f(x_k)\rangle \to 0.$$

Again we find a subnet of $\{(x_k^*, y_k^*)\}$ weak* converging to some $(x^*, y^*) \in X^* \times \text{cl}^*\Xi$ and conclude that $x^* \in \breve{D}_{N,\Theta}^* f(\bar{x})(y^*)$ and $y^* \in \text{cl}^*\Xi$ with $\langle y^*, f(\bar{x})\rangle = 0$. This gives us $a(\bar{x}) = \|x^*\|^{-1}$ and thus deduce the claimed upper estimate (8.106) from that given in (8.97).

To justify the equality in (8.106) with the corresponding representations of $b(\bar{x})$, observe that the weak* closedness of Ξ ensures the formula

$$\Xi_0 = \big\{y^* \in \text{cl}^*\Xi \big|\ \langle y^*, f(\bar{x})\rangle = 0\big\},$$

where Ξ_0 is taken from (8.105). If f is Θ-convex, we easily get from (8.74) that $x^* \in \breve{D}_{N,\Theta}^* f(\bar{x})(y^*)$ with $y^* \in \Xi_0$ if and only if $(x^*, -y^*) \in N((\bar{x}, 0); \text{gph}\, F)$ with $y^* \in S_{Y^*}$. By (8.106) it gives us the conditions

$$\text{reg}\, F(\bar{x}, 0) \le b(\bar{x}) = \sup\left\{\frac{1}{\|x^*\|}\ \middle|\ (x^*, -y^*) \in N\big((\bar{x}, 0); \text{gph}\, F\big),\ \|y^*\| = 1\right\}.$$

On the other hand, we have from Theorem 3.2 held in any Banach space and estimate (3.61) under the assumptions made that

$$\text{reg}\, F(\bar{x}, 0) \ge \sup\big\{\|y^*\|\ \big|\ (y^*, -x^*) \in N\big((0, \bar{x}); \text{gph}\, F^{-1}\big),\ \|x^*\| = 1\big\},$$

which yields the equality in (8.106) with $b(\bar{x})$ calculated by (8.107). The specification of (8.107) in the case of Θ-convex mappings follows directly from the structure of the normal cone to convex sets.

It remains to justify the equality case for mappings f strictly differentiable at $\bar{x}$. In this case we have from (8.73) that

$$\begin{aligned}\breve{D}_{N,\Theta}^* f(\bar{x})(y^*) &= \big\{\nabla f(\bar{x})^* y^*\big\} = \big\{x^* \in X^* \big|\ (x^*, -y^*) \in \widehat{N}\big((\bar{x}, 0); \text{gph}\, F\big)\big\} \\ &= \big\{x^* \in X^* \big|\ (x^*, -y^*) \in N\big((\bar{x}, 0); \text{gph}\, F\big)\big\} \quad \text{for any } \ y^* \in \Xi_0.\end{aligned}$$

Combining it with (8.106) and the lower estimate of the regularity bound $\operatorname{reg} F(\bar{x}, 0)$ above gives us the relationships

$$\begin{aligned}\operatorname{reg} F(\bar{x},0) \le b(\bar{x}) \le \sup\Big\{\frac{1}{\|x^*\|}\Big|\ (x^*,-y^*) \in \widehat{N}\big((\bar{x},0);\operatorname{gph} F\big),\ \|y^*\|=1\Big\}\\ = \sup\Big\{\frac{1}{\|x^*\|}\Big|\ (x^*,-y^*) \in N\big((\bar{x},0);\operatorname{gph} F\big),\ \|y^*\|=1\Big\}\\ \le \sup\Big\{\|y^*\|\Big|\ (y^*,-x^*) \in \widehat{N}\big((0,\bar{x});\operatorname{gph} F^{-1}\big),\ \|x^*\|=1\Big\} \le \operatorname{reg} F(\bar{x},0),\end{aligned}$$

which imply the equality in (8.106) and formula (8.107) for representing $b(\bar{x})$ in this case. The explicit calculation of $b(\bar{x})$ for strictly differentiable mappings follows from (8.106) with $\breve{D}^*_{N,\Theta} f(\bar{x})(y^*) = \{\nabla f(\bar{x})^* y^*\}$. $\triangle$

Observe that the weak* closedness assumption imposed on $\Xi \subset Y^*$ for ensuring the equality in Theorem 8.54 seems to be restrictive in infinite dimensions, since Ξ is a part of the unit sphere S_{Y^*}, which is never weak* closed in infinite-dimensional Banach spaces by the classical Josefson-Nissenzweig theorem. However, we show in the next subsection that the weak* closed assumption on Ξ is satisfied for the space $Y = l^\infty(T)$ with $\Theta = l^\infty_+(T)$ where T is an arbitrary index set, as well as for the space $Y = \mathcal{C}(T)$ with $\Theta = \mathcal{C}_+(T)$ where T is compact. Both spaces naturally appear in applications to the corresponding models of semi-infinite programming considered below.

8.3.5 Optimality and Well-Posedness for Nonconvex SIPs

Here we apply the obtained results for conic programs and cone-constrained systems to derive optimality conditions for nonsmooth SIPs and certifications of metric regularity for infinite Lipschitzian inequality systems. Let us first consider the following SIP with infinite inequality and geometric constraints:

$$\begin{cases}\text{minimize } \ \varphi(x) \ \text{ subject to}\\ f(x,t) \le 0,\ t \in T, \ \text{ and } \ x \in \Omega \subset X,\end{cases} \tag{8.108}$$

where $\dim X < \infty$ (for simplicity), $\varphi\colon X \to \overline{\mathbb{R}}$, $f\colon X \times T \to \overline{\mathbb{R}}$, and T is an arbitrary index set. In the absence of geometric constraints, SIP (8.108) with Lipschitzian data has been studied in Section 8.2 from the viewpoint of deriving optimality conditions by using another approach. Besides covering geometric constraints, our approach here applies to SIPs under significantly weaker assumptions and constraint qualifications in comparison with Section 8.2 and leads us to generally different optimality conditions. Furthermore, we establish also pointbased sufficient conditions and complete characterizations of metric regularity of infinite nonconvex inequality systems.

It is more convenient for us to use in (8.108) the notation for the inequality constraint functions different from (8.31), while we require for them the same *local*

Lipschitzian property with respect to x around $\bar{x}$ *uniformly* in $t \in T$: there are $K, \rho > 0$ such that

$$|f(x,t) - f(u,t)| \le K\|x - u\| \text{ for all } x, u \in B_\rho(\bar{x}),\ t \in T. \tag{8.109}$$

At the same time, contrary to Section 8.2, we assume here that the cost function φ is merely *l.s.c.* around the reference point $\bar{x}$ and that Ω is an arbitrary *locally closed* set around $\bar{x}$. Consider the collection of ε-*active indices*

$$T_\varepsilon(\bar{x}) := \left\{t \in T \middle|\ f(\bar{x},t) \ge -\varepsilon\right\} \text{ as } \varepsilon \ge 0 \text{ with } T(\bar{x}) := T_0(\bar{x}).$$

It follows from the uniform Lipschitz property of f in (8.109) that for any $\varepsilon > 0$ there is $\delta > 0$ sufficiently small such that $f(x,t) < 0$ whenever $x \in B_\delta(\bar{x})$ and $t \notin T_\varepsilon(\bar{x})$. This observation allows us to restrict the inequality constraints in (8.108) to the set $T_\varepsilon(\bar{x})$ with keeping all the local properties assumed around $\bar{x}$. Observe further from (8.108) that the function $t \mapsto f(x,t)$ is bounded on $T_\varepsilon(\bar{x})$ for each x around $\bar{x}$. These discussions show that there is no restriction to suppose that $f(x,\cdot)$ for $x \in X$ are elements of $l^\infty(T)$.

Using the function $f(x,t)$ of two variables from (8.108), define the mapping $f: X \to l^\infty(T)$ by $f(x)(\cdot) := f(x,\cdot) \in l^\infty(T)$ for all $x \in X$. It follows from (8.109) that this mapping is locally Lipschitzian around $\bar{x}$ as in (8.65). Further, it is easy to see that f is $l^\infty_+(T)$-convex if and only if all the functions $f(\cdot,t)$ as $t \in T$ are convex with respect to the variable x. Moreover, the strict differentiability of the mapping $f: X \to l^\infty(T)$ at $\bar{x}$ corresponds to the uniform strict differentiability of $x \mapsto f(x,t)$ at $\bar{x}$ for all $t \in T$ in the sense of Subsection 8.1.1. When the index set T is a compact Hausdorff space and the functions $p(\cdot) \in l^\infty(T)$ are restricted to be continuous on T, $l^\infty(T)$ reduces to the space of continuous functions $\mathcal{C}(T)$ with the maximum norm.

As discussed in Section 7.1, the Banach spaces $l^\infty(T)$ and $\mathcal{C}(T)$ are not Asplund. Furthermore, it is well known that $l^\infty(T)$ is never separable unless T is finite, while the space $\mathcal{C}(T)$ is separable provided that T is a compact metric space. Similarly to Section 7.1, we identify the dual space $l^\infty(T)^*$ with the space $ba(T)$ of bounded and additive measures $\mu(\cdot)$ on T satisfying

$$\langle \mu, p\rangle = \int_T p(t)\mu(dt) \text{ for any } \mu \in ba(T),\ p \in l^\infty(T)$$

with the dual norm on $ba(T)$ defined as the total variation of $\mu(\cdot)$ on T by

$$\|\mu\| := \sup_{A \subset T} \mu(A) - \inf_{B \subset T} \mu(B).$$

Denoting by $ba_+(T)$ the collection of nonnegative bounded and additive measures on T, we can easily check that

$$ba_+(T) = \left\{\mu \in ba(T) \middle|\ \int_T p(t)\mu(dt) \ge 0,\ p \in l^\infty_+(T)\right\},$$

where $l^\infty_+(T) := \{p \in l^\infty(T) \big|\ p_t \ge 0,\ t \in T\}$ is the positive cone in $l^\infty(T)$.

When T is a compact topological space, denote by $\mathcal{B}(T)$ the σ-algebra of all the Borel sets on T. As well known, the topologically dual space to $\mathcal{C}(T)$ is the space $rca(T)$ of all the regular finite real-valued Borel measures on T equipped with the total variation norm $\|\mu\|$. We define the collection of all the nonnegative regular Borel measures on T by

$$rca_+(T) := \big\{\mu \in rca(T) \big|\ \mu(A) \ge 0 \ \text{ as } \ A \in \mathcal{B}(T)\big\},$$

which can be equivalently described as

$$rca_+(T) = \Big\{\mu \in rca(T) \Big|\ \int_T p(t)\mu(dt) \ge 0 \ \text{ as } \ p \in \mathcal{C}_+(T)\Big\},$$

where $\mathcal{C}_+(T)$ is the set of all the nonnegative continuous functions on T. Recall that a Borel measure $\mu(\cdot)$ is *supported* on $A \in \mathcal{B}(T)$ if $\mu(B) = 0$ for all the sets $B \in \mathcal{B}(T)$ with $B \cap A = \emptyset$ and then observe the following statement.

Proposition 8.55 (Supported Measures). *Let T be a compact Hausdorff space, and let $p \in \mathcal{C}_+(T)$. If the measure $\mu \in rca_+(T)$ satisfies the relationship $\int_T p(t)\mu(dt) = 0$, then it is supported on the set $\{t \in T|\ p(t) = 0\}$.*

Proof. Define $A := \{t \in T|\ p(t) = 0\}$ and pick any $B \in \mathcal{B}(T)$ such that $B \cap A = \emptyset$. Since $\mu(\cdot)$ is a regular measure, we have

$$\mu(B) = \sup\big\{\mu(C) \big|\ C \subset B,\ C \text{ compact}\big\}.$$

To verify that $\mu(B) = 0$, we only need to show that $\mu(C) = 0$ for all the compact sets C contained in B. To proceed, define $\delta := \max\{p(t)|\ t \in C\} \ge 0$ and observe that $\delta > 0$ due to $C \cap A = \emptyset$. It follows that

$$0 = \int_T p(t)\mu(dt) = \int_{T\setminus C} p(t)\mu(dt) + \int_C p(t)\mu(dt) \ge \int_C p(t)\mu(dt) \ge \delta\mu(C) \ge 0,$$

which implies that $\mu(C) = 0$ and thus justifies the claimed result. $\triangle$

As discussed above, SIP (8.108) can be formulated as a cone-constrained program (8.64) with $Y = l^\infty(T)$ and $\Theta = l^\infty_+(T)$. Applying Theorem 8.45 yields the following optimality conditions for nonsmooth and nonconvex SIPs.

Theorem 8.56 (Necessary Optimality Conditions for Nonconvex SIPs with Arbitrary Index Sets). *Let $\bar{x}$ be a local optimal solution to SIP* (8.108) *under the standing assumptions of this subsection. For the constraint function $f(x,t)$ in* (8.108)*, define the collections of measures*

$$ba_+(T)(f) := \Big\{\mu \in ba_+(T) \Big|\ \mu(T) = 1,\ \int_T f(\bar{x},t)\mu(dt) = 0\Big\}$$

and suppose that the qualification conditions (8.77) *and*

$$\Big(\partial^\infty\varphi(\bar{x}) + N(\bar{x};\Omega)\Big) \bigcap \Big(-\breve{D}^*_{N,\Theta} f(\bar{x})\big(ba_+(T)(f)\big)\Big) = \emptyset \qquad (8.110)$$

are satisfied. Then there is a measure $\mu \in ba_+(T)$ *such that*

$$0 \in \partial\varphi(\bar{x}) + \breve{D}^*_{N,\Theta} f(\bar{x})(\mu) + N(\bar{x};\Omega) \text{ and } \int_T f(\bar{x},t)\mu(dt) = 0. \quad (8.111)$$

Proof. To deduce this result from Theorem 8.45, recall the remarkable classical fact from the geometry of Banach spaces telling us that $\text{int}\, l^\infty_+(T) \neq \emptyset$. It follows from the above discussions that in the notation of Corollary 8.47 specified to problem (8.108), we get $\text{int}\,\Theta \neq \emptyset$ and $\Theta^+ = ba_+(T)$. Furthermore,

$$\mu(T) \geq \|\mu\| \geq \langle \mu, e\rangle = \int_T \mu(dt) = \mu(T) \text{ for all } \mu \in ba_+(T),$$

where $e(\cdot)$ is the unit function of $l^\infty(T)$. This shows that $\Xi_0 = ba_+(T)(f)$, and hence the qualification condition (8.83) of Corollary 8.47 reduces to (8.110) for SIP (8.108). Then following the arguments of Corollary 8.47 in the setting under consideration, we arrive at (8.111) and thus complete the proof. △

When T is *compact*, the underlying space $Y = \mathcal{C}(T)$ is separable, and thus the unit ball of $\mathcal{C}^*(T) = rca(T)$ is *sequentially* weak* compact. This allows us to use the (sequential) normal coderivative $D^*_{N,\Theta} f(\bar{x})$ from Definition 8.43 to derive the corresponding necessary optimality conditions for SIP (8.108).

Corollary 8.57 (Optimality Conditions for Nonconvex SIPs with Compact Index Sets). *In the setting of Theorem 8.56, suppose that the index set* T *is a compact metric space and that the function* $t \mapsto f(x,t)$ *is continuous on* T *for each* $x \in X$. *Assume also that the qualification conditions* (8.77) *and*

$$\Big(\partial^\infty\varphi(\bar{x}) + N(\bar{x};\Omega)\Big) \bigcap \Big(- D^*_{N,\Theta} f(\bar{x})\big(rca_+(T)(f)\big)\Big) = \emptyset \quad (8.112)$$

are satisfied, where the coderivative argument in (8.112) *is defined by*

$$rca_+(T)(f) := \big\{\mu \in rca_+(T)\big|\ \mu(T) = 1,\ \mu \text{ is supported on } T(\bar{x})\big\}.$$

Then there is a measure $\mu \in rca_+(T)$ *supported on* $T(\bar{x})$ *such that*

$$0 \in \partial\varphi(\bar{x}) + D^*_{N,\Theta} f(\bar{x})(\mu) + N(\bar{x};\Omega). \quad (8.113)$$

Proof. Since the closed unit ball of $\mathcal{C}^*(T)$ is sequentially weak* compact, combining the last part in Corollary 8.47 with Proposition 8.55 ensures the existence of the claimed measure $\mu(\cdot)$ satisfying (8.113). △

Let us present a simple example illustrating the application of the qualification and optimality conditions from Corollary 8.57 as well as their behavior under the convexification of the involved generalized differential constructions.

Example 8.58 (Illustration of Qualification and Optimality Conditions for SIPs Over Compact Index Sets). Consider the following one-dimensional SIP (with $x \in \mathbb{R}$), where the cost function is smooth:

$$\text{minimize } \varphi(x) := x^2 \text{ subject to } f(x,t) := -|x| - t \le 0,\ t \in T := [0,1].$$

It is obvious that $\bar{x} = 0$ is the only minimizer for this problem with $T(\bar{x}) = \{0\}$. We can directly calculate the regular normal cone in this setting by

$$\widehat{N}\big((x, f(x)); \operatorname{epi}_\Theta f\big) = \begin{cases} \big\{(r, -\mu) \in \mathbb{R} \times rca_+(T) \,\big|\, r = -\mu(T)\big\} & \text{if } x > 0, \\ \big\{(r, -\mu) \in \mathbb{R} \times rca_+(T) \,\big|\, r = \mu(T)\big\} & \text{if } x < 0, \end{cases}$$

which tells us that $D^*_{N,\Theta} f(\bar{x})(\mu) = \{-\mu(T), \mu(T)\}$ for all $\mu \in rca_+(T)$. Thus the qualification condition (8.112) reads as

$$\partial^\infty \varphi(\bar{x}) \cap \Big(- D^*_{N,\Theta} f(\bar{x})\big(rca_+(T)(f)\big)\Big) = \{0\} \cap \{-1, 1\},$$

which obviously holds together with the necessary condition (8.113). This allows us to confirm the optimality of $\bar{x} = 0$ by Corollary 8.57. On the other hand, the corresponding convexified qualification condition

$$\operatorname{co} \partial^\infty \vartheta(\bar{x}) \cap \Big(- \operatorname{co} D^*_{N,\Theta} f(\bar{x})\big(rca_+(T)(f)\big)\Big) = \{0\} \cap [-1, 1] = \{0\} \neq \emptyset$$

fails and prevents using convexification to get rid of this nonoptimal solution.

The last result of this subsection presents applications of the metric regularity conditions for cone-constrained systems from Theorem 8.54 to the case of infinite inequality constraints from (8.108) under parameter perturbations.

Theorem 8.59 (Pointbased Characterizations of Metric Regularity for Infinite Inequality Systems). *Suppose that in the setting of Theorem 8.54, we have the infinite inequality system $F : X \rightrightarrows l^\infty(T)$ given by*

$$F(x) := \big\{p \in l^\infty(T) \,\big|\, f(x,t) \le p(t),\ t \in T\big\}, \quad x \in X,$$

with an arbitrary index set T. Pick $\bar{x} \in \ker F$ so that the qualification condition

$$\big(\ker \breve{D}^*_{N,\Theta} f(\bar{x})\big) \cap \big(ba_+(T)(f)\big) = \emptyset$$

is satisfied. Then F is metrically regular around $(\bar{x}, 0)$ and its exact regularity bound at $(\bar{x}, 0)$ is estimated from the above by

$$\operatorname{reg} F(\bar{x}, 0) \le \sup\left\{ \frac{1}{\|x^*\|} \,\middle|\, x^* \in \breve{D}^*_{N,\Theta} f(\bar{x})(\mu),\ \mu \in ba_+(T)(f)\right\}. \quad (8.114)$$

If further for all $t \in T$ the functions $x \mapsto f(x,t)$ are either convex or uniformly strictly differentiable at $\bar{x}$, then the equality holds in (8.114) and we have

$$\operatorname{reg} F(\bar{x}, 0) = \sup\big\{\|\mu\| \;\big|\; (\mu, -x^*) \in N\big((0, \bar{x}); \operatorname{gph} F^{-1}\big),\ \|x^*\| = 1\big\}$$

with the specifications of this formula, which are similar to Theorem 8.54.

Proof. Recall that $\operatorname{int} l^\infty_+(T) \neq \emptyset$. By Theorem 8.54 and the discussions above it is sufficient to check that the set $\Xi = \{\mu \in ba_+(T) |\ \|\mu\| = 1\}$ is weak* closed in $ba(T)$. To proceed, take any net $\{\mu_\nu\}_{\nu\in\mathcal{N}} \subset \Xi$ weak* converging to μ and show that $\mu \in \Xi$. Indeed, it follows that

$$1 = \lim_{\nu\in\mathcal{N}} \|\mu_\nu\| = \lim_{\nu\in\mathcal{N}} \mu_\nu(T) = \lim_{\nu\in\mathcal{N}} \langle\mu_\nu, e\rangle = \langle\mu, e\rangle = \mu(T) = \|\mu\|,$$

where $e(\cdot)$ is the unit function in $l^\infty(T)$. This verifies the claimed weak* closedness of the set $\Xi \subset ba(T)$ and thus completes the proof of the theorem. $\triangle$

8.4 Nonconvex SIPs with Countable Constraints

In the concluding section of this and the previous chapters devoted to SIPs, we develop yet another approach to SIPs based on applying the *extremal principles* for *countable* set systems obtained in Chapter 2. This approach allows us to establish necessary optimality conditions for SIPs with geometric constraints given by countably many nonconvex sets and then apply them to infinite inequality constraints described by general nonsmooth (may not be Lipschitzian) functions. The conditions obtained in this way provide new results even for smooth, convex, and Lipschitzian SIPs in comparison with those established in the previous sections of Chapters 7 and 8. To proceed, we derive some *calculus properties* of tangents and normals to countable nonconvex set intersections, which are of their own values in variational analysis. Throughout this section we suppose that the decision space X is *finite-dimensional* with the *Euclidean norm*, although a number of the obtained results can be extended to infinite-dimensional spaces. Recall also that our standing assumptions are, unless otherwise stated, that all the sets under consideration are locally *closed* and all the functions are *l.s.c.* around the reference points. It is possible to observe from the given proofs that the latter assumptions are not always needed; we leave it for the reader to check this as an exercise.

8.4.1 CHIP Properties for Countable Set Intersections

We start with the study of the so-called conical hull intersection property (CHIP) for countable nonconvex set intersections, which has been intensively investigated and applied for the case of finite intersections of convex sets; see more discussions and references in Section 8.6. In what follows we keep the terminology of convex analysis while replacing the classical tangent cone by its contingent counterpart (1.11) in the nonconvex setting. Furthermore, we formulate also the strong version of CHIP for nonconvex intersections expressed via our basic normal cone (1.4).

Definition 8.60 (CHIP for Countable Intersections). *Given an arbitrary set system $\{\Omega_i\}_{i\in\mathbb{N}} \subset X$ and $\bar{x} \in \bigcap_{i=1}^\infty \Omega_i$, it is said that:*

(i) *The system* $\{\Omega_i\}_{i\in\mathbb{N}}$ *has the* CONICAL HULL INTERSECTION PROPERTY (CHIP) *at the point* $\bar{x}$ *if we have*

$$T\Big(\bar{x};\bigcap_{i=1}^{\infty}\Omega_i\Big)=\bigcap_{i=1}^{\infty}T(\bar{x};\Omega_i). \tag{8.115}$$

(ii) *The system* $\{\Omega_i\}_{i\in\mathbb{N}}$ *has the* STRONG CHIP *at* $\bar{x}$ *if we have*

$$N\Big(\bar{x};\bigcap_{i=1}^{\infty}\Omega_i\Big)=\Big\{\sum_{i\in I}x_i^*\Big|\ x_i^*\in N(\bar{x};\Omega_i),\ \ I\in\mathcal{L}\Big\}, \tag{8.116}$$

where $\mathcal{L}$ *is the collection of all the finite subsets of the natural series* $\mathbb{N}$.

It can be checked that assuming the convexity of Ω_i in (8.116) allows us to equivalently represent the strong CHIP for $\{\Omega_i\}_{i\in\mathbb{N}}$ in the form

$$N\Big(\bar{x};\bigcap_{i=1}^{\infty}\Omega_i\Big)=\operatorname{co}\bigcup_{i=1}^{\infty}N(\bar{x};\Omega_i). \tag{8.117}$$

Further, we say that a countable set system $\{\Omega_i\}_{i\in\mathbb{N}}$ has the *asymptotic strong CHIP* at $\bar{x}\in\bigcap_{i=1}^{\infty}\Omega_i$ if the latter representation is replaced by

$$N\Big(\bar{x};\bigcap_{i=1}^{\infty}\Omega_i\Big)=\operatorname{cl}\operatorname{co}\bigcup_{i=1}^{\infty}N(\bar{x};\Omega_i). \tag{8.118}$$

The next result reveals the *equivalence* between CHIP and the asymptotic strong CHIP for countable intersections of convex sets.

Theorem 8.61 (Characterization of CHIP for Intersections of Convex Sets). *Let* $\{\Omega_i\}_{i\in\mathbb{N}}\subset X$ *be a countable system of convex sets with* $\bar{x}\in\bigcap_{i=1}^{\infty}\Omega_i$. *The following properties are equivalent:* **(a)** *The system* $\{\Omega_i\}_{i\in\mathbb{N}}$ *has CHIP at* $\bar{x}$.

(b) *The system* $\{\Omega_i\}_{i\in\mathbb{N}}$ *has the asymptotic strong CHIP at* $\bar{x}$.

In particular, the strong CHIP implies CHIP but not vice versa.

Proof. It is well known in convex analysis that the *full duality*

$$T(\bar{x};\Omega)=N^*(\bar{x};\Omega)\ \text{ and }\ N(\bar{x};\Omega)=T^*(\bar{x};\Omega),\quad x\in\Omega, \tag{8.119}$$

holds for convex sets. Let us now justify the equality

$$\Big(\bigcap_{i=1}^{\infty}T(\bar{x};\Omega_i)\Big)^*=\operatorname{cl}\operatorname{co}\bigcup_{i=1}^{\infty}N(\bar{x};\Omega_i). \tag{8.120}$$

The inclusion "$\supset$" in (8.120) follows from the second formula in (8.119) by

$$N(\bar{x};\Omega_i)=T^*(\bar{x};\Omega_i)\subset\Big(\bigcap_{i=1}^{\infty}T(\bar{x};\Omega_i)\Big)^*$$

due to the closedness and convexity of the polar set on the right-hand side of the latter inclusion. To prove the opposite inclusion, pick $x^*\notin \operatorname{cl}\operatorname{co}\bigcup_{i=1}^{\infty}N(\bar{x};\Omega_i)$ and find by convex separation such $v\neq 0$ that

$$\langle x^*,v\rangle>0\ \text{ and }\ \langle u^*,v\rangle\le 0\ \text{ for all }\ u^*\in\operatorname{cl}\operatorname{co}\bigcup_{i=1}^{\infty}N(\bar{x};\Omega_i). \tag{8.121}$$

Hence for each $i\in\mathbb{N}$ we get $\langle u^*,v\rangle\le 0$ whenever $u^*\in N(\bar{x};\Omega_i)$, which yields $v\in N^*(\bar{x};\Omega_i)$ and therefore $v\in T(\bar{x};\Omega_i)$ by the first formula in (8.119). This tells us that $v\in\bigcap_{i=1}^{\infty}T(\bar{x};\Omega_i)$, and so $x^*\notin\big(\bigcap_{i=1}^{\infty}T(\bar{x};\Omega_i)\big)^*$ by $\langle x^*,v\rangle>0$ in (8.121) verifying the equality in (8.120). Since the set $\cap_{i=1}^{\infty}T(\bar{x};\Omega_i)$ is closed and convex, it agrees with its second dual, which ensures the representation

$$\bigcap_{i=1}^{\infty}T(\bar{x};\Omega_i)=\Big(\operatorname{cl}\operatorname{co}\bigcup_{i=1}^{\infty}N(\bar{x};\Omega_i)\Big)^*. \tag{8.122}$$

Assuming that CHIP in (a) holds and employing (8.119) together with (8.120) for the set intersection $\Omega:=\bigcap_{i=1}^{\infty}\Omega_i$ yield the equalities

$$N(\bar{x};\Omega)=T^*(\bar{x};\Omega)=\Big(\bigcap_{i=1}^{\infty}T(\bar{x};\Omega_i)\Big)^*=\operatorname{cl}\operatorname{co}\bigcup_{i=1}^{\infty}N(\bar{x};\Omega_i),$$

which justify the validity of the asymptotic strong CHIP in (b). Conversely, having (b) and using the relationships in (8.119) and (8.122) give us

$$T(\bar{x};\Omega)=N^*(\bar{x};\Omega)=\Big(\operatorname{cl}\operatorname{co}\bigcup_{i=1}^{\infty}N(\bar{x};\Omega_i)\Big)^*=\bigcap_{i=1}^{\infty}T(\bar{x};\Omega_i),$$

which verifies CHIP in (a) and thus establishes the equivalence statement claimed in the theorem. Since the strong CHIP implies the asymptotic strong CHIP due to the closedness of $N(\bar{x};\Omega)$, it also implies CHIP. The converse implication doesn't hold even for finitely many sets; see Exercise 8.111(ii). △

As a direct consequence of Theorem 8.61, we obtain an unconditional representation of the normal cone to solutions of infinite linear systems.

Corollary 8.62 (Normal Cone to Solution Sets for Countable Linear Inequality Systems). *The normal cone at the origin to the solution set*

$$\Omega:=\big\{x\in X\big|\ \langle a_i,x\rangle\le 0,\ i\in\mathbb{N}\big\},$$

of countable linear inequalities is calculated by

$$N(0;\Omega)=\mathrm{cl}\,\mathrm{co}\Big[\bigcup_{i=1}^{\infty}\big\{\lambda a_i\big|\ \lambda\ge 0\big\}\Big]. \tag{8.123}$$

Proof. It is easy to see that Ω is represented as a countable intersection of sets having CHIP. The asymptotic strong CHIP for this system is obviously (8.123). Thus the result follows immediately from Theorem 8.61. △

Of course, we cannot expect extending the equivalence of Theorem 8.61 to nonconvex sets. Let us now derive some sufficient conditions ensuring CHIP for countable intersections of nonconvex sets. The first step in this direction is to employ the notion of *bounded linear regularity* for countable systems of nonconvex sets, which goes beyond its conventional study and applications for convex systems. This notion is certainly important for its own sake, regardless of its subsequent applications to CHIP; see Section 8.6.

Definition 8.63 (Bounded Linear Regularity for Countable Systems of Nonconvex Sets). *Given a set system $\{\Omega_i\}_{i\in\mathbb{N}}$, we say that it is* BOUNDEDLY LINEARLY REGULAR *at $\bar x\in\Omega:=\bigcap_{i=1}^{\infty}\Omega_i$ if there exist a neighborhood U of $\bar x$ and a number $C>0$ such that*

$$\mathrm{dist}\,(x;\Omega)\le C\sup_{i\in\mathbb{N}}\big\{\mathrm{dist}\,(x;\Omega_i)\big\}\ \textit{ for all }\ x\in U. \tag{8.124}$$

In the next proposition, the notation $d_\Omega(x)$:=dist$(x;\Omega)$ is used for convenience.

Proposition 8.64 (Sufficient Conditions for CHIP of Countable Set Systems in Terms of Bounded Linear Regularity). *Let $\{\Omega_i\}_{i\in\mathbb{N}}\subset X$ be a countable system of sets with $\bar x\in\Omega:=\bigcap_{i=1}^{\infty}\Omega_i$. Assume that this system is boundedly linearly regular at $\bar x$ with some $C>0$ in (8.124) and that the family of functions $\{d_{\Omega_i}(\cdot)\}_{i\in\mathbb{N}}$ is* EQUIDIRECTIONALLY DIFFERENTIABLE *at $\bar x$ in the sense that for any $h\in X$ the functions*

$$\Big\{\frac{d_{\Omega_i}(\bar x+th)}{t},\ i\in\mathbb{N}\Big\}$$

of $t>0$ converge as $t\downarrow 0$ to the corresponding directional derivatives $d'_{\Omega_i}(\bar x;h)$ uniformly in $i\in\mathbb{N}$. Then for all $h\in X$ we have the estimate

$$\mathrm{dist}\,(h;\Lambda)\le C\sup_{i\in\mathbb{N}}\big\{\mathrm{dist}\,(h;\Lambda_i)\big\}\ \textit{with}\ \Lambda:=T(\bar x;\Omega)\ \ \textit{and}\ \ \Lambda_i:=T(\bar x;\Omega_i)$$

as $i\in\mathbb{N}$. In particular, the set system $\{\Omega_i\}_{i\in\mathbb{N}}$ enjoys CHIP at $\bar x$.

Proof. Recalling definition (1.11) of $T(\bar x;\Omega)$, we get (see Exercise 8.114) that

$$\mathrm{dist}\,(h;\Lambda)=\liminf_{t\downarrow 0}\mathrm{dist}\Big(h;\frac{\Omega-\bar x}{t}\Big)=\liminf_{t\downarrow 0}\frac{\mathrm{dist}\,(\bar x+th;\Omega)}{t}. \tag{8.125}$$

When t is small, the assumed bounded linear regularity yields

$$\frac{\operatorname{dist}(\bar{x}+th;\Omega)}{t}\le C\sup_{i\in\mathbb{N}}\frac{\operatorname{dist}(\bar{x}+th;\Omega_i)}{t}.$$

Applying further the equidirectional differentiability ensures the convergence

$$\frac{\operatorname{dist}(\bar{x}+th;\Omega_i)}{t}\to d'_{\Omega_i}(\bar{x};h)=\operatorname{dist}(h;\Lambda_i)\ \text{ uniformly in }\ i\ \text{ as }\ t\downarrow 0,$$

i.e., for any $\varepsilon>0$ there exists $\delta>0$ such that whenever $t\in(0,\delta)$ we have

$$\left|\frac{\operatorname{dist}(\bar{x}+th;\Omega_i)}{t}-\operatorname{dist}(h;\Lambda_i)\right|\le\varepsilon\ \text{ for all }\ i\in\mathbb{N}.$$

Hence it follows for any $t\in(0,\delta)$ that

$$\sup_{i\in\mathbb{N}}\frac{\operatorname{dist}(\bar{x}+th;\Omega_i)}{t}\le\sup_{i\in\mathbb{N}}\big\{\operatorname{dist}(h;\Lambda_i)\big\}+\varepsilon.$$

Combining all the above, we get the estimates

$$\operatorname{dist}(h;\Lambda)\le C\liminf_{t\downarrow 0}\sup_{i\in\mathbb{N}}\frac{\operatorname{dist}(\bar{x}+th;\Omega_i)}{t}\le C\sup_{i\in\mathbb{N}}\big\{\operatorname{dist}(h;\Lambda_i)\big\}+C\varepsilon,$$

which yield (8.124) by the arbitrary choice of ε and thus verify CHIP. △

The next result simplifies checking bounded linear regularity.

Corollary 8.65 (CHIP via Simplified Bounded Linear Regularity of Nonconvex Sets). *Without assuming bounded linear regularity in the framework of Proposition 8.64, suppose that there are numbers $C>0$, $j\in\mathbb{N}$, and a neighborhood U of $\bar{x}$ such that*

$$\operatorname{dist}(x;\Omega)\le C\sup_{i\ne j}\big\{\operatorname{dist}(x;\Omega_i)\big\}\ \textit{for all } x\ \in\Omega_j\cap U.$$

Then the set system $\{\Omega_i\}_{i\in\mathbb{N}}$ enjoys CHIP at $\bar{x}$.

Proof. Employing Proposition 8.64, it suffices to show that the system $\{\Omega_i\}_{i\in\mathbb{N}}$ is boundedly linearly regular at $\bar{x}$. Take $r>0$ so small that

$$\operatorname{dist}(x;\Omega)\le C\sup_{i\ne j}\big\{\operatorname{dist}(x;\Omega_i)\big\}\ \text{ for all }\ x\in\Omega_j\cap(\bar{x}+3r\mathbb{B}).$$

Since the distance function is nonexpansive (i.e., Lipschitzian with modulus $\ell=1$), for every $y\in\Omega_j\cap(\bar{x}+3r\mathbb{B})$ and $x\in X$ we have

$$\begin{aligned} 0 &\le C \sup_{i\neq j}\big\{\text{dist}\,(y;\Omega_i)\big\} - \text{dist}\,(y;\Omega)\\ &\le C \sup_{i\neq j}\Big(\big\{\text{dist}\,(x;\Omega_i)\big\} + \|x-y\|\Big) - \text{dist}\,(x;\Omega) + \|x-y\|\\ &\le C \sup_{i\neq j}\big\{\text{dist}\,(x;\Omega_i)\big\} - \text{dist}\,(x;\Omega) + (C+1)\|x-y\|, \end{aligned}$$

which readily ensures the estimate

$$\text{dist}\,(x;\Omega) \le (2C+1)\max\Big[\sup_{i\neq j}\big\{\text{dist}\,(x;\Omega_i)\big\},\ \text{dist}\big(x;\Omega_j\cap(\bar{x}+3r\mathbb{B})\big)\Big].$$

Thus the bounded linear regularity of $\{\Omega_i\}_{i\in\mathbb{N}}$ at $\bar{x}$ in the form of

$$\text{dist}\,(x;\Omega) \le (2C+1)\sup_{i\in\mathbb{N}}\big\{\text{dist}\,(x;\Omega_i)\big\}$$

would follow now from the relationship

$$\text{dist}\big(x;\Omega_j\cap(\bar{x}+3r\mathbb{B})\big) = \text{dist}\,(x;\Omega_j)\ \text{ for all }\ x\in\bar{x}+r\mathbb{B}. \tag{8.126}$$

To verify (8.126), fix a vector $x\in\bar{x}+r\mathbb{B}$ above and pick any $y\in\Omega_j\setminus(\bar{x}+3r\mathbb{B})$. This gives us $\|x-y\|\ge\|y-\bar{x}\|-\|\bar{x}-x\|\ge 3r-r=2r$ and implies that

$$\text{dist}\big(x;\Omega_j\setminus(\bar{x}+3r\mathbb{B})\big)\ge 2r\ \text{ while }\ \text{dist}\big(x;\Omega_j\cap(\bar{x}+3r\mathbb{B})\big)\le\|x-\bar{x}\|\le r.$$

Hence we get the equalities

$$\begin{aligned} \text{dist}\,(x;\Omega_j) &= \min\Big\{\text{dist}\big(x;\Omega_j\setminus(\bar{x}+3r\mathbb{B})\big),\ \text{dist}\big(x;\Omega_j\cap(\bar{x}+3r\mathbb{B})\big)\Big\}\\ &= \text{dist}\big(x;\Omega_j\cap(\bar{x}+3r\mathbb{B})\big), \end{aligned}$$

which readily justify (8.126) and thus complete the proof of the corollary. △

The next proposition, which holds in fact for arbitrary (not only countable) set intersections, provides a new kind of sufficient conditions for CHIP. Define the *tangential rank* of the intersection $\Omega := \bigcap_{i=1}^{\infty}\Omega_i$ at $\bar{x}\in\Omega$ by

$$\rho_\Omega(\bar{x}) := \inf_{i\in\mathbb{N}}\left\{\limsup_{\substack{x\to\bar{x}\\ x\in\Omega_i\setminus\{\bar{x}\}}}\frac{\text{dist}\,(x;\Omega)}{\|x-\bar{x}\|}\right\},$$

where we put $\rho_\Omega(\bar{x}) := 0$ if $\Omega_i=\{\bar{x}\}$ for at least one $i\in\mathbb{N}$.

Proposition 8.66 (Sufficient Condition for CHIP via Tangential Rank of Intersections). *Given a countable set system $\{\Omega_i\}_{i\in\mathbb{N}}\subset X$ with a common point $\bar{x}$, suppose that $\rho_\Omega(\bar{x})=0$ for the tangential rank of $\Omega := \bigcap_{i=1}^{\infty}\Omega_i$ at $\bar{x}\in\Omega$. Then this system exhibits CHIP at the point $\bar{x}$.*

Proof. Since the result holds trivially if $\Omega_i=\{\bar{x}\}$ for some $i\in\mathbb{N}$, suppose that $\Omega_i\setminus\{\bar{x}\}\neq\emptyset$ for all $i\in\mathbb{N}$ and observe that $T(\bar{x};\Omega)\subset T(\bar{x};\Omega_i)$ whenever $i\in\mathbb{N}$. Thus we always have the inclusion

$$T(\bar{x};\Omega)\subset\bigcap_{i\in\mathbb{N}}T(\bar{x};\Omega_i).$$

To verify the opposite inclusion, pick $0\neq v\in\bigcap_{i=1}^{\infty}T(\bar{x};\Omega_i)$ and deduce from $\rho_\Omega(\bar{x})=0$ by the rank definition that for any fixed $k\in\mathbb{N}$ there is Ω_k with

$$\limsup_{\substack{x\to\bar{x}\\ x\in\Omega_k\setminus\{\bar{x}\}}}\frac{\text{dist}\,(x;\Omega)}{\|x-\bar{x}\|}<\frac{1}{k}.$$

Since $v\in T(\bar{x};\Omega_k)$, there exist sequences $\{x_j\}_{j\in\mathbb{N}}\subset\Omega_k$ and $t_j\downarrow 0$ satisfying

$$x_j\to\bar{x}\ \text{ and }\ \frac{x_j-\bar{x}}{t_j}\to v\ \text{ as }\ j\to\infty,$$

which in turn yields the limiting estimate

$$\limsup_{j\to\infty}\frac{\text{dist}\,(x_j;\Omega)}{\|x_j-\bar{x}\|}<\frac{1}{k}.$$

The latter gives us $x_k\in\{x_j\}_{j\in\mathbb{N}}$ with $\|x_k-\bar{x}\|\le 1/k$ and $t_k\le 1/k$ with

$$\left\|\frac{x_k-\bar{x}}{t_k}-v\right\|\le\frac{1}{k}\ \text{ and }\ \frac{\text{dist}\,(x_k;\Omega)}{\|x_k-\bar{x}\|}<\frac{1}{k}.$$

Then it follows that there exists $z_k\in\Omega$ satisfying the relationships

$$\|z_k-x_k\|<\frac{1}{k}\|x_k-\bar{x}\|\le\frac{1}{k^2}.$$

Combining the estimates above, we arrive at

$$\left\|\frac{z_k-\bar{x}}{t_k}-v\right\|\le\left\|\frac{z_k-x_k}{t_k}\right\|+\left\|\frac{x_k-\bar{x}}{t_k}-v\right\|\le\frac{1}{k}\Big(\|v\|+\frac{1}{k}\Big)+\frac{1}{k}$$

for all $k\in\mathbb{N}$. Now letting $k\to\infty$ gives us $z_k\xrightarrow{\Omega}\bar{x}$, $t_k\downarrow 0$, and a set $\left\|\frac{z_k-\bar{x}}{t_k}-v\right\|\longrightarrow 0$. Thus $v\in T(\bar{x};\Omega)$, which completes the proof. △

To conclude our discussions on CHIP, let us establish yet another verifiable condition ensuring the fulfillment of this property for countable intersections. We say that $A\subset X$ is of the *invex type* if it can be represented as the complement to a union with respect to $t\in T$ of some open convex sets A_t, i.e.,

$$A=X\setminus\bigcup_{t\in T}A_t. \tag{8.127}$$

Proposition 8.67 (CHIP for Countable Intersections of Invex-Type Sets). *Given a countable system $\{\Omega_i\}_{i\in\mathbb{N}}\subset\mathbb{R}^n$, assume that there is a (possibly infinite) index* *subset $J \subset \mathbb{N}$ such that each Ω_i for $i \in J$ is the complement to an open and convex set in X and that for some $\bar{x} \in X$ we have*

$$\bar{x} \in \Big(\bigcap_{i \in J} \operatorname{bd} \Omega_i\Big) \cap \operatorname{int} \bigcap_{i \notin J} \Omega_i. \tag{8.128}$$

Then the system $\{\Omega_i\}_{i \in \mathbb{N}}$ enjoys CHIP at $\bar{x}$.

Proof. Let us first show that for A of the invex type (8.127) the inclusion

$$\bar{x} + T(\bar{x}; A) \subset A \quad \text{whenever} \quad \bar{x} \in \bigcap_{t \in T} \operatorname{bd} A_t \cap \operatorname{bd} A \tag{8.129}$$

involving the contingent cone $T(\bar{x}; A)$ holds. On the contrary, suppose that there is $v \in T(\bar{x}; A)$ with $\bar{x} + v \notin A$. By (1.11) we find sequences $s_k \downarrow 0$ and $x_k \in A$ such that $\frac{x_k - \bar{x}}{s_k} \to v$. Since $\bar{x} + v \notin A$, the invexity assumption (8.127) gives us an index $t_0 \in T$ for which $\bar{x} + v \in A_{t_0}$. Thus we get

$$\bar{x} + \frac{x_k - \bar{x}}{s_k} \in A_{t_0} \quad \text{for all} \quad k \in \mathbb{N} \quad \text{sufficiently large.}$$

Then employing the convexity of A_{t_0} tells us that

$$x_k = (1 - s_k)\bar{x} + s_k\Big(\bar{x} + \frac{x_k - \bar{x}}{s_k}\Big) \in A_{t_0}$$

for the fixed index $t_0 \in T$ and all large numbers $k \in \mathbb{N}$. This contradicts the fact that of $x_k \in A$ and thus justifies the claimed inclusion (8.129).

To verify that $\{\Omega_i\}_{i \in \mathbb{N}}$ enjoys CHIP at $\bar{x}$ satisfying (8.128), take any Ω_i with $i \in J$ and consider $A \subset X$ with $\Omega = X \setminus A$. Then $\bar{x} \in \operatorname{bd} A \cap \operatorname{bd} \Omega_i$ by (8.128), and thus (8.129) ensures that $\bar{x} + T(\bar{x}; \Omega_i) \subset \Omega_i$ for this index $i \in J$. By the choice of $\bar{x}$ in (8.128), we have furthermore that

$$\bigcap_{i=1}^{\infty} T(\bar{x}; \Omega_i) = \bigcap_{i \in J} T(\bar{x}; \Omega_i) \subset \bigcap_{i \in J} (\Omega_i - \bar{x}).$$

Since the set on the left-hand side above is a cone, it follows that

$$\bigcap_{i=1}^{\infty} T(\bar{x}; \Omega_i) \subset T\Big(0; \bigcap_{i \in J} (\Omega_i - \bar{x})\Big) = T\Big(\bar{x}; \bigcap_{i \in J} \Omega_i\Big) = T\Big(\bar{x}; \bigcap_{i=1}^{\infty} \Omega_i\Big).$$

The opposite inclusion is obvious, and thus we justify CHIP at $\bar{x}$. △

The following consequence of Proposition 8.67 holds for linear systems.

Corollary 8.68 (CHIP for Countable Linear Systems). *Consider the set system* $\{\Omega_i\}_{i\in\mathbb{N}}$ *defined by countably many linear inequalities*

$$\Omega_i := \big\{x \in X\big|\ \langle a_i, x\rangle \le b_i\big\}.$$

Given a point $\bar{x} \in \Omega$ *and the active index set* $J(\bar{x})$*, suppose that*

$$\bar{x} \in \operatorname{int}\big\{x \in X\big|\ \langle a_i, x\rangle \le b_i,\ \ i \in \mathbb{N}\setminus J(\bar{x})\big\}.$$

Then the countable linear system $\{\Omega_i\}_{i\in\mathbb{N}}$ *enjoys CHIP at* $\bar{x}$*.*

Proof. It follows directly from Proposition 8.67. △

8.4.2 Generalized Normals to Countable Set Intersections

Now we proceed with another ingredient needed for applications to SIPs with countable geometric and nonsmooth inequality constraints. It relates to calculus rules for generalized normals to countable intersections of nonconvex sets under appropriate dual-space qualification conditions. Needless to say that results of this type are of their independent interest in variational analysis. Some developments in this direction are given in Sections 7.5 and 8.1 for sets with special convex and smooth structure. Our approach here is based on the *conic extremal principle* for countable systems of sets established in Theorem 2.9. It leads us, in particular, to new results in comparison with the aforementioned ones even for sets with the convex and smooth structures investigated therein.

First we formulate and discuss appropriate qualification conditions for countable systems of sets in terms of basic normals. In what follows the symbol $\mathcal{L}$ signifies the collection of all the *finite* subsets of the natural series $\mathbb{N}$. Recall also we are in the finite-dimensional Euclidean setting of $X = X^*$.

Definition 8.69 (Normal Closedness and Qualification Conditions for Countable Set Systems). *Let* $\{\Omega_i\}_{i\in\mathbb{N}} \subset X$ *be a countable system of nonempty sets, and let* $\bar{x} \in \bigcap_{i=1}^{\infty}\Omega_i$*. We say that:*

(a) *The set system* $\{\Omega_i\}_{i\in\mathbb{N}}$ *satisfies the* NORMAL CLOSEDNESS CONDITION (NCC) *at* $\bar{x}$ *if the combination of basic normals*

$$\Big\{\sum_{i\in I} x_i^*\Big|\ x_i^* \in N(\bar{x};\Omega_i),\ \ I \in \mathcal{L}\Big\}\ \ \text{is closed in}\ \ X^*. \tag{8.130}$$

(b) *The system* $\{\Omega_i\}_{i\in\mathbb{N}}$ *satisfies the* NORMAL QUALIFICATION CONDITION (NQC) *at* $\bar{x}$ *if the following implication holds:*

$$\Big[\sum_{i=1}^{\infty} x_i^* = 0,\ \ x_i^* \in N(\bar{x};\Omega_i)\Big] \Longrightarrow \Big[x_i^* = 0\ \ \text{for all}\ \ i \in \mathbb{N}\Big]. \tag{8.131}$$

The normal closedness condition (8.130) is of the Farkas-Minkowski type considered above for linear, convex, and differentiable infinite systems. The normal qualification condition (8.131) is an extension to countable systems of the condition with the same name for two and finitely many sets used in Section 2.4 to derive representations of basic normals to finite set intersections.

The next proposition presents a simple sufficient condition for the validity of NQC in the case of countable systems of convex sets.

Proposition 8.70 (NQC for Countable Systems of Convex Sets). *Let $\{\Omega_i\}_{i\in\mathbb{N}}$ be a system of convex sets for which there is an index $i_0 \in \mathbb{N}$ with*

$$\Omega_{i_0} \cap \bigcap_{i\neq i_0} \operatorname{int}\Omega_i \neq \emptyset. \tag{8.132}$$

Then NQC (8.131) is satisfied for the system $\{\Omega_i\}_{i\in\mathbb{N}}$ at any $\bar{x} \in \bigcap_{i=1}^{\infty} \Omega_i$.

Proof. Suppose without loss of generality that $i_0 = 1$ and fix some $w \in \Omega_1 \cap \bigcap_{i=2}^{\infty} \operatorname{int}\Omega_i$. Taking any normals $x_i^* \in N(\bar{x}; \Omega_i)$ as $i \in \mathbb{N}$ with

$$\sum_{i=1}^{\infty} x_i^* = 0,$$

we get by the convexity of Ω_i that $\langle x_i^*, w - \bar{x}\rangle \le 0$ for all $i \in \mathbb{N}$. It shows that

$$\langle x_i^*, w - \bar{x}\rangle = -\sum_{j\neq i}\langle x_j^*, w - \bar{x}\rangle \ge 0, \quad i \in \mathbb{N},$$

which yields $\langle x_i^*, w - \bar{x}\rangle = 0$ whenever $i \in \mathbb{N}$. Picking $u \in X$ with $\|u\| = 1$ and taking into account that $w \in \cap_{i=2}^{\infty}(\operatorname{int}\Omega_i)$ give us

$$\lambda\langle x_i^*, u\rangle = \langle x_i^*, w + \lambda u - \bar{x}\rangle \le 0, \quad i = 2, 3, \ldots,$$

if $\lambda > 0$ is sufficiently small. Due to the arbitrary choice of the unit vector u, it follows that $x_i^* = 0$ for $i = 2, 3, \ldots$ and therefore $x_i^* = 0$ for all $i \in \mathbb{N}$. △

Our next goal is to establish a certain "fuzzy" representation of regular normals to countable intersections of nonconvex cones via basic normals to the cones in question. To proceed in this direction, we first observe a simple while useful relationship between regular and basic normals to arbitrary cones.

Lemma 8.71 (Generalized Normals to Cones). *Let $\Lambda \subset X$ be a cone with $w \in \Lambda$. Then we have the inclusion $\widehat{N}(w; \Lambda) \subset N(0; \Lambda)$.*

Proof. Pick a regular normal $x^* \in \widehat{N}(w; \Lambda)$ and get its definition that

$$\limsup_{x \xrightarrow{\Lambda} w} \frac{\langle x^*, x - w\rangle}{\|x - w\|} \le 0.$$

Fix $x \in \Lambda$, $t > 0$ and let $u := x/t$. Then $(x/t) \in \Lambda$, $tw \in \Lambda$, and

$$\limsup_{x \xrightarrow{\Lambda} tw} \frac{\langle x^*, x - tw\rangle}{\|x - tw\|} = \limsup_{x \xrightarrow{\Lambda} w} \frac{t\langle x^*, (x/t) - w\rangle}{t\|(x/t) - w\|} = \limsup_{u \xrightarrow{\Lambda} w} \frac{\langle x^*, u - w\rangle}{\|u - w\|} \le 0,$$

which yields $x^* \in \widehat{N}(tw; \Lambda)$. Letting $t \to 0$, we arrive at $x^* \in N(0; \Lambda)$. $\triangle$

Now we are ready to obtain the aforementioned fuzzy representation.

Theorem 8.72 (Fuzzy Representation of Regular Normals to Countable Intersections of Cones). *Let $\{\Lambda_i\}_{i\in\mathbb{N}}$ be a countable system of cones in X satisfying the normal qualification condition (8.131) at $\bar{x} = 0$. Then given a regular normal $x^* \in \widehat{N}\big(0; \bigcap_{i=1}^{\infty} \Lambda_i\big)$ and a number $\varepsilon > 0$, there are basic normals $x_i^* \in N(0; \Lambda_i)$ as $i \in \mathbb{N}$ such that we have the inclusion*

$$x^* \in \sum_{i=1}^{\infty} \frac{1}{2^i} x_i^* + \varepsilon \mathbb{B}^*. \tag{8.133}$$

Proof. Fix $x^* \in \widehat{N}\big(0; \bigcap_{i=1}^{\infty} \Lambda_i\big)$, $\varepsilon > 0$ and get by the choice of x^* that

$$\langle x^*, x\rangle - \varepsilon\|x\| < 0 \;\text{ whenever }\; x \in \bigcap_{i=1}^{\infty} \Lambda_i \setminus \{0\}. \tag{8.134}$$

Define a countable system of closed cones in $X \times \mathbb{R}$ by

$$\begin{aligned} O_1 &:= \big\{(x,\alpha)\big|\; x \in \Lambda_1,\; \alpha \le \langle x^*, x\rangle - \varepsilon\|x\|\big\}, \\ O_i &:= \Lambda_i \times \mathbb{R}_+ \quad \text{for} \quad i \ge 2. \end{aligned} \tag{8.135}$$

Let us check that all the assumptions needed for the validity of the conic extremal principle in Theorem 2.9 are satisfied for the system $\{O_i\}_{i\in\mathbb{N}}$. Picking any $(x,\alpha) \in \bigcap_{i=1}^{\infty} O_i$, we have $x \in \bigcap_{i=1}^{\infty} \Lambda_i$ and $\alpha \ge 0$ from the construction of Ω_i as $i \ge 2$. This implies in fact that $(x,\alpha) = (0,0)$. Indeed, supposing $x \ne 0$ gives us by (8.134) that

$$0 \le \alpha \le \langle x^*, x\rangle - \varepsilon\|x\| < 0,$$

which is a contradiction. On the other hand, we get from $(0,\alpha) \in O_1$ by (8.135) that $\alpha \le 0$, i.e., $\alpha = 0$. Thus the nonoverlapping condition

$$\bigcap_{i=1}^{\infty} O_i = \{(0,0)\}$$

holds for $\{O_i\}_{i\in\mathbb{N}}$. Similarly we check that

$$\Big(O_1 - (0,\gamma)\Big) \cap \bigcap_{i=2}^{\infty} O_i = \emptyset \;\text{ for any fixed }\; \gamma > 0, \tag{8.136}$$

which says that $\{O_i\}_{i\in\mathbb{N}}$ is a *conic extremal system* at the origin. Indeed, violating of (8.136) means the existence of $(x,\alpha)\in X\times\mathbb{R}$ such that

$$(x,\alpha)\in\Big[O_1-(0,\gamma)\Big]\cap\bigcap_{i=2}^{\infty}O_i,$$

which yields $x\in\bigcap_{i=1}^{\infty}O_i$ and $\alpha\ge 0$. This tells us by (8.135) that

$$\gamma+\alpha\le\langle x^*,x\rangle-\varepsilon\|x\|\le 0,$$

a contradiction due to the positivity of γ in (8.136). Applying now Theorem 2.9 to the system $\{O_i\}_{i\in\mathbb{N}}$ gives us the pairs $(w_i,\alpha_i)\in O_i$ and $(x_i^*,\lambda_i)\in\widehat{N}\big((w_i,\alpha_i);O_i\big)$ as $i\in\mathbb{N}$ satisfying the relationships

$$\sum_{i=1}^{\infty}\frac{1}{2^i}\big(x_i^*,\lambda_i\big)=0\ \text{ and }\ \sum_{i=1}^{\infty}\frac{1}{2^i}\big\|(x_i^*,\lambda_i)\big\|^2=1. \tag{8.137}$$

It follows from the constructions of O_i as $i\ge 2$ that $\lambda_i\le 0$ and $x_i^*\in\widehat{N}(w_i;\Lambda_i)$; thus $x_i^*\in N(0;\Lambda_i)$ for $i=2,3,\ldots$ by Lemma 8.71. Furthermore, we get

$$\limsup_{(x,\alpha)\xrightarrow{O_1}(w_1,\alpha_1)}\frac{\langle x_1^*,x-w_1\rangle+\lambda_1(\alpha-\alpha_1)}{\|x-w_1\|+|\alpha-\alpha_1|}\le 0, \tag{8.138}$$

which readily implies by the construction of O_1 in (8.135) that $\lambda_1\ge 0$ and

$$\alpha_1\le\langle x^*,w_1\rangle-\varepsilon\|w_1\|. \tag{8.139}$$

Let us further examine the two possible cases: $\lambda_1=0$ and $\lambda_1>0$.

Case 1: $\lambda_1=0$. If inequality (8.139) is strict, we have

$$\alpha_1<\langle x^*,x\rangle-\varepsilon\|x\|\ \text{ for all }\ x\in U$$

for some neighborhood U of w_1, which yields $(x,\alpha_1)\in O_1$ for all $x\in\Lambda_1\cap U$. Plugging (x,α_1) into (8.138) gives us

$$\limsup_{x\xrightarrow{\Lambda_1}w_1}\frac{\langle x_1^*,x-w_1\rangle}{\|x-w_1\|}\le 0,$$

i.e., $x_1^*\in\widehat{N}(w_1;\Lambda_1)$. If (8.139) holds as equality, we get

$$|\alpha-\alpha_1|=\big|\langle x^*,x-w_1\rangle+\varepsilon(\|w_1\|-\|x\|)\big|\le\big(\|x^*\|+\varepsilon\big)\|x-w_1\|$$

by putting $\alpha := \langle x^*, x\rangle - \varepsilon\|x\|$. Furthermore, it follows from (8.138) that

$$\limsup_{(x,\alpha)\xrightarrow{O_1}(w_1,\alpha_1)} \frac{\langle x_1^*, x - w_1\rangle}{\|x - w_1\| + |\alpha - \alpha_1|} \le 0,$$

and hence for any $\nu > 0$ sufficiently small and α chosen above we have

$$\langle x_1^*, x - w_1\rangle \le \nu\big(\|x - w_1\| + |\alpha - \alpha_1|\big) \le \nu\big(1 + \|x^*\| + \varepsilon\big)\|x - w_1\|$$

whenever $x \in \Lambda_1$ is sufficiently closed to w_1. The latter implies that

$$\limsup_{x\xrightarrow{\Lambda_1}w_1} \frac{\langle x_1^*, x - w_1\rangle}{\|x - w_1\|} \le 0, \quad \text{i.e., } x_1^* \in \widehat{N}(w_1; \Lambda_1).$$

Thus in both possible cases in (8.139) we obtain $x_1^* \in \widehat{N}(w_1; \Lambda_1)$ and so $x_1^* \in N(0; \Lambda_1)$ by Lemma 8.71. Summarizing the above relationships yields

$$x_i^* \in N(0; \Lambda_i) \;\text{ and }\; \lambda_i = 0 \;\text{ for all }\; i \in \mathbb{N}.$$

Hence it follows from (8.137) that there are $\widetilde{x}_i^* := (1/2^i)x_i^* \in N(0; \Lambda_i)$ as $i \in \mathbb{N}$, not equal to zero simultaneously, satisfying

$$\sum_{i=1}^{\infty} \widetilde{x}_i^* = 0.$$

This contradicts the normal qualification condition (8.131) and thus shows that the case of $\lambda_1 = 0$ is actually *not possible* in (8.139).

Case 2: $\lambda_1 > 0$. If inequality (8.139) is strict, put $x = w_1$ in (8.138) and deduce from there that $\lambda_1 = 0$, a contradiction. Hence it remains to consider the case where (8.139) holds as equality. To proceed, take $(x, \alpha) \in O_1$ satisfying

$$x \in \Lambda_1 \setminus \{w_1\} \;\text{ and }\; \alpha = \langle x^*, x\rangle - \varepsilon\|x\|.$$

By the equality in (8.139) we have

$$\alpha - \alpha_1 = \langle x^*, x - w_1\rangle + \varepsilon(\|w_1\| - \|x\|) \;\text{ and thus }\; |\alpha - \alpha_1| \le (\|x^*\| + \varepsilon)\|x - w_1\|.$$

On the other hand, it follows from (8.138) that for any $\gamma > 0$ sufficiently small there exists a neighborhood V of w_1 such that

$$\langle x_1^*, x - w_1\rangle + \lambda_1(\alpha - \alpha_1) \le \lambda_1\gamma\varepsilon\big(\|x - w_1\| + |\alpha - \alpha_1|\big)$$

whenever $x \in \Lambda_1 \cap V$. Substituting there (x, α) with $x \in \Lambda_1 \cap V$ gives us

$$\begin{aligned}\langle x_1^*, x - w_1\rangle + \lambda_1(\alpha - \alpha_1) &= \langle x_1^* + \lambda_1 x^*, x - w_1\rangle + \lambda_1\varepsilon(\|w_1\| - \|x\|)\\ &\le \lambda_1\gamma\varepsilon(\|x - w_1\| + |\alpha - \alpha_1|)\\ &\le \lambda_1\gamma\varepsilon\big[\|x - w_1\| + (\|x^*\| + \varepsilon)\|x - w_1\|\big]\\ &= \lambda_1\gamma\varepsilon\big(1 + \|x^*\| + \varepsilon\big)\|x - w_1\|.\end{aligned}$$

It follows from the above that for small $\gamma > 0$ we have

$$\langle x_1^* + \lambda_1 x^*, x - w_1\rangle + \lambda_1\varepsilon(\|w_1\| - \|x\|) \le \lambda_1\varepsilon\|x - w_1\|$$

and thus arrive at the estimates

$$\langle x_1^* + \lambda_1 x^*, x - w_1\rangle \le \lambda_1\varepsilon\|x - w_1\| + \lambda_1\varepsilon(\|x\| - \|w_1\|) \le 2\lambda_1\varepsilon\|x - w_1\|$$

for all $x \in \Lambda_1 \cap V$. The latter implies by definition (1.6) of ε-normals that

$$x_1^* + \lambda_1 x^* \in \widehat{N}_{2\lambda_1\varepsilon}(w_1; \Lambda_1).$$

Furthermore, it is easy to observe from the above choice of λ_1 and the structure of O_1 in (8.135) that $\lambda_1 \le 2 + 2\varepsilon$. Employing now the representation of ε-normals from Exercise 1.42(i), we find $v \in \Lambda_1 \cap (w_1 + 2\lambda_1\varepsilon\mathbb{B})$ such that

$$x_1^* + \lambda_1 x^* \in \widehat{N}(v; \Lambda_1) + 2\lambda_1\varepsilon\mathbb{B}^* \subset N(0; \Lambda_1) + 2\lambda_1\varepsilon\mathbb{B}^*. \tag{8.140}$$

Since $\lambda_1 > 0$ and $-x_1^* = 2\sum_{i=2}^{\infty}\frac{1}{2^i}x_i^*$ by (8.137), it follows from (8.140) that

$$x^* \in N(0; \Lambda_1) + \frac{2}{\lambda_1}\sum_{i=2}^{\infty}\frac{1}{2^i}x_i^* + 2\varepsilon\mathbb{B}^*.$$

Hence there exists $\widetilde{x}_1^* \in N(0; \Lambda_1)$ such that

$$x^* \in \sum_{i=1}^{\infty}\frac{1}{2^i}\widetilde{x}_i^* + 2\varepsilon\mathbb{B}^* \ \text{ with } \ \widetilde{x}_i^* := \frac{2x_i^*}{\lambda_1} \in N(0; \Lambda_i) \ \text{ for } \ i = 2, 3, \ldots.$$

This justifies (8.133) and thus completes the proof of the theorem. △

Our next result presents an additional assumption under which we can put $\varepsilon = 0$ in (8.133) and hence get an exact representation of the normal x^*.

Theorem 8.73 (Exact Representation of Regular Normals to Countable Intersections of Cones). *Let $\{\Lambda_i\}_{i\in\mathbb{N}}$ be a countable system of cones in X under the validity of the normal qualification condition (8.131) at the origin. Then for any regular normal $x^* \in \widehat{N}\big(0; \bigcap_{i=1}^{\infty}\Lambda_i\big)$ satisfying*

$$\langle x^*, x\rangle < 0 \ \textit{whenever} \ x \in \bigcap_{i=1}^{\infty} \Lambda_i \setminus \{0\} \tag{8.141}$$

there are basic normals $x_i^* \in N(0; \Lambda_i)$, $i = 1, 2, \ldots$, *such that*

$$x^* = \sum_{i=1}^{\infty} \frac{1}{2^i} x_i^*. \tag{8.142}$$

Proof. Fix $x^* \in \widehat{N}\big(0; \bigcap_{i=1}^{\infty} \Lambda_i\big)$ satisfying condition (8.141) and construct a countable system of closed cones in $X \times \mathbb{R}$ by

$$O_1 := \big\{(x, \alpha)\big|\ x \in \Lambda_1,\ \alpha \le \langle x^*, x\rangle\big\},\ \ O_i := \Lambda_i \times \mathbb{R}_+ \text{ for } i \ge 2. \tag{8.143}$$

Similarly to the proof of Theorem 8.72 with taking (8.141) into account, we can verify that all the assumptions of Theorem 2.9 hold. Applying the conic extremal principle established therein gives us pairs $(w_i, \alpha_i) \in O_i$ and $(x_i^*, \lambda_i) \in \widehat{N}\big((w_i, \alpha_i); O_i\big)$ such that the extremality conditions in (8.137) are satisfied. We obviously get $\lambda_i \le 0$ and $x_i^* \in \widehat{N}(w_i; \Lambda_i)$ for $i = 1, 2, \ldots$, which ensures that $x_i^* \in N(0; \Lambda_i)$ as $i \ge 2$ by Lemma 8.71. It follows furthermore that for $i = 1$ the limiting inequality (8.138) holds. The latter implies by the structure of O_1 in (8.143) that

$$\lambda_1 \ge 0 \ \text{ and } \ \alpha_1 \le \langle x^*, w_1\rangle. \tag{8.144}$$

Similarly to the proof of Theorem 8.72, we consider the two possible cases $\lambda_1 = 0$ and $\lambda_1 > 0$ in (8.144) and show that the first case contradicts (8.131). In the second case, we arrive at representation (8.142) based on the extremality conditions in (8.137) and the structures of the sets O_i in (8.143). △

The final goal in this subsection is obtaining a constructive upper estimate of the regular normal cone to countable intersections of arbitrary closed sets via basic normals to each of these sets. To proceed in this direction, we first consider countable systems of closed cones in X.

Lemma 8.74 (Upper Estimate of the Regular Normal Cone to Countable Cone Intersections). *Let* $\{\Lambda_i\}_{i\in\mathbb{N}}$ *be a countable system of cones in* X *satisfying the normal qualification condition* (8.131) *at the origin. Then*

$$\widehat{N}\Big(0; \bigcap_{i=1}^{\infty} \Lambda_i\Big) \subset \text{cl}\,\Big\{\sum_{i\in I} x_i^* \Big|\ x_i^* \in N(0; \Lambda_i),\ \ I \in \mathcal{L}\Big\}. \tag{8.145}$$

Proof. To verify (8.145), pick $x^* \in \widehat{N}(0; \Lambda)$ and for any fixed $\varepsilon > 0$ apply Theorem 8.72. In this way we find $x_i^* \in N(0; \Lambda_i)$, $i \in \mathbb{N}$, satisfying (8.133). Since $\varepsilon > 0$ was chosen arbitrarily, it follows that

$$x^* \in A := \text{cl}\,\left\{\sum_{i=1}^{\infty} \frac{1}{2^i} x_i^* \right|\left. x_i^* \in N(0; \Lambda_i)\right\}.$$

It remains to justify the inclusion

$$A \subset \mathrm{cl}\, C \quad \text{with} \quad C := \Big\{ \sum_{i \in I} x_i^* \Big| \ x_i^* \in N(0; \Lambda_i), \ I \in \mathcal{L} \Big\}.$$

To proceed, pick $z^* \in A$ and for any fixed $\varepsilon > 0$ find $x_i^* \in N(0; \Lambda_i)$ satisfying

$$\Big\| z^* - \sum_{i=1}^{\infty} \frac{1}{2^i} x_i^* \Big\| \le \frac{\varepsilon}{2}. \tag{8.146}$$

Then choose $k \in \mathbb{N}$ so large that (8.146) holds with replacing the series $\sum_{i=1}^{\infty}$ by the sum $\sum_{i=1}^{k}$ therein. The latter sum clearly belongs to C, and hence $(z^* + \varepsilon \mathbb{B}^*) \cap C \neq \emptyset$, which yields $z^* \in \mathrm{cl}\, C$ and thus justifies (8.145). △

Now we are ready to derive the aforementioned upper estimate of the regular normal cone to countable intersections of arbitrary closed sets, which is important for applications to necessary optimality conditions for non-Lipschitzian SIPs in the next subsection.

Theorem 8.75 (Upper Estimates of the Regular Normal Cone to Countable Set Intersections). *Let $\{\Omega_i\}_{i \in \mathbb{N}}$ be a countable system of nonempty sets in X, and let $\bar{x} \in \Omega := \bigcap_{i=1}^{\infty} \Omega_i$. Assume that CHIP in Definition 8.60 and NQC in (8.131) are satisfied for $\{\Omega_i\}_{i \in \mathbb{N}}$ at $\bar{x}$. Then we have the inclusion*

$$\widehat{N}(\bar{x}; \Omega) \subset \mathrm{cl} \Big\{ \sum_{i \in I} x_i^* \Big| \ x_i^* \in N(\bar{x}; \Omega_i), \ I \in \mathcal{L} \Big\}. \tag{8.147}$$

If in addition NCC (8.130) holds for $\{\Omega_i\}_{i \in \mathbb{N}}$ at $\bar{x}$, then the closure operation can be omitted on the right-hand side of (8.147).

Proof. Using the definitions of the regular normal and contingent cones and then the assumed CHIP gives us, respectively, the following two equalities:

$$\widehat{N}(\bar{x}; \Omega) = \widehat{N}\big(0; T(\bar{x}; \Omega)\big) = \widehat{N}\Big(0; \bigcap_{i=1}^{\infty} T(\bar{x}; \Omega_i)\Big).$$

By passing to basic normals, we have the inclusions (see Exercise 1.49)

$$N\big(0; T(\bar{x}; \Omega_i)\big) \subset N(\bar{x}; \Omega_i) \quad \text{for all} \quad i \in \mathbb{N}.$$

This shows that the imposed NQC for $\{\Omega_i\}_{i \in \mathbb{N}}$ at $\bar{x}$ ensures it for the cones $\{T(\bar{x}; \Omega_i)\}_{i \in \mathbb{N}}$ at the origin. Applying Lemma 8.74 to the latter system yields

$$\widehat{N}\Big(0; \bigcap_{i=1}^{\infty} T(\bar{x}; \Omega_i)\Big) \subset \mathrm{cl} \Big\{ \sum_{i \in I} x_i^* \Big| \ x_i^* \in N\big(0; T(\bar{x}; \Omega_i)\big), \ I \in \mathcal{L} \Big\}.$$

Combining the above, we arrive at (8.147) and can obviously drop the closure operation therein under the additional NCC assumption. △

8.4.3 Optimality Conditions Under Countable Constraints

This subsection is devoted to deriving necessary optimality conditions of diverse types for nonsmooth SIPs with *countable* constraints. Although such problems naturally arise in various applications (in particular, to control systems on the infinite horizon and dynamical models of macroeconomics), they are much less investigated in comparison with SIPs indexed by compact sets, which offer more possibilities to implement a variety of mathematical techniques. The absence of the index set compactness, as in the case of countable constraints, creates significant mathematical difficulties, which are comparable for SIPs with countable index sets and those with arbitrary index sets considered above. In contrast to the previous material that concerns SIPs with smooth, convex, and Lipschitzian data, now we are able to deal with general geometric and nonsmooth inequality constraints. In this way, mainly based on the extremal principle for countable set systems, we establish verifiable conditions in broader frameworks that are independent from the previous ones in their special settings (even for linear system) as illustrated by examples.

Let us start with SIPs involving countable *geometric* constraints:

$$\text{minimize } \varphi(x) \text{ subject to } x \in \Omega_i \text{ as } i \in \mathbb{N}, \tag{8.148}$$

where $\varphi\colon X \to \overline{\mathbb{R}}$ is an extended-real-valued cost function, and where $\{\Omega_i\}_{i\in\mathbb{N}}$ is a countable system of sets. Following the general scheme of deriving necessary conditions in nonsmooth optimization presented in Section 6.1 and using the calculus rules for countable set intersections developed in the preceding subsection, we obtain *qualified* necessary optimality conditions for SIP (8.148) of both *upper subdifferential* and *lower subdifferential* types.

Theorem 8.76 (Upper Subdifferential Conditions for SIPs with Countable Geometric Constraints). *Let $\bar{x}$ be a local optimal solution to problem* (8.148)*, where $\varphi\colon X \to \overline{\mathbb{R}}$ is finite at $\bar{x} \in \Omega_i$ as $i \in \mathbb{N}$. Assume that the system $\{\Omega_i\}_{i\in\mathbb{N}}$ enjoys CHIP from Definition 8.60 at $\bar{x}$ and satisfies NQC* (8.131) *at this point. Then we have the set inclusion*

$$-\widehat{\partial}^{+}\varphi(\bar{x}) \subset \text{cl}\left\{\sum_{i\in I} x_i^*\,\middle|\, x_i^* \in N(\bar{x};\Omega_i),\ I \in \mathcal{L}\right\}, \tag{8.149}$$

which reduces to the simplified condition

$$0 \in \nabla\varphi(\bar{x}) + \text{cl}\left\{\sum_{i\in I} x_i^*\,\middle|\, x_i^* \in N(\bar{x};\Omega_i),\ I \in \mathcal{L}\right\} \tag{8.150}$$

if φ is Fréchet differentiable at $\bar{x}$. If further NCC (8.130) holds for $\{\Omega_i\}_{i\in\mathbb{N}}$ at $\bar{x}$, then the closure operations can be dropped in (8.149) and (8.150).

Proof. It follows from Theorem 6.1(i) that

$$-\widehat{\partial}^+\varphi(\bar{x}) \subset \widehat{N}\Big(\bar{x}; \bigcap_{i=1}^{\infty}\Omega_i\Big).$$

Applying now the upper estimate of $\widehat{N}\big(\bar{x}; \bigcap_{i=1}^{\infty}\Omega_i\big)$ from Theorem 8.75 under the assumed CHIP and NQC, we arrive at (8.149), where the closure operation can be omitted when NCC holds at $\bar{x}$. If φ is Fréchet differentiable at $\bar{x}$, it follows that $\widehat{\partial}^+\varphi(\bar{x}) = \{\nabla\varphi(\bar{x})\}$, and thus (8.149) reduces to (8.150). △

Theorem 8.77 (Lower Subdifferential Conditions for SIPs with Countable Geometric Constraints). *Let $\bar{x}$ be a local optimal solution to SIP (8.148), where the set $\Omega := \bigcap_{i=1}^{\infty}\Omega_i$ is normally regular at $\bar{x}$, where the system $\{\Omega_i\}_{i\in\mathbb{N}}$ enjoys CHIP and NQC at $\bar{x}$, and where*

$$\mathrm{cl}\left\{\sum_{i\in I}x_i^*\,\Big|\, x_i^*\in N(\bar{x};\Omega_i),\ I\in\mathcal{L}\right\}\bigcap\big(-\partial^{\infty}\varphi(\bar{x})\big)=\{0\}; \tag{8.151}$$

the latter is satisfied when φ is locally Lipschitzian around $\bar{x}$. Then

$$0\in\partial\varphi(\bar{x})+\mathrm{cl}\left\{\sum_{i\in I}x_i^*\,\Big|\, x_i^*\in N(\bar{x};\Omega_i),\ I\in\mathcal{L}\right\}. \tag{8.152}$$

The closure operations can be dropped in (8.151), (8.152) if NCC holds at $\bar{x}$.

Proof. It follows from Theorem 6.1(ii) that

$$0\in\partial\varphi(\bar{x})+N(\bar{x};\Omega)\ \text{ if }\ \partial^{\infty}\varphi(\bar{x})\cap\big(-N(\bar{x};\Omega)\big)=\{0\} \tag{8.153}$$

for the optimal solution $\bar{x}$ to (8.148) with the only geometric constraint written as $x\in\Omega=\bigcap_{i=1}^{\infty}\Omega_i$. Since Ω is normally regular at $\bar{x}$, we can replace $N(\bar{x};\Omega)$ by $\widehat{N}(\bar{x};\Omega)$ in (8.153). Applying now Theorem 8.75 to the countable set intersection Ω in (8.153), we arrive at all the claimed conclusions. △

Remark 8.78 (Passing to SIPs with Structural Constraints). Since the sets Ω_i are arbitrary (closed) in Theorems 8.76 and 8.77, we may consider various constraint settings, where Ω_i are given in some structural forms via, e.g., operator, functional, and other types of constraints. Proceeding from the general geometric results to their implementations for particular situations requires appropriate *calculus rules* for the generalized differential constructions therein while ensuring the *preservation* of the properties and qualification conditions used in Theorems 8.76 and 8.77. No troubles arise in this direction for the conditions and properties involved in Theorem 8.76 and 8.77 that are expressed in terms of our basic normal cone, which enjoys full calculus. More challenging and underinvestigated issues relate to the CHIP preservation due to limited calculus available for the contingent cone to nonconvex sets. Exercise 8.118 describes a particular constraint setting, where such an implementation completely goes through.

Next let us consider SIPs with *countable inequality constraints* given by:

$$\text{minimize } \varphi(x) \text{ subject to } \varphi_i(x) \le 0,\ i \in \mathbb{N}, \tag{8.154}$$

where the cost function φ is the same as above while φ_i are supposed to be merely l.s.c. around the reference local solution for (8.154) to ensure the local closedness of the sets $\Omega_i := \operatorname{epi} \varphi_i$ in Theorems 8.76 and 8.77, which is our standing assumption. We know that the normal cone to each set $\operatorname{epi} \varphi_i$ is fully described by collections of both basic and singular subgradients of φ_i allowing us therefore to transform the normal cone conditions of Theorems 8.76 and 8.77 into the basic and singular subdifferential conditions for problem (8.154). Just to simplify the expressions obtained in this way, we suppose that φ_i are locally Lipschitzian and thus exclude the singular subgradients of φ_i from the consideration in this case. The corresponding constraint qualifications of Definition 8.69 reduce now to the following ones.

Definition 8.79 (Subdifferential Closedness and Qualification Conditions for Countable Inequality Constraints). *Consider the constraint sets*

$$\Omega_i := \big\{x \in X \big|\ \varphi_i(x) \le 0\big\}, \quad i \in \mathbb{N}, \tag{8.155}$$

where φ_i are locally Lipschitzian around $\bar{x} \in \bigcap_{i=1}^{\infty} \Omega_i$. We say that:

(a) *The system $\{\Omega_i\}_{i\in\mathbb{N}}$ in (8.155) satisfies the* SUBDIFFERENTIAL CLOSEDNESS CONDITION (*SCC*) *at $\bar{x}$ if the set*

$$\Big\{\sum_{i\in I} \lambda_i \partial\varphi_i(\bar{x}) \Big|\ \lambda_i \ge 0, \lambda_i\varphi_i(\bar{x}) = 0,\ I \in \mathcal{L}\Big\} \ \textit{is closed in}\ \ X^*. \tag{8.156}$$

(b) *The system $\{\Omega_i\}_{i\in\mathbb{N}}$ in (8.155) satisfies the* SUBDIFFERENTIAL QUALIFICATION CONDITION (*SQC*) *at $\bar{x}$ if the trivial multiplier collection $\lambda_i = 0$ for all $i \in \mathbb{N}$ is the only one for which we have the relationships*

$$\sum_{i=1}^{\infty} \lambda_i x_i^* = 0,\ x_i^* \in \partial\varphi_i(\bar{x}),\ \lambda_i \ge 0,\ \lambda_i\varphi_i(\bar{x}) = 0.$$

Using the constraint qualifications from Definition 8.79 and subdifferential calculus implies the following consequences of Theorems 8.76 and 8.77.

Corollary 8.80 (Upper and Lower Subdifferential Conditions for SIPs with Countable Inequality Constraints). *Let $\bar{x}$ be a local optimal solution to* (8.154)*, where the constraint functions φ_i, $i \in \mathbb{N}$, are locally Lipschitzian around $\bar{x}$. Assume that the constraint set system $\{\Omega_i\}_{i\in\mathbb{N}}$ in* (8.155) *has CHIP at $\bar{x}$ and that SQC from Definition* 8.79 *is satisfied at this point. Then we have:*

(i) *The upper subdifferential optimality condition*

$$-\widehat{\partial}^+\varphi(\bar{x}) \subset \mathrm{cl}\Big\{\sum_{i\in I}\lambda_i\partial\varphi_i(\bar{x})\Big|\ \lambda_i\ge 0,\ \lambda_i\varphi_i(\bar{x})=0,\ I\in\mathcal{L}\Big\}, \tag{8.157}$$

where the closure operation can be omitted if SCC (8.156) *is satisfied at* $\bar{x}$.

(ii) *Let in addition the feasible set in* (8.154) *be normally regular at* $\bar{x}$ *and*

$$\mathrm{cl}\Big\{\sum_{i\in I}\lambda_i\partial\varphi_i(\bar{x})\Big|\ \lambda_i\ge 0, \lambda_i\varphi_i(\bar{x})=0, I\in\mathcal{L}\Big\}\bigcap\big(-\partial^\infty\varphi(\bar{x})\big)=\{0\},$$

which is automatic if φ *is locally Lipschitzian around* $\bar{x}$. *Then*

$$0\in\partial\varphi(\bar{x})+\mathrm{cl}\Big\{\sum_{i\in I}\lambda_i\partial\varphi_i(\bar{x})\Big|\ \lambda_i\ge 0,\ \lambda_i\varphi_i(\bar{x})=0,\ I\in\mathcal{L}\Big\} \tag{8.158}$$

with removing the closure operations above when SCC holds at $\bar{x}$.

Proof. It follows from Exercise 2.51(i) that

$$N(\bar{x};\Omega)\subset\mathbb{R}_+\partial\vartheta(\bar{x})\ \text{ for }\ \Omega:=\big\{x\in X\big|\ \vartheta(x)\le 0\big\} \tag{8.159}$$

provided that ϑ is locally Lipschitzian around $\bar{x}$ and that $0\notin\partial\vartheta(\bar{x})$, which is ensured by the assumed SQC. Now we apply (8.159) to each set Ω_i in (8.155) and substitute this into NQC (8.131) as well as into the qualification condition (8.151) and the optimality conditions (8.149) and (8.152) for problem (8.148) with the constraint sets (8.155). It follows in this way that SQC and the assumptions in (ii) imply the aforementioned conditions of Theorems 8.76 and 8.77 in the setting (8.154) under consideration. Furthermore, SCC clearly yields NCC (8.130) for the sets in (8.155), which completes the proof. △

Now we consider the case of *convex* constraint functions φ_i in (8.154). Note that the validity of SQC is ensured in this case by the interior-type condition (8.132) of Proposition 8.70. The next result establishes necessary optimality conditions for problems with countable convex inequality constraints, which don't require either interiority-type or SQC assumptions while containing the following qualification condition that implies both CHIP and SCC. Recall that the symbol "cone" stands here for the *convex* conic hull of a set.

Definition 8.81 (Local Farkas-Minkowski Property). *We say that the countable system of convex inequalities* (8.155) *satisfies the* LOCAL FARKAS-MINKOWSKI (LFM) *property at* $\bar{x}\in\Omega:=\cap_{i=1}^\infty\Omega_i$ *if*

$$N(\bar{x};\Omega)=\mathrm{cone}\bigcup_{i\in J(\bar{x})}\partial\varphi_i(\bar{x})=:A(\bar{x}), \tag{8.160}$$

where $J(\bar{x}):=\{i\in\mathbb{N}|\ \varphi_i(\bar{x})=0\}$ *is the the collection of active indices at* $\bar{x}$.

The LFM terminology is supported by the fact that FMCQ⇒LFM for the *Farkas-Minkowski* property (constraint qualification) defined for convex inequality systems in (7.159); see Exercise 8.120. Having this in hand, we get the following results for infinite convex inequality systems, where it is assumed for simplicity that the cost function in (8.154) is locally Lipschitzian.

Proposition 8.82 (Upper and Lower Subdifferential Conditions for SIP with Convex Inequality Constraints). *Let all the general assumptions but SQC of Corollary 8.80 be fulfilled at the local optimal solution $\bar{x}$ to (8.154). Suppose in addition that the cost function φ is locally Lipschitzian around $\bar{x}$, that the constraint functions φ_i, $i \in \mathbb{N}$, are convex, and that the LFM property (8.160) holds at $\bar{x}$. Then both SCC and CHIP also hold for this system, and the necessary optimality conditions (8.157) and (8.158) are satisfied without the closure operation therein.*

Proof. Observe that SCC (8.156) is nothing else but the closedness of the set $A(\bar{x})$, and hence we get LFM⇒SCC by the closedness of the normal cone $N(\bar{x};\Omega)$. Furthermore, we always have the inclusions

$$A(\bar{x}) \subset \operatorname{co} \bigcup_{i \in J(\bar{x})} N(\bar{x};\Omega_i) \subset N(\bar{x};\Omega).$$

Hence the LFM property combined with the latter inclusions yields the strong CHIP. By Theorem 8.61 we have CHIP as well due to $N(\bar{x};\Omega_i) = \{0\}$ whenever $i \notin J(\bar{x})$. Taking all this into account gives us the relationships

$$-\widehat{\partial}^{+}\varphi(\bar{x}) \subset N(\bar{x};\Omega) \ \text{ and } \ 0 \in \partial\varphi(\bar{x}) + N(\bar{x};\Omega),$$

which imply in turn the validity of the inclusions

$$-\widehat{\partial}^{+}\varphi(\bar{x}) \subset A(\bar{x}) \ \text{ and } \ 0 \in \partial\varphi(\bar{x}) + A(\bar{x})$$

and thus complete the proof of the proposition. △

The next specification of Proposition 8.82 for SIPs with linear inequality constraints agrees with the corresponding results of Section 7.2 *without* imposing the strong Slater condition and the coefficient boundedness.

Corollary 8.83 (Upper and Lower Subdifferential Conditions for SIPs with Linear Inequality Constraints). *Let $\bar{x} = 0$ locally solve the SIP:*

$$\textit{minimize} \ \ \varphi(x) \ \ \textit{subject to} \ \ \langle a_i, x\rangle \le 0 \ \ \textit{for all} \ \ i \in \mathbb{N},$$

where $\varphi\colon X \to \overline{\mathbb{R}}$ is finite at the origin. Then we have the inclusions

$$-\widehat{\partial}^{+}\varphi(0) \subset \operatorname{cl}\operatorname{co}\Big[\bigcup_{i=1}^{\infty}\big\{\lambda a_i \,\big|\, \lambda \ge 0\big\}\Big].$$

$$0 \in \partial\varphi(0) + \text{cl}\,\text{co}\Big[\bigcup_{i=1}^{\infty}\big\{\lambda a_i \big|\, \lambda \geq 0\big\}\Big],$$

where the latter one holds provided that

$$\Big(\text{cl}\,\text{co}\Big[\bigcup_{i=1}^{\infty}\big\{\lambda a_i \big|\, \lambda \geq 0\big\}\Big]\Big) \cap \big(-\partial^{\infty}\varphi(0)\big) = \{0\}.$$

Furthermore, the LFM property implies that the closure operations can be omitted in all the conditions above.

Proof. It follows from Proposition 8.82 by the normal cone representation for solutions to linear inequality systems given in Corollary 8.68. △

Finally in this section, we present several examples illustrating the qualification conditions imposed in Proposition 8.82 and their comparison with the corresponding results of Chapter 7 for fully convex SIPs.

Example 8.84 (Comparison of Qualification Conditions). All the examples below concern lower subdifferential conditions for *fully convex* SIPs (8.154), i.e., those with convex cost and constraint functions.

(i) *CHIP* (8.115) *and SCC* (8.156) *are independent*. Consider a linear constraint system in (8.155) at $\bar{x} = (0,0) \in \mathbb{R}^2$ for $\varphi_i(x) = \langle a_i, x\rangle$ with $a_i = (1,i)$ as $i \in \mathbb{N}$, which clearly enjoys CHIP at the origin. On the other hand, the set

$$\text{co}\bigcup_{i=0}^{\infty}\mathbb{R}_+\partial\varphi_i(\bar{x}) = \text{co}\big\{\lambda(1,i) \in \mathbb{R}^2 \big|\, \lambda \geq 0,\ i \in \mathbb{N}\big\} = \mathbb{R}^2_+ \setminus \big\{(0,\lambda) \big|\, \lambda > 0\big\}$$

is not closed; see Fig. 8.3; and hence SCC doesn't hold in this setting. If we consider now the quadratic inequality constraint functions $\varphi_i(x) = ix_1^2 - x_2$ with $x = (x_1, x_2) \in \mathbb{R}^2$ and $i \in \mathbb{N}$, then $\partial\varphi_i(\bar{x}) = \nabla\varphi_i(\bar{x}) = (0,-1)$, and hence SCC is satisfied at the origin. However, it is easy to check by the direct calculation that CHIP fails at $\bar{x}$ for the latter constraint system.

(ii) *CHIP and SCC vs. FMCQ and CQC.* Besides FMCQ (7.159), in Section 7.5 we study and apply to deriving necessary optimality conditions for fully convex SIPs yet another property, which is called the *closedness qualification condition* (CQC) and is formulated in the case of SIPs (8.154) with convex data in $X = \mathbb{R}^n$ as follows: the set

$$\text{epi}\,\varphi^* + \text{cone}\bigcup_{i=1}^{\infty}\text{epi}\,\varphi_i^*$$

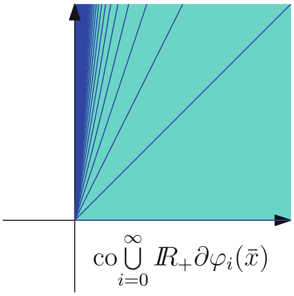


Fig. 8.3 Failure of SCC.

defined via the conjugate functions is closed in $\mathbb{R}^{n+1}$; see Definition 7.48. This property is strictly weaker (better) than FMCQ if, in particular, the cost function φ is continuous at $\bar{x}$; see Exercise 7.97(i).

The next system reveals more: it presents a fully convex SIP satisfying both CHIP and SCC but not CQC, and hence not FMCQ. This shows that Proposition 8.82 holds in this case to produce the KKT optimality condition while the corresponding results of Subsection 7.5.1 are not applicable.

Consider SIP (8.154) in $\mathbb{R}^2$ with $\bar{x} = (0,0)$, $\varphi(x) := -x_2$, and

$$\varphi_i(x_1,x_2) := \begin{cases} ix_1^3 - x_2 & \text{if } x_1 < 0, \\ -x_2 & \text{if } x_1 \geq 0, \end{cases} \qquad i \in \mathbb{N}.$$

We have $\partial\varphi_i(\bar{x}) = \{\nabla\varphi_i(\bar{x})\} = (0,-1)$ for all $i \in \mathbb{N}$, and hence SCC holds. It is easy to check that CHIP also holds at $\bar{x}$, since

$$T\Big(\bar{x}; \bigcap_{i=1}^{\infty}\Omega_i\Big) = T(\bar{x};\Omega_i) = \mathbb{R}\times\mathbb{R}_+ \ \text{ for } \ \Omega_i := \big\{x \in \mathbb{R}^2 \big|\ \varphi_i(x) \leq 0\big\}, \quad i \in \mathbb{N}.$$

On the other hand, for $(\lambda_1,\lambda_2) \in \mathbb{R}^2$ we calculate

$$\varphi^*(\lambda_1,\lambda_2) = \begin{cases} 0 & \text{if } (\lambda_1,\lambda_2) = (0,-1), \\ \infty & \text{otherwise} \end{cases} \quad \text{and } \varphi_i^*(\lambda_1,\lambda_2) = \begin{cases} 0 & \text{if } \lambda_1 \leq 0, \lambda_2 = -1, \\ \infty & \text{otherwise.} \end{cases}$$

This shows that the convex sets

$$\text{cone}\bigcup_{i=0}^{\infty}\text{epi}\,\varphi_i^* \quad \text{and} \quad \text{epi}\,\varphi^* + \text{cone}\bigcup_{i=0}^{\infty}\text{epi}\,\varphi_i^*$$

are not closed in $\mathbb{R}^3$, and hence both FMCQ and CQC are violated.

(iii) *SQC doesn't imply CHIP for countable systems.* Comparing the results of Corollary 8.80 and Proposition 8.82 for convex SIPs, a natural question arises: whether SQC always yields CHIP in the case of constraint sets Ω_i in (8.155) defined by $\mathcal{C}^1$-smooth convex functions. The following simple example shows that it is not the case for countably many constraints in $\mathbb{R}^2$. Indeed, we obviously have the validity of SQC at $\bar{x} = (0, 0)$ for the family of functions

$$\varphi_i(x_1, x_2) := ix_1^2 - x_2 \text{ with } x = (x_1, x_2) \in \mathbb{R}^2,\ i \in \mathbb{N},$$

while it is easy to check by the direct calculation that CHIP fails at the origin for the countable system of sets $\Omega_i := \{x \in \mathbb{R}^2 |\ \varphi_i(x_1, x_2) \leq 0\}$, $i \in \mathbb{N}$.

8.5 Exercises for Chapter 8

Exercise 8.85 (Uniform Differentiability and EMFCQ Assumptions Over Noncompact Index Sets). Give examples of infinite systems (8.2) with noncompact index sets in both cases of finite-dimensional and infinite-dimensional spaces X for which all the assumptions in (SA), (8.3), and (8.4) as well as the EMFCQ property from Definition 8.2 are satisfied.

Exercise 8.86 (Violation of KKT for SIPs with Compact Index Sets and Discontinuity with Respect to Index Variables). Give an example of SIP (8.1) in finite dimensions with $T = [0, 1]$ and a discontinuous mapping $(x, t) \mapsto \nabla\varphi_t(x)$ for which the KKT condition (8.7) fails.

Exercise 8.87 (Equivalent Form of NFMCQ). Show that NFMCQ from Definition 8.8 is equivalent to the requirement that the convex conic hull of the set $\{(\nabla\varphi_t(\bar{x}), \varphi_t(\bar{x}))|\ t \in T\}$ is weak* closed in $X^* \times \mathbb{R}$.

Exercise 8.88 (NFMCQ from MFCQ for Finite Inequality Systems). Give a detailed proof of Proposition 8.9(i).

Exercise 8.89 (Closedness of Conic Convex Hulls in Finite Dimensions). Let $\emptyset \neq S \subset \mathbb{R}^n$ be a compact set such that $0 \notin \operatorname{co} S$. Show that cone S is closed in $\mathbb{R}^n$.

Exercise 8.90 (Regular Normal Cone Representation for Infinite Systems with Unbounded Gradients). Formulate and prove a counterpart of Corollary 8.15 for the regular normal cone to the constraint set Ω in (8.2).

Exercise 8.91 (Normal Cone Representation Under Equicontinuity of Gradients). Considering the infinite inequality system in (8.2), assume that the gradients $\nabla\varphi_t$ of the constraint functions are *equicontinuous* at $\bar{x}$ in the sense of [686]: for each $\gamma > 0$ there is $\eta > 0$ such that

$$\|\nabla\varphi_t(x) - \nabla\varphi_t(\bar{x})\| \leq \gamma \text{ for all } x \in B_\eta(\bar{x}),\ t \in T.$$

Hint: Show by using the mean value theorem that this assumption together with the Fréchet differentiability of φ_t around $\bar{x}$ implies condition (b) in Corollary 8.15.

Exercise 8.92 (Normals to Infinite Convex Sets and Farkas-Minkowski Conditions). Consider the infinite constraint set Ω from (8.2) with convex functions φ_t and $h \equiv 0$. As we see, representation (8.21) of the normal cone to Ω is the same as that obtained in Corollary 7.53 in this case.

(i) Show that assumptions imposed in Corollaries 7.53 and 8.19 are generally independent even in the case of $\dim X < \infty$ and the validity of (SA).

(ii) Find relationships between convex FMCQ (7.159) and NFMCQ from Definition 8.8 for convex infinite systems in both finite and infinite dimensions.

(iii) Modify the proof of Corollary 8.19 by using the arguments in the proof of Corollary 8.15 to avoid the boundedness requirement on $\{\nabla\varphi_t(\bar{x})|\ t \in T\}$ from the standing assumptions in (SA).

Exercise 8.93 (Normals to Infinite Linear Inequality Systems). Using the approach developed in the proof of Theorem 7.5 for infinite linear inequality systems, derive the results of Corollary 8.20 in the case of $h \equiv 0$ without the boundedness assumption on the coefficient set $\{a_t^*|\ t \in T\} \subset X^*$.

Exercise 8.94 (Upper Subdifferential Necessary Optimality Conditions for SIPs with Differentiable Constraints). Derive a general upper subdifferential counterpart of Proposition 8.21 in Banach spaces.

Exercise 8.95 (Comparison Between Necessary and Sufficient Optimality Conditions for Convex SIPs). Observe that the optimality conditions (7.168) for $\Theta = 0$ and (8.30) for $h \equiv 0$ agree while they are derived under different assumptions.

(i) Show that these results of Corollary 7.54 and Theorem 8.26 are generally independent. Are they the same if the space X is finite-dimensional?

(ii) Find a general formulation of necessary and sufficient conditions for convex SIPs that covers both results obtained in these statements.

Exercise 8.96 (Evaluations of Generalized Gradients for Supremum Functions). Let $\overline{\partial}\psi(\bar{x})$ be Clarke's generalized gradient (1.78) of the supremum function (8.32) on an Asplund space X.

(i) Using the representation of $\overline{\partial}\psi(x)$ from Exercise 4.36(i) and the results for $\partial\psi(\bar{x})$ obtained in Theorem 8.30 and its corollaries, evaluate $\overline{\partial}\psi(x)$ in the case of arbitrary index sets T. *Hint:* Compare it with [550, Theorem 4.1].

(ii) Establish specifications of (i) in the case of compact sets T in metrizable spaces and compare them with [165, Theorem 2.8.2] and [550, Corollary 4.2].

(iii) Prove the equality representation

$$\overline{\partial}\psi(\bar{x}) = \text{cl}^*\text{co}\Big[\bigcup\big\{\overline{\partial}\varphi_t(\bar{x})\big|\ t \in T(\bar{x})\big\}\Big]$$

provided that $T_\varepsilon(\bar{x})$ is a compact subset of a metrizable index set T for some $\varepsilon > 0$, that $t \mapsto \varphi_t(x)$ is u.s.c. on $T_\varepsilon(\bar{x})$ for each x sufficiently closed to $\bar{x}$, and that the functions $\varphi_t(\cdot)$, $t \in T_\varepsilon(\bar{x})$, are UNIFORMLY SUBSMOOTH at $\bar{x} \in X$ in the sense that whenever $\widetilde{\varepsilon} > 0$ there is $\delta > 0$ for which

$$\varphi_t(x) - \varphi_t(u) \geq \langle u^*, x - u\rangle - \widetilde{\varepsilon}\|x - u\| \ \text{ if } \ x, u \in B_\delta(\bar{x}),\ u^* \in \overline{\partial}\varphi_t(u).$$

Show furthermore that the representation can be replaced by

$$\overline{\partial}\psi(\bar{x}) = \text{co}\Big[\bigcup\big\{\overline{\partial}\varphi_t(\bar{x})\big|\ t \in T(\bar{x})\big\}\Big]$$

if $\dim X < \infty$; cf. [778, Theorems 3.1, 3.2] and [550, Corollaries 4.3, 4.4].

Exercise 8.97 (Generalized Gradients for Suprema of Equicontinuously Subdifferentiable Functions). Let the functions φ_t, $t \in T$, in (8.32) equicontinuously subdifferentiable at $\bar{x}$, where T is an arbitrary index set. Then we have the generalized gradient estimate

$$\overline{\partial}\psi(\bar{x}) \subset \bigcap_{\varepsilon>0} \text{cl}^*\text{co}\Big[\bigcup\big\{\overline{\partial}\varphi_t(\bar{x})\big|\ t \in T_\varepsilon(\bar{x})\big\}\Big].$$

Hint: Argue as in the proof of Corollary 8.34 and compare with [550, Proposition 4.5].

Exercise 8.98 (Relationships Between Qualification and Optimality Conditions for Lipschitzian SIPs). Consider the Lipschitzian SIP (8.31) under the equicontinuous subdifferentiability of φ_t at a feasible point $\bar{x}$.

(i) Find relationships between the generalized PMFCQ and the constraint qualification of Corollary 8.38 in finite and infinite dimensions.

(ii) How do the corresponding statements of Corollary 8.38 and Theorem 8.40 relate to each other when $\bar{x}$ is a local minimizer of (8.31)?

Exercise 8.99 (Supremum Marginal Functions). Consider the class of *supremum marginal functions* $\vartheta\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ defined by

$$\vartheta(x) := \sup_{y\in G(x)} \varphi(x, y), \quad x \in \mathbb{R}^n, \tag{8.161}$$

where the sets $G(x) \subset \mathbb{R}^m$ are nonempty, uniformly bounded around some point $\bar{x} \in \mathbb{R}^n$, and admit the representation

$$G(x) = \big\{y \big|\ \varphi_i(x, y) \le 0, i = 1, \ldots, r;\ \varphi_i(x, y) = 0,\ i = r+1, \ldots, s\big\} \tag{8.162}$$

via continuous real-valued functions φ_i on $\mathbb{R}^n \times \mathbb{R}^m$ with $i = 1, \ldots, s$ and $r \le s$.

(i) Verify that $\vartheta(x)$ is upper semicontinuous around $\bar{x}$.

(ii) Define the *argmaximum sets*

$$S(x) := \big\{y \in G(x) \big|\ \varphi(x, y) = \vartheta(x)\big\}, \quad x \in \mathbb{R}^n, \tag{8.163}$$

and show that $\vartheta(x)$ is continuous at $\bar{x}$ if $\mathrm{dist}(S(x); G(x)) \to 0$ as $x \to \bar{x}$. It holds, in particular, when $G(x)$ from (8.162) is inner semicontinuous at $\bar{x}$.

Hint: Proceed by the definitions and compare it with [466, Propositions 3.1, 3.2].

Exercise 8.100 (Marginal Mangasarian-Fromovitz Constraint Qualification). Given $\bar{x} \in \mathbb{R}^n$, $G(\bar{x})$ from (8.162), and $\Omega \subset G(\bar{x})$, assume that φ_i are strictly differentiable at $(\bar{x}, y)$ for all $y \in \Omega$ and $i = 1, \ldots, s$. Then we say that the MARGINAL MANGASARIAN-FROMOVITZ CONSTRAINT QUALIFICATION (MMFCQ) holds at $\bar{x}$ relative to Ω if there is a vector $\xi \in \mathbb{R}^n$ such that

$$\Big\langle \sum_{i=1}^{s} \lambda_i \nabla_x \varphi_i(\bar{x}, y), \xi \Big\rangle > 0 \ \text{ for any } \ y \in \Omega$$

whenever Lagrange multipliers $(\lambda_1, \ldots, \lambda_s) \ne 0 \in \mathbb{R}^s$ satisfy the conditions

$$\sum_{i=1}^{s} \lambda_i \nabla_y \varphi_i(\bar{x}, y) = 0 \ \text{ and } \ \lambda_i \ge 0,\ \lambda_i \varphi_i(\bar{x}, y) = 0 \ \text{ as } \ i = 1, \ldots, r.$$

(i) Compare MMFCQ with the extended Mangasarian-Fromovitz constraint qualification introduced in [396] for the so-called generalized semi-infinite programs (GSIPs); cf. also EMFCQ for standard SIPs formulated in Definition 8.2.

(ii) Show that in the case where Ω is the argmaximum set (8.163) the introduced MMFCQ at $\bar{x}$ is *robust* in the sense that there is $\delta > 0$ such that MMFCQ holds on $B_\delta(\bar{x})$ relative to $S(x)$ provided that the supremum marginal function (8.161) is l.s.c. at $\bar{x}$ as in the setting of Exercise 8.99(ii).

Exercise 8.101 (Basic Subgradients of Supremum Marginal Functions). Assume that in the setting of Exercise 8.99 the supremum marginal function $\vartheta(x)$ from (8.161) is l.s.c. around $\bar{x}$ and MMFCQ holds at $\bar{x}$ relative to $S(\bar{x})$. Taking $\lambda := (\lambda_0, \lambda_1, \ldots \lambda_s) \in \mathbb{R}^{s+1}$, consider the Lagrangian

$$\mathcal{L}(x, y, \lambda) := \sum_{i=0}^{s} \lambda_i \varphi_i(x, y) \ \text{ with } \ \lambda_0 \le 0$$

due to the maximization in (8.161). Prove that there is $\varepsilon > 0$ such that for any $x \in B_\varepsilon(\bar{x})$ and $v \in \partial\vartheta(x)$ we can find $y_j \in S(x)$ as $j = 1, \ldots, n+1$ with $\sum_{j=1}^{n+1} y_j = 1$ and $\lambda = (\lambda_0, \ldots, \lambda_s) \ne 0$ with $\lambda_i \ge 0$ as $i = 1, \ldots, r$ satisfying the conditions

$$v = \sum_{j=1}^{n+1} \nabla_x \mathcal{L}(x, y_j, \lambda),\ \ \nabla_y \mathcal{L}(x, y, \lambda) = 0,\ \ \lambda_i \varphi_i(x, y) = 0 \ \text{ as } \ i = 1, \ldots, r.$$

Hint: Proceed as in the proof of [466, Theorem 3.7] by using the necessary optimality conditions for GSIPs developed in [396, Theorem 1.1].

Exercise 8.102 (Mixed Limiting Θ-Coderivatives). Let $f: X \to Y$ be a mapping between Banach spaces, where the image space Y is partially ordered by a closed and convex cone $\Theta \subset Y$ as in (8.71).

(i) Define the mixed limiting Θ-coderivatives of f at $\bar{x}$ with using the strong convergence on Y^* in the framework of Definition 8.43 and provide examples illustrating the relationships between them as well as between the corresponding mixed and normal Θ-coderivative constructions.

(ii) Investigate the possibility to improve the corresponding results of Section 8.3.4 by replacing the normal Θ-coderivatives by their mixed counterparts.

Exercise 8.103 (Θ-Coderivative Scalarization). Let $f: X \to Y$ be a locally Lipschitzian mapping between Banach spaces, and let $\Theta \subset Y$ be a closed and convex ordering cone. Consider the corresponding Θ-coderivatives from Definition 8.43 and Exercise 8.102 and do the following:

(i) Prove representation (8.72) in the general ordered Banach space setting.

(ii) Obtain scalarization formulas for the mixed limiting Θ-coderivatives for ordered mappings between Banach spaces. *Hint:* Proceed similarly to the proof of Theorem 1.32, which works in general Banach spaces; cf. [522, Theorem 1.90].

(iii) Establish scalarization formulas for the limiting normal Θ-coderivatives under appropriate *strict* Lipschitzian assumptions in the case where X is Asplund. *Hint:* Proceed similarly to [522, Subsection 3.1.3] where $\Theta = \{0\}$.

(iv) Verify the scalarization formulas in (8.73) for strictly differentiable mappings between Banach spaces. *Hint:* Proceed similarly to the proof of Theorem 1.32.

(v) Verify formula (7.4) for Θ-convex mappings f between Banach spaces.

Exercise 8.104 (Θ-Coderivative Calculus). Let $F: X \rightrightarrows Y$ be a set-valued mapping between Banach spaces, where the image space Y is partially ordered by a nonempty set $\Theta \subset Y$ (which may not be cone) as in (8.71).

(i) Define appropriate versions of the normal and mixed Θ-coderivatives considered above in the single-valued case and establish relationships between them.

(ii) Develop appropriate calculus rules for such coderivatives of single-valued and set-valued mappings partly in general Banach and mainly in Asplund space settings. *Hint:* Proceed similarly to the coderivatives of mappings without ordering structures by using suitable versions of the extremal principle and associated results.

(iii) Investigate the possibility of employing Θ-coderivative calculus to the study of structural optimization-related problems and constrained systems described by set-valued mappings with Θ-ordered image spaces.

Exercise 8.105 (Sequential Optimality Conditions for Conic Programs for Θ-Convex Constraint Mappings). Proceeding as in the proof of Theorem 8.50 show that its conclusions hold if the local Lipschitz continuity of f is replaced by that of the scalarized function $x \mapsto \langle y^*, f(x)\rangle$ for all $y^* \in \Theta^+$, which is always the case where $f: X \to Y$ is a continuous Θ-convex mapping. Compare this result with the corresponding one in [388] for reflexive Banach spaces X.

Exercise 8.106 (Covering and Lipschitzian Stability of Cone-Constrained and Infinite Nonconvex Inequality Systems).

(i) Establish counterparts of the results in Theorems 8.53 and 8.54 for the covering and Lipschitz-like properties of cone-constrained systems.

(ii) Establish counterparts of the results in Theorem 8.59 for the covering and Lipschitz-like properties of the infinite parametric inequality systems under consideration in Subsection 8.3.5.

Exercise 8.107 (Well-Posedness Properties of Nonconvex Cone-Constrained and Infinite Inequality Systems in Infinite Dimensions).

(i) Do formulation and proof of Theorem 8.53 need any change when the domain/parameter space X is Asplund?

(ii) Extend Theorem 8.54 to conic systems with Asplund domain spaces.

(iii) Establish a counterpart of Theorem 8.54 for the Lipschitz-like property of cone-constrained systems with Asplund image and Banach domain spaces.

(iv) Apply the results from (ii) and (iii) to derive the corresponding counterparts of Theorem 8.59 for infinite nonconvex inequality systems.

Exercise 8.108 (Comparison with Characterizing Well-Posedness Properties for Linear and Convex Inequality Systems). Present specifications for the cases of linear and convex infinite inequality systems of the results on well-posedness properties obtained in Subsection 8.3.5 and the exercises above and then compare them with those established in Sections 7.1 and 7.3.

Exercise 8.109 (Optimality Conditions for Nonconvex and Nonsmooth SIPs in Infinite Dimensions). Extend necessary optimality conditions from Theorem 8.56 and Corollary 8.57 to SIPs with Asplund decision spaces.

Exercise 8.110 (Comparison Between Necessary Optimality Conditions for Lipschitzian SIPs). Considering SIPs in finite-dimensional spaces with Lipschitzian cost and uniformly Lipschitzian inequality constraint functions in the absence of geometric constraints, clarify relationships between the necessary optimality conditions for them obtained in Subsections 8.2.3 and 8.3.5, respectively.

Exercise 8.111 (CHIP Versions for Convex Sets). Let $\{\Omega_i\}_{i\in\mathbb{N}}$ be a countable system of convex sets with a common point $\bar{x}$.

(i) Prove that CHIP for this system can be equivalently represented by (8.117).

(ii) Give an example showing that CHIP doesn't imply the strong CHIP at $\bar{x}$ even for finitely many convex sets. *Hint:* Compare it with [69, 253].

(iii) Does Proposition 8.61 hold for arbitrary (not just countable) intersections?

Exercise 8.112 (Full Duality for Convex Sets). Verify that the full duality properties in (8.119) are satisfied for convex sets.

Exercise 8.113 (Violation of CHIP for Systems of Convex Sets). Construct examples of convex systems containing finitely many as well as countably many sets in $\mathbb{R}^2$ for which CHIP is violated.

Exercise 8.114 (Distance to Contingent Directions). Prove the representations in (8.125). *Hint:* Compare it with [678, Exercise 4.8] and the guides therein.

Exercise 8.115 (CHIP for Countable Linear Inequality Systems via the Farkas Lemma). Give an alternative proof of Corollary 8.68 based on the classical Farkas lemma extended to countable linear inequalities.

Exercise 8.116 (Comparison Between the Normal Closedness and Farkas-Minkowski Properties of Countable Systems). Establish relationships between the normal closedness condition in (8.130) and the versions of the Farkas-Minkowski property for infinite linear, convex, and differentiable systems (see Sections 7.2, 7.5, and 8.1) in finite and infinite dimensions. The closure operation in (8.130) in infinite dimensions is understood in the weak* topology of X^*.

Exercise 8.117 (Interior of the Regular Normal Cone for Countable Cone Intersections). In the setting of Lemma 8.74, derive an upper estimate of the interior of $\widehat{N}\big(0;\bigcap_{i=1}^{\infty}\Lambda_i\big)$. *Hint:* Use Theorem 8.73.

Exercise 8.118 (Necessary Conditions for SIPs with Countable Operator Constraints). Consider the SIP with countable operator constraints

$$\text{minimize } \varphi(x) \text{ subject to } x\in f^{-1}(\Theta_i) \text{ as } i\in\mathbb{N},$$

where $\varphi\colon X\to\overline{\mathbb{R}}$, $\Theta_i\subset Y$ for $i\in\mathbb{N}$, and $f\colon X\to Y$ is strictly differentiable at $\bar{x}$ with the surjective derivative. Derive upper subdifferential and lower subdifferential optimality conditions for this problem in terms of its given data. *Hint:* Use the representation for the normal cone to inverse images from Exercise 1.54(ii) and the one for contingent cone from [678, Exercise 6.7] while extending the latter to the case of strictly differentiable mappings.

Exercise 8.119 (Necessary Conditions for SIPs with Countable l.s.c. Inequality Constraints). Establish an extension of Corollary 8.80 to SIP (8.154) with l.s.c. constraint functions φ_i. *Hint:* Proceed as in the proof of Corollary 8.80 with the usage of the calculus result of Exercise 2.51 for l.s.c. functions.

Exercise 8.120 (Relationships Between the Farkas-Minkowski and Local Farkas-Minkowski Properties for Countable Convex Inequalities). Prove that FMCQ⇒LFM and show that the opposite implication fails.

Exercise 8.121 (Relationships Between SCQ and CHIP for Finite Convex Inequality Systems). Consider the convex inequality system

$$\Omega_i := \big\{x \in X \big|\ \varphi_i(x) \le 0\big\}, \quad i = 1, \ldots, m,$$

where all the functions $\varphi_i : \mathbb{R}^n \to \overline{\mathbb{R}}$ are convex.

(i) Verify that SCQ⇒CHIP at $\bar{x}$ if the functions φ_i are smooth around this point. Does it hold in infinite dimensions?

(ii) Does a counterpart of (i) hold for nonsmooth convex functions?

8.6 Commentaries to Chapter 8

Section 8.1. This section is based on author's joint paper with Nghia [549] devoted by variational analysis of infinite constraint systems described via inequalities and equalities by nonconvex while differentiable functions. The main emphases in [549] was on obtaining precise normal cone representations for such infinite set intersections over arbitrary collections of indices. This was done under new constraint qualifications and subsequently applied to deriving necessary optimality conditions for SIPs with nonconvex infinite constraints and generally nonsmooth cost functions. Note that the *uniform strict differentiability* property of infinite families of functions φ_t from our standing assumptions was introduced in [549] as a natural extension of the strict differentiability of finitely many functions at the reference point. It is more general than the *equicontinuity* of the gradients $\nabla\varphi_t(x)$ defined by Seidman [686] for SIPs with compact index sets; see [549] for more discussions.

From the very beginning of semi-infinite programming, conventional SIPs concerned infinite systems over *compact* index sets with the *continuous* dependence of inequality constraint functions on index variables. The major constraint qualification for nonconvex while differentiable problems of this kind was introduced by Jongen, Twilt, and Weber [394] as *EMFCQ* from Definition 8.2. Since that this condition has been widely used in many publications to study various issues (including necessary optimality conditions of the KKT type) for SIPs with compact index sets; see, e.g., [96, 157, 394, 396, 418, 420, 475, 686, 688]. As shown in Example 8.3, KKT necessary optimality conditions fail under the validity of EMFCQ even for simple two-dimensional SIPs with countably many inequality constraints.

In the case of general differentiable SIPs with arbitrary index sets, we suggested in [549] a more appropriate SIP counterpart of MFCQ labeled as *PMFCQ* in Definition 8.4. For convex SIPs this new constraint qualification is equivalent to the *strong Slater condition* used in Chapter 7, while in the general setting of Section 8.1, the introduced PMFCQ is crucial to establish the desired normal cone representations for the infinite constraint set under consideration. The finest normal cone representations were obtained in [549] when PMFCQ was combined with the closedness-type condition *NFMCQ* from Definition 8.8. Corollary 8.15 for systems with unbounded gradients of inequality constraint functions was inspired by the corresponding result by Seidman [686]. The necessary optimality conditions for SIPs presented in Subsection 8.1.3 were derived in [549] from the obtained normal cone representations for infinite constraint sets and subdifferential sum rules of variational analysis.

To conclude the commentaries on the results presented in Section 8.1, we mention the claim in [374] about the possibility of an easier device of necessary optimality conditions obtained first in [549] while using instead a preliminary convexification and then applying the corresponding conditions for convex SIPs. Besides the incorrect usage on [374, p. 428] of Sion's minimax theorem similarly to [373] (see Section 7.7), the approach of [374] to deriving necessary optimality conditions for SIPs doesn't seem to be easier overall than our device in [549] via establishing normal cone calculus of its own interest. Indeed, the reduction in [374] was based on the nontrivial formula by López and Volle [476] on subdifferentiation of maximum functions together with the rather involved result given in [374, Lemma 19.29].

Section 8.2. This section is devoted to developing another approach to SIPs with infinitely many inequality constraints $\varphi_t(x) \le 0$, $t \in T$, indexed generally by an arbitrary set T. Such constraints can be equivalently reduced to the single constraint

$$\psi(x) := \sup\left\{\varphi_i(x)\middle|\ t \in T\right\} \le 0 \tag{8.164}$$

given by the *supremum function* $\psi(x)$, which is intrinsically nonsmooth even when all φ_t are differentiable. In this section we follow the author's joint paper with Nghia [550] and consider the case when the functions φ_t are locally *Lipschitzian*. The study of subdifferential properties of the supremum functions (8.164) generated by *convex* (or locally convex; in particular, smooth) functions φ_t has been an old topic of nonsmooth analysis; see, e.g., [203, 234, 352, 378, 476, 331, 728, 757] and the references therein. The precise subdifferential formula of convex analysis

$$\partial\psi(\bar{x}) = \mathrm{cl}^*\mathrm{co}\Big[\bigcup\Big\{\partial\varphi_t(\bar{x})\Big|t \in T(\bar{x})\Big\}\Big],\ \ T(\bar{x}) := \Big\{t \in T\Big|\varphi_t(\bar{x}) = \psi(\bar{x})\Big\} \tag{8.165}$$

has been established by Ioffe and Tikhomirov [378, Theorem 4.2.3] provided that T is a Hausdorff compact, that the mapping $t \mapsto \varphi_t(x)$ is upper semicontinuous (u.s.c.) for each x, and that the functions φ_t are continuous at $\bar{x}$.

Several counterparts of (8.165) in the inclusion and equality forms were obtained for convex functions with no assumptions imposed on the topological structure of T and on behavior of φ_t with respect to t by using the perturbation

$$T_\varepsilon(\bar{x}) := \big\{t \in T\big|\varphi_t(\bar{x}) \ge \psi(\bar{x}) - \varepsilon\big\},\quad \varepsilon \ge 0,$$

of the active index set first introduced by Valadier [728]. To the best of our knowledge, the most powerful results in this direction were obtained by Hantoute, López, and Zălinescu [331] and by López and Volle [476] via the ε-subdifferentials of convex analysis for the functions φ_t at $\bar{x}$ with no (semi)continuity requirements on $\varphi_t(\cdot)$. The functions $\varphi_t(\cdot)$ were not even assumed to be convex in [476], but the situation was actually reduced to convexity under the relaxation assumption

$$\psi^{**}(x) = \sup_{t\in T}\psi_t^{**}(x)$$

via the biconjugate functions imposed in both papers [331, 476].

If the functions φ_t are uniformly Lipschitzian around $\bar{x}$, then the inclusion

$$\overline{\partial}\psi(\bar{x}) \subset \mathrm{cl}^*\mathrm{co}\Big[\bigcup\big\{\overline{\partial}^{[T]}\varphi_t(\bar{x})\big|\ t \in T(\bar{x})\big\}\Big] \tag{8.166}$$

for the generalized gradient of the supremum function ψ over the metrizable compact T under the u.s.c. assumption on $t \mapsto \varphi_t(x)$ was derived by Clarke [165, Theorem 2.8.2] by reducing it to the convex case of (8.165). The extended subdifferential construction $\overline{\partial}^{[T]}\varphi_t(\bar{x})$ in (8.166) was defined in [165] by

$$\overline{\partial}^{[T]}\varphi_t(\bar{x}) := \text{cl}^*\text{co}\Big\{x^* \in X^* \;\Big|\; \text{there exist } t_k \overset{T_\varepsilon(\bar{x})}{\to} t,\; x_k \to \bar{x},\; \text{and } x_k^* \in \overline{\partial}\varphi_{t_k}(x_k) \text{ such that } x^* \text{ is a weak}^* \text{ cluster point of } x_k^*\Big\}.$$

The upper estimate (8.166) has been widely applied to various problems in SIP, control theory, etc. over compact index/continuous time sets; see, e.g., [165, 778, 783] and the references therein. The first results to evaluate $\overline{\partial}\psi(\bar{x})$ with no topological requirements on T (in fact for arbitrary index sets) were given in our paper [550]; see, in particular, Exercises 8.96 and 8.97. The notion of *subsmoothness* and its modifications used in Exercise 8.96(iii) have been largely studied in the literature; see [40, 673, 523, 529, 609, 775, 778] among other publications. Generalized gradients evaluations for maxima of such functions over compact index sets obtained in [550] strengthened the previous ones established in [778].

The major emphasis in [550] as well as in Section 8.2 is on evaluating the basic subdifferential $\partial\psi(\bar{x})$ of the supremum functions (8.164), with taking into account its nonconvexity, via subgradients of φ_t in the case of arbitrary index sets. The proof of the preliminary while important technical result of Lemma 8.27 on fuzzy upper estimating the regular subdifferential follows the approach to fuzzy calculus via optimization techniques developed by Borwein and Zhu [113]. The most efficient estimates of the basic subdifferential of the supremum functions are established for the new and rather broad class of *equicontinuously subdifferentiable* functions introduced in [550] as a nonsmooth extension of the uniform strict differentiability for infinite systems taken from [549]. The results obtained for evaluating the subdifferential constructions $\partial\psi(\bar{x})$ and $\overline{\partial}\psi(\bar{x})$ are applied then to deriving various forms of necessary optimality conditions for Lipschitzian SIPs under appropriate constraint qualifications of the generalized PMFCQ and NFMCQ types as defined in Subsection 8.2.3. Note also that other constraint qualifications were employed in [405, 406, 778] for deriving some optimality conditions for Lipschitzian SIPs in the case of finite-dimensional spaces X and compact sets T in terms of generalized gradients in [405, 778] and basic subgradients in [406].

The class of *supremum marginal functions* (8.161), introduced and studied in the joint author's paper with Li, Nghia, and Pham [466], happens to be essentially more complicated from the viewpoint of (lower) generalized differentiation than the standard (infimum) marginal functions investigated in Chapter 4 as well as the supremum functions of the SIP type (8.164) considered above in Section 8.2. The main challenge, besides the supremum operation, comes from the *variable* constraint set $G(x)$ under maximization in (8.161). The basic subdifferential evaluation, obtained in [466, Theorem 3.7], is presented in Exercise 8.101 under the new *marginal MFCQ* (MMFCQ). The proof of this result relies, besides other things, on the necessary optimality conditions for the so-called *generalized SIPs* taken from the paper by Jongen, Rückmann and Stein [396]. Deriving the subdifferential formula for (8.161) was motivated in [466] by a subgradient extension of the classical Łojasiewicz gradient inequality from semialgebraic geometry [473] and its subsequent applications to error bounds of parametric polynomial systems, higher-order stability analysis, and explicit convergence rates of various algorithms. However, the spectrum of other potential applications of this result is much broader; see commentaries to Section 8.5.

Section 8.3. This section is based on yet another joint paper with Nghia [553] devoted to problems of *conic programming* (or *cone-constrained optimization*) in infinite-dimensional spaces that encompass a class of Lipschitzian SIPs with infinitely many inequalities as well as additional geometric constraints. Problems of this type are important and challenging from the viewpoint of optimization theory, while they are motivated by a large variety of practical applications including those in operations research, engineering and financial management, systems control, best approximation, portfolio optimization, etc. Among the most remarkable special classes in cone-constrained optimization, there are problems of semidefinite programming, second-order cone programming, and copositive programming; see [12, 95, 96, 128, 208, 209, 494, 556, 561, 562, 625, 627, 688, 700, 730, 738, 747, 766, 784, 785] and the references therein for more details, discussions, and various applications.

Note that the vast majority of publications on conic programming concerns the settings where the underlying convex cone Θ in (8.64) is finite-dimensional. It is not the case of our applications to SIPs in [553], where Θ are positive cones of the Banach (non-Asplund unless T is a finite set)) spaces $\mathcal{C}(T)$ and $l^\infty(T)$, independently on dimensionality of the decision space X as well on compactness or noncompactness of the index set T. This is taken into account in the conic programming theory developed in [553] and reproduced in Section 8.3 with the subsequent applications to Lipschitzian SIPs (8.108) indexed by an arbitrary set T. In this theory we address the general Banach space setting for Y in (8.64) and study not only necessary optimality conditions of different kinds for conic programs and SIPs but also subdifferentiation of the scalarized supremum functions (8.66) and metric regularity of general cone-constrained systems together with the implementation of the latter for the case of infinite inequality constraints in SIPs.

Note that the *fuzzy* necessary optimality conditions for cone-constrained problems derived in the *qualified/KKT* form while without *constraint qualifications*. They seem to be new even for nonlinear programs, since the previous results produced only conditions of the Fritz John type for smooth and Lipschitzian problems; see, e.g., [111, 523, 587, 611]. As observed by Nghia [614], the local Lipschitz continuity of the functions φ_i in Corollary 8.52 can be relaxed to their lower semicontinuity for $i = 0, \ldots, m$ and their continuity for $i = m+1, \ldots, m+r$. Note also that the obtained fuzzy optimality conditions yield the so-called *sequential* ones known for special classes of optimization problems as in [388, 708].

To derive the *pointbased* necessary optimality conditions as well as the pointbased criteria for metric regularity for cone-constrained systems, we introduce by following [553] several modifications of the limiting Θ-*coderivatives* for mappings with values in *ordered Banach* spaces by using different types of convergence in dual spaces. They can also be useful for other applications; in particular, to multiobjective optimization and economic modeling (cf. Chapters 9 and 10). As mentioned, the limiting Θ-coderivative constructions in Definition 8.43 employed above are the *normal* type. Their *mixed* counterparts can be defined similarly; see Exercise 8.102.

Observe that the result of Corollary 8.57 for SIPs with compact index sets significantly improves the previous result for the same model obtained by Zheng and Yang [783] by a completely different device in a weaker form under a stronger qualification condition. The principal difference between Corollary 8.57 and the corresponding result of [783] is that the latter employs in the formulations of qualification and optimality conditions the so-called "Clarke epi-coderivative" defined in [783], which is always larger (much larger as a rule) than our basic sequential limiting coderivative used in Corollary 8.57. In particular, the qualification condition from [783] fails in Example 8.58 while the result of Corollary 8.57 holds true and confirms the optimality of the reference feasible solution.

Section 8.4. The last section of this chapter is based on the author's joint papers with Phan [568, 569] and is mainly devoted to evaluation generalized normals to *countable* set intersections and their applications to necessary optimality conditions for SIPs with *countably many* set and inequality constraints described by l.s.c. functions. The major machinery of our study relies on the conic and contingent *extremal principles* for countable systems of sets [568] presented in Section 2.2. Note that countable constraint systems are much less investigated in comparison with those indexed by compact sets being actually of the same level of difficulties as constraint systems with arbitrary index collections. However, the methods and results presented in this section essentially exploit the countable structure of constraints.

First we address by following [569] the *CHIP* and *strong CHIP* notions (the terminology is taken from [156]), which have been mainly investigated for of convex sets; see, e.g., [69, 131, 156, 204, 253, 463]. Observe that the notion of strong CHIP for finitely many convex sets postulates the conclusion of the classical Moreau-Rockafellar theorem on representing the normal cone to set intersections [667]. Our major attention in this direction turns to countable intersections of convex and nonconvex sets in finite dimensions and concerns also the new property of the *asymptotic strong CHIP*, which happens to be equivalent to strong CHIP for convex sets while playing an independent role in nonconvex settings. Corollary 8.62 for linear inequality systems reduces to the result by Cánovas et al. [140]. The notion of *bounded linear regularity* used in Proposition 8.64 and its corollary for CHIP of nonconvex countable set intersections is an extension of the corresponding property introduced and investigated by Bauschke, Borwein, and Li [69] for finite intersections of convex sets; see also [253, 463, 694] among other publications.

The qualification conditions from Definition 8.69 (together with the related ones in Definition 8.79) and the normal cone formula for countable intersections from Theorem 8.75 first appeared in [569], while the other material of Subsection 8.4.2 is taken from [568]. These results are crucial for applications to necessary optimality conditions for SIPs with countable set and inequality constraints presented in Subsection 8.4.3 based on [569]. Note that for the convex inequality systems in (8.154), both CHIP and SCC (8.156) are implied by the *local Farkas-Minkowski* property from Definition 8.81 that follows [297, 298]. We refer the reader to [262] for a detailed study of conventional qualification conditions for systems of convex inequalities. Relationships between the qualification and optimality conditions for countable convex systems presented in Proposition 8.82 and their previous counterparts established in Chapter 7 are illustrated in various settings of Example 8.84.

Let us finally mention in our comments here the recent results by Movahedian [596] concerning SIPs with *countable equality* constraints described by Lipschitzian functions on Asplund spaces, where necessary optimality conditions are derived in a pointbased KKT form under a new *boundary MFCQ*. The major role of the latter constraint qualification is to justify the validity of the required *coderivative calculus* rules and the *calmness* property of the corresponding set-valued mapping.

Section 8.6. As for the preceding chapters of the book, this section contains various exercises on the material presented in Chapter 8 of different levels of difficulties. Some of them can be fulfilled by just following the given basic results and proofs with clarifying the constructions and facts therein while other ones require a significant additional work by using supplied hints and references.

There are also some items in Section 8.6, which contain challenging issues. First we mention further developments and applications of the class of *supremum marginal functions* defined in (8.161). Functions of this type appear in many areas of variational analysis, optimization, control, etc. In particular, they describe *Hamiltonians* for control systems with control sets depending on state variables, which appear, e.g., in *feedback control* with a variety of applications to engineering design, mechanics, economics, etc. Pointbased subdifferentiation of such functions brings a lot of information for the further implementation in these and related areas of applications. Besides this, the lower subdifferentiation of the supremum marginal functions, essentially more challenging that than of the infimum ones (4.1), provides efficient tools for analysis of both *optimistic* and *pessimistic* models in *bilevel programming* in vein of the value function approach discussed in Chapter 6.

Another promising area of further extensions and applications concerns developing comprehensive calculus rules for both normal and mixed limiting Θ-*coderivatives* for single-valued and set-valued *ordered* mappings; see Exercises 8.102–8.104. This has great potential for applications not only to the topics discussed in Section 8.3 but also to *multiobjective optimization* problems considered in Chapters 9 and 10. Furthermore, the efficient *calculation* of the Θ-coderivative constructions involved in particular classes of the constrained problems under consideration is a challenging issue for both theoretical and numerical aspects of stability and optimization.

Chapter 9
Variational Analysis in Set Optimization

Here we start studying problems of *set optimization* and interrelated ones of *multiobjective optimization*, where optimal solutions are understood in various *Pareto-type* sense with respect to general preference relations. Our study equally applies to the cases of *set-valued* and *single-valued* objectives (problems of the latter type are usually considered belonging to *vector optimization*) by reducing both of them to *minimal/efficient* points of sets and employing geometric ideas of variational analysis. Note that the areas of set and set-valued optimization are relatively new and their developments have been strongly motivated by applications. Some applications to economics, where the objective set-valuedness is crucial, are presented in the final Chapter 10.

This chapter is devoted to formulations of general problems of set and set-valued optimization and to their investigation by using powerful techniques of variational analysis and generalized differentiation. Developing *variational principles* for set-valued mappings, we then apply them, together with the tools and calculus of generalized differentiation presented above, to establishing *existence theorems* for multiobjective optimal/efficient solutions and deriving *necessary optimality conditions* in unconstrained and constrained frameworks. Our approach to these issues is different from conventional ones in vector and set-valued optimization. It employs *geometric dual-space* techniques mainly based on the *extremal principle* and related developments without using any scalarization, tangential approximations, and the like. To emphasize the dual-space ideas and having in mind a variety of applications, we work in this and the next chapters in infinite-dimensional settings. Unless otherwise stated, all the spaces are assumed to be *Banach*.

9.1 Minimizers and Subdifferentials Induced by Cones

We first consider some notions of minimal points of sets with respect to preference relations induced by convex cones and then use them to define optimal solutions to problems of multiobjective optimization.

B. S. Mordukhovich, *Variational Analysis and Applications*, Springer Monographs in Mathematics, https://doi.org/10.1007/978-3-319-92775-6_9

9.1.1 Minimal Points of Sets

Let $\Theta \subset Z$ be a *closed* and *convex cone* of a space Z. We associate with this cone a *preference* $\preceq$ on Z defined by

$$z_1 \preceq z_2 \Longleftrightarrow z_2 - z_1 \in \Theta, \tag{9.1}$$

where the dependence of $\preceq$ on Θ is omitted for simplicity.

We begin with recalling standard notions of Pareto-type *minimal/efficient* and *weak minimal/weak efficient* points of sets in linear topological spaces partially ordered by cone-generated preferences of type (9.1).

Definition 9.1 (Pareto Minimal and Weak Minimal Points of Sets). *Given* $\Xi \subset Z$ *and preference* $\preceq$ (9.1) *generated by the cone* Θ*, we say that:*

(i) *The point* $\bar{z} \in \Xi$ *is* PARETO MINIMAL *for* Ξ *if*

$$(\bar{z} - \Theta) \cap \Xi = \{\bar{z}\}. \tag{9.2}$$

(ii) *The point* $\bar{z} \in \Xi$ *is* WEAK PARETO MINIMAL *for* Ξ *if*

$$(\bar{z} - \operatorname{int}\Theta) \cap \Xi = \emptyset, \quad \operatorname{int}\Theta \neq \emptyset. \tag{9.3}$$

Note that the minimality property (9.2) is formally different from the conventional definition of minimal points of sets given by

$$(\bar{z} - \Theta) \cap \Xi \subset \bar{z} + \Theta \tag{9.4}$$

while the notions (9.2) and (9.4) are obviously equivalent when the ordering cone Θ is *pointed*, i.e., $\Theta \cap (-\Theta) = \{0\}$. In fact, even in the case of nonpointed cones Θ, the conventional construction (9.4) reduces to (9.2) by considering therein the pointed ordering cone defined by

$$\widetilde{\Theta} := \big(\Theta \cap (Z \setminus (-\Theta))\big) \cup \{0\}.$$

Indeed, it is easy to check that the collection of minimal points of Ξ in the sense of (9.4) with respect to Θ is the same as the one for Ξ in the sense of (9.2) with respect to the pointed cone $\widetilde{\Theta}$. We prefer to use in what follows the minimal point definition (9.2) in both pointed and nonpointed cases, since it allows us to reduce (see below) Pareto minimality and related notions to set *extremality* in the basic sense of variational analysis as considered in Chapter 2.

A visible disadvantage of weak minimal points (9.3) is the *nonempty interior* requirement on the ordering cone Θ, which seems to be a serious restriction from both viewpoints of optimization theory and applications. In particular, various vector optimization problems can be formalized by using convex ordering cones having empty interiors in both finite-dimensional and infinite-dimensional frameworks. In settings of this type, the usage of appropriate *relative* interior points of the corresponding ordering cones seems to be reasonable provided, of course, that such points exist.

Recall that the standard *relative interior* of $\Theta \subset Z$, denoted by $\mathrm{ri}\,\Theta$, is the interior of Θ relative to the closed affine hull of Θ. It is well known that $\mathrm{ri}\,\Theta \neq \emptyset$ for every nonempty convex set Θ in *finite dimensions*. However, it is not the case in many infinite-dimensional settings. In particular, it is well known that the natural ordering cones in the standard Lebesgue spaces of sequences l^p and functions $L^p[0, 1]$ for $1 \leq p < \infty$ as well as in a number of other classical infinite-dimensional spaces have *empty relative interiors*.

To improve this situation, consider the following extensions of the relative interior notion for sets in infinite dimensions.

Definition 9.2 (Quasi-Relative and Intrinsic Relative Interiors of Convex Sets). *Let $\Theta \subset Z$ be a convex set. Then:*

(i) *The* QUASI-RELATIVE INTERIOR *of Θ, denoted by* $\mathrm{qri}\,\Theta$, *is the collection of those $z \in \Theta$ for which the closed conic hull* $\mathrm{cl\,cone}\,(\Theta - z)$ *of the shifted set $\Theta - z$ is a linear subspace of Z.*

(ii) *The* INTRINSIC RELATIVE INTERIOR *of Θ, denoted by* $\mathrm{iri}\,\Theta$, *is the collection of those $z \in \Theta$ for which the conic hull* $\mathrm{cone}\,(\Theta - z)$ *of the shifted set $\Theta - z$ is a linear subspace of Z.*

Note that the definition of $\mathrm{iri}\,\Theta$, in contrast to $\mathrm{ri}\,\Theta$ and $\mathrm{qri}\,\Theta$, is pure algebraic without involving any topology. It is obvious that

$$\mathrm{ri}\,\Theta \subset \mathrm{iri}\,\Theta \subset \mathrm{qri}\,\Theta, \tag{9.5}$$

where both inclusions hold as equalities if $\mathrm{ri}\,\Theta \neq \emptyset$, in particular, when Z is finite-dimensional. Some properties associated with these notions are listed in Exercise 9.24; see also the corresponding commentaries in Section 9.6. A truly remarkable property of quasi-relative interiors is that $\mathrm{qri}\,\Theta \neq \emptyset$ for any closed and convex subsets of *separable* Banach spaces.

Utilizing the above relative interior notions in the case of the ordering cone Θ in (9.1), we now introduce the corresponding notions of relative minimal points of sets that intermediate positions between Pareto and weak Pareto minimal/efficient points from Definition 9.1.

Definition 9.3 (Relative Minimal Points of Sets). Let a set Ξ be a nonempty subset of a linear topological space Z partially ordered by the closed and convex cone $\{0\} \neq \Theta \subset Z$ as in (9.1). We say that:

(i) $\bar{z} \in \Xi$ is (PRIMARY) RELATIVE MINIMAL POINT of Ξ if

$$(\bar{z} - \mathrm{ri}\,\Theta) \cap \Xi = \emptyset, \quad \mathrm{ri}\,\Theta \neq \emptyset.$$

(ii) $\bar{z} \in \Xi$ is an INTRINSIC RELATIVE MINIMAL POINT of Ξ if

$$(\bar{z} - \mathrm{iri}\,\Theta) \cap \Xi = \emptyset, \quad \mathrm{iri}\,\Theta \neq \emptyset.$$

(iii) $\bar{z} \in \Xi$ is a QUASI-RELATIVE MINIMAL POINT of Ξ if

$$(\bar{z} - \operatorname{qri}\Theta) \cap \Xi = \emptyset, \quad \operatorname{qri}\Theta \neq \emptyset.$$

It follows from Exercise 9.24(iii) that all the relative minimality notions above agree if the set Ξ admits a relative minimal point. All these notions clearly reduce to weak efficiency provided that $\operatorname{int}\Theta \neq \emptyset$, which is a restrictive assumption. In general, any quasi-relative minimal point of Ξ is an intrinsic minimal point of this set (but not vice versa), and the existence of the latter doesn't imply the existence of primary relative minimal points of Ξ and hence the existence of weak efficient points of this set; see Exercise 9.25.

9.1.2 Minimizers and Subdifferentials for Mappings

Consider next a set-valued mapping $F\colon X \rightrightarrows Z$ with values in a linear topological space, which is partially ordered by a closed and convex cone $\Theta \subset Z$. We define the notions of minimizers for it induced by cone ordering on the image space Z that is generated by the minimal points of the image set $F(X) := \bigcup_{x \in X} F(x)$ in the senses of Definitions 9.1 and 9.2.

Definition 9.4 (Global Minimizers of Set-Valued Mappings). *Given a set-valued mapping $F\colon X \rightrightarrows Z$ with values ordered by a cone Θ and given a pair $(\bar{x}, \bar{z}) \in$* gph F *from the graph of F, we say that:*

(i) $(\bar{x}, \bar{z})$ *is a* (*Pareto*) MINIMIZER *of the mapping F if*

$$(\bar{z} - \Theta) \cap F(X) = \{\bar{z}\}.$$

(ii) $(\bar{x}, \bar{z})$ *is a* WEAK MINIMIZER *for F if*

$$\big(\bar{z} - \operatorname{int}\Theta) \cap F(X) = \emptyset \ \textit{ provided that } \ \operatorname{int}\Theta \neq \emptyset.$$

(iii) $(\bar{x}, \bar{z})$ *is a* (PRIMARY) RELATIVE MINIMIZER *for F if*

$$\big(\bar{z} - \operatorname{ri}\Theta\big) \cap F(X) = \emptyset \ \textit{ provided that } \ \operatorname{ri}\Theta \neq \emptyset.$$

(iv) $(\bar{x}, \bar{z})$ *is an* INTRINSIC RELATIVE MINIMIZER *for F if*

$$\big(\bar{z} - \operatorname{iri}\Theta\big) \cap F(X) = \emptyset \ \textit{ provided that } \ \operatorname{iri}\Theta \neq \emptyset.$$

(v) $(\bar{x}, \bar{z})$ *is a* QUASI-RELATIVE MINIMIZER *for F if*

$$\big(\bar{z} - \operatorname{qri}\Theta\big) \cap F(X) = \emptyset \ \textit{ provided that } \ \operatorname{qri}\Theta \neq \emptyset.$$

Since the mapping F may take empty values for some $x \in X$, we actually have constraints $x \in \operatorname{dom} F$ for the minimizers in Definition 9.4. On the other hand, explicit constraints of the type $x \in \Omega$ and their specifications can be reduced to minimizing unconstrained set-valued mappings by imposing $F(x) := \emptyset$ for $x \notin \Omega$.

If the mapping F in Definition 9.4 is *single-valued* $F = f : X \to Z$, there is the unique choice of $\bar{z} = f(\bar{x})$ therein, and so we may speak about the corresponding minimizers $\bar{x}$ while having in mind the pair $(\bar{x}, f(\bar{x}))$ in all the properties (i)–(v) listed above.

Example 9.5 (Global Pareto and Weak Pareto Minimizers). Consider the following Fig. 9.1 to illustrate the difference between global Pareto and weak Pareto minimizers. This example concerns minimizing a mapping $F : X \rightrightarrows \mathbb{R}^2$ defined on some nonempty subset $X \subset \mathbb{R}^2$ with the range space $Z = \mathbb{R}^2$ ordered by the cone $\Theta = \mathbb{R}^2_+$. The red parts in the boundary of the image set $F(X)$ depict the global Pareto points of $F(X)$, while the yellow parts show the weak Pareto but not Pareto points of this set.

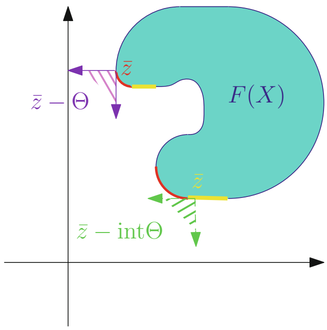


Fig. 9.1 Pareto vs. weak Pareto minimal points for $F(X)$ with $\Theta = \mathbb{R}^2_+$.

The corresponding *local* versions of all the Pareto-type global optimality notions from Definition 9.4 are formulated similarly to the above by replacing the whole space image $F(X)$ by that of a neighborhood $U \subset X$ for the domain component $\bar{x}$ of the local minimizer $(\bar{x}, \bar{z})$ in question. The next figure, Fig. 9.2, illustrates the difference between global and local Pareto minimizers in the same setting as in Example 9.5.

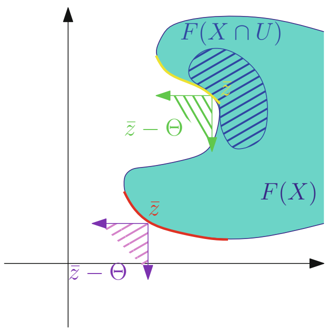


Fig. 9.2 Global Pareto vs. local Pareto minimal points for $F(X)$ with $\Theta = \mathbb{R}^2_+$.

To proceed further with the study of variational properties of set-valued mappings $F\colon X \rightrightarrows Z$ with values in partially ordering spaces and optimization problems for them, we need to introduce some notation and subdifferential notions for mappings of this type with respect to the ordering cone $\Theta \subset Z$. In addition to the closedness and convexity requirements on Θ mentioned above, suppose in what follows that Θ is *proper*, i.e., $\Theta \neq \emptyset$ and $\Theta \neq Z$. Let us associate with F and Θ the *epigraph*

$$\text{epi}_\Theta F := \big\{(x,z) \in X \times Z \big|\ z \in F(x) + \Theta\}$$

and the corresponding *epigraphical multifunction* $\mathcal{E}_{F,\Theta}\colon X \rightrightarrows Z$ defined by

$$\mathcal{E}_{F,\Theta}(x) := \big\{z \in Z \big|\ z \in F(x) + \Theta\big\} \ \text{ with } \ \text{gph}\,\mathcal{E}_{F,\Theta} = \text{epi}_\Theta F.$$

Using now the *coderivatives* of the epigraphical multifunction, we define appropriate extensions of *subdifferentials* from extended-real-valued functions to vector-valued and set-valued mappings with values in partially ordered spaces.

Definition 9.6 (Subdifferentials of Ordered Set-Valued Mappings). *Given $F\colon X \rightrightarrows Z$ with Z ordered by Θ, define the following:*

(i) *The* REGULAR SUBDIFFERENTIAL *of F at $(\bar{x},\bar{z}) \in \text{epi}_\Theta F$ is*

$$\widehat{\partial}_\Theta F(\bar{x},\bar{z}) := \big\{x^* \in X^* \big|\ x^* \in \widehat{D}^*\mathcal{E}_{F,\Theta}(\bar{x},\bar{z})(z^*),\ -z^* \in N(0;\Theta),\ \|z^*\| = 1\big\},$$

where $\widehat{D}^*$ stands for the regular coderivative/precoderivative (1.16).

(ii) *The* BASIC SUBDIFFERENTIAL *of F at $(\bar{x},\bar{z}) \in \text{epi}_\Theta F$ is*

$$\partial_\Theta F(\bar{x},\bar{z}) := \big\{x^* \in X^* \big|\ x^* \in D^*\mathcal{E}_{F,\Theta}(\bar{x},\bar{z})(z^*),\ -z^* \in N(0;\Theta),\ \|z^*\| = 1\big\},$$

where D^* stands for the (*normal*) *coderivative from Definition* 1.11*; see* (7.2).

Observe that the *range condition* $-z^* \in N(0;\Theta)$ in the constructions of Definition 9.6 is not a restriction; in fact, it automatically follows from each of the inclusions $x^* \in \widehat{D}^*\mathcal{E}_{F,\Theta}(\bar{x},\bar{z})(z^*)$ and $x^* \in D^*\mathcal{E}_{F,\Theta}(\bar{x},\bar{z})(z^*)$. This condition is presented just to reveal the possible value range of z^*. In particular, we have $z^* = 1$ for the usual order $\Theta = \mathbb{R}_+$ on the real line.

It is important to emphasize that, similarly to the case of extended-real-valued functions, the basic subdifferential $\partial_\Theta F(\bar{x},\bar{z})$ of ordered set-valued mappings enjoys a pointbased *full calculus* due to the comprehensive calculus rules available for our basic coderivative D^*; see Exercise 9.27.

9.2 Variational Principles for Ordered Mappings

The main goal of this section is to derive two variational principles for set-valued mappings with values in partially ordered spaces. The first result is a set-valued Banach space version of *Ekeland's variational principle* for extended-real-valued functions formulated in Corollary 2.13 (see Exercise 2.38), while the second one is a set-valued Asplund space version of the (lower) *subdifferential variational principle* for extended-real-valued functions formulated in Theorem 2.38; see also Exercise 2.39. Both results play a significant role in further developments in this chapter.

9.2.1 Limiting Monotonicity for Set-Valued Mappings

Given a mapping $F\colon X \rightrightarrows Z$ between Banach spaces and a set $\Xi \subset Z$, denote by $\operatorname{Min}\Xi$ *collections of Pareto minimal points* (9.2) of $\Xi \subset Z$ with respect to the ordering cone Θ on Z. This can be equivalently written as

$$\operatorname{Min}\Xi = \operatorname{Min}_{\Theta}\Xi := \big\{\bar z \in \Xi\big|\ \bar z - z \notin \Theta \ \text{ whenever } \ z \in \Xi,\ z \neq \bar z\big\}. \tag{9.6}$$

Next we recall several notions and conventional terminology from set-valued analysis and multiobjective optimization regarding the ordering cone $\Theta \subset Z$ and the mapping $F\colon X \rightrightarrows Z$ that are broadly used in what follows:

- Θ has the *normality property* if the set $(\mathbb{B}+\Theta)\cap(\mathbb{B}-\Theta)$ is bounded.
- F is *epiclosed* if its epigraph is closed in $X \times Z$.
- F is *level-closed* if its z-level sets

$$\mathcal{L}(z) := \big\{x \in X\big|\ \exists\, v \in F(x) \ \text{ with } \ v \preceq z\big\} = \big\{x \in X|\ F(x)\cap(z-\Theta) \neq \emptyset\big\}$$

are closed in X for all $z \in Z$.

- F is Θ-*quasibounded* (or simply *quasibounded*) *from below* if there is a bounded subset $M \subset Z$ with $F(X) \subset M + \Theta$. A *set* $\Omega \subset Z$ is quasibounded from below if the constant mapping $F(x) \equiv \Omega$ enjoys this property.
- F has the *domination property* at $\bar x$ if $F(\bar x) \subset \operatorname{Min} F(\bar x) + \Theta$, i.e.,

$$\text{for every } \ z \in F(\bar x)\ \text{, there is } \ v \in \operatorname{Min} F(\bar x) \ \text{ with } \ v \preceq z. \tag{9.7}$$

Note that this property is automatic for single-valued mappings F.

Let us discuss some relationships between the above properties. It is easy to see that every epiclosed mapping is level-closed, but the opposite may not be true in the case of set-valued mappings as, e.g., for $F\colon \mathbb{R} \rightrightarrows \mathbb{R}$ given by $F(x) := 0$ if $x \neq 0$ and $F(x) := (-1, 1]$ if $x = 0$. Sufficient conditions for the validity of the opposite implication are presented in Exercise 9.28. Note also that the normality property of Θ yields the pointedness property $\Theta \cap (-\Theta) = \{0\}$, while the opposite doesn't hold even for convex cones in finite dimensions as, e.g., for $\Theta := \{(z_1, z_2) \in \mathbb{R}^2|\ z_1 > 0\} \cup \{0\}$; see Exercise 9.29.

The following property is important for our extension of Ekeland's variational principle to ordered set-valued mappings.

Definition 9.7 (Limiting Monotonicity Condition). *Given* $F\colon X \rightrightarrows Z$ *and* $\bar{x} \in \operatorname{dom} F$*, we say that* F *satisfies the* LIMITING MONOTONICITY CONDITION *at* $\bar{x}$ *if any sequence of pairs* $\{(x_k, z_k)\} \subset \operatorname{gph} F$ *with* $x_k \to \bar{x}$ *satisfies*

$$\big[z_{k+1} \preceq z_k,\ k \in \mathbb{N}\big] \Longrightarrow \big[\exists\, \bar{z} \in \operatorname{Min} F(\bar{x}) \text{ with } \bar{z} \preceq z_k,\ k \in \mathbb{N}\big]. \tag{9.8}$$

Observe that the *limiting monotonicity condition* (9.8) *always implies the domination property* (9.7). Indeed, let $z \in F(\bar{x})$ and take the constant sequences $x_k \equiv \bar{x}$ and $z_k \equiv z$ as $k \in \mathbb{N}$. This clearly yields that $(x_k, z_k) \in \operatorname{gph} F$, $x_k \to \bar{x}$, and $z_{k+1} \preceq z_k$ for all $k \in \mathbb{N}$. The limiting monotonicity condition (9.8) of F at $\bar{x}$ ensures the existence of $v \in \operatorname{Min} F(\bar{x})$ with $v \preceq z = z_k$, i.e., F has the domination property at this point.

Further, it is easy to see that every level-closed and *single-valued* mapping enjoys the limiting monotonicity condition (9.8). Let us present some properties, which ensure (9.8) for *set-valued* mappings. Recall that $C \subset Z$ is a *base* for the cone Θ if $0 \notin C$ and $\Theta = \mathbb{R}_+ C$.

Proposition 9.8 (Sufficient Conditions for Limiting Monotonicity). *Let* $F\colon X \rightrightarrows Z$ *be level-closed, and let* $\bar{x} \in \operatorname{dom} F$*. Then* F *satisfies the limiting monotonicity condition at* $\bar{x}$ *if it has the domination property at this point and either one of the following assumptions is fulfilled:*

(a) *The minimum set* $\operatorname{Min} F(\bar{x})$ *is compact.*

(b) *The mapping* F *is quasibounded from below, the set* $\operatorname{Min} F(\bar{x})$ *is closed, and the ordering cone* Θ *has a compact base.*

Proof. To verify (9.8), take a sequence $\{(x_k, z_k)\} \subset \operatorname{gph} F$ such that $x_k \to \bar{x}$ as $k \to \infty$ and $z_{k+1} \preceq z_k$ for all $k \in \mathbb{N}$ and then define the sets

$$\Lambda_k := \operatorname{Min} F(\bar{x}) \cap \big(z_k - \Theta\big) = \big\{v \in \operatorname{Min} F(\bar{x})\big|\ v \preceq z_k\big\}, \tag{9.9}$$

which are closed due to the closedness of the ordering cone Θ and of the minimal set $\operatorname{Min} F(\bar{x})$. Furthermore, we have $\Lambda_{k+1} \subset \Lambda_k$ due to $z_{k+1} \in z_k - \Theta$ as $k \in \mathbb{N}$ and the convexity of Θ. Let us show that $\Lambda_k \neq \emptyset$ for all $k \in \mathbb{N}$. Indeed, fixing $k \in \mathbb{N}$ and using the monotonicity of $\{z_k\}$ give us the inclusions

$$x_{k+n} \in \mathcal{L}(z_k) \ \text{ for all } \ n \in \mathbb{N},$$

which imply that $\bar{x} \in \mathcal{L}(z_k)$ by the level-closedness of F. Thus there is $u_k \in F(\bar{x})$ with $u_k \preceq z_k$. Employing the domination property of F at $\bar{x}$, find $v_k \in \operatorname{Min} F(\bar{x})$ such that $v_k \preceq u_k \preceq z_k$, which therefore justifies the desired nonemptiness $\Lambda_k \neq \emptyset$ as $k \in \mathbb{N}$.

Next we prove that any sequence $\{v_k\} \subset \Lambda_k$ contains a subsequence converging to some $\bar{z} \in \operatorname{Min} F(\bar{x})$ if the assumptions made in either (a) or (b) are fulfilled. Observing that $\{v_k\} \subset \Lambda_1$ by the established set decreasing $\Lambda_{k+1} \subset \Lambda_k$, it remains to verify the compactness of Λ_1 under (a) and (b). It immediately follows from

(a) due to the structure of Λ_1 in (9.9). To proceed in case (b), we first recall an easily checkable fact that the compact-based property of the ordering cone Θ in (b) is equivalent to the simultaneous fulfillment of the normality property of this cone and the compactness property of the set $\Theta \cap \mathbb{B}$; see Exercise 9.30. By the quasiboundedness of F from below assumed in (b), there is a bounded set $M \subset Z$ and hence a number $m \in \mathbb{N}$ such that

$$\text{Min}\, F(\bar{x}) \subset M + \Theta \subset m\mathbb{B} + \Theta.$$

It follows directly from the structure of Λ_1 in (9.9) that

$$\Lambda_1 \subset \big(m\mathbb{B} + \Theta\big) \cap \big(\|z_1\|\mathbb{B} - \Theta\big),$$

which yields the boundedness of Λ_1 due to the normality property of Θ. Therefore we get from (9.9) that the set $z_1 - \Lambda_1 \subset \Theta$ is bounded as well. Since $\Theta \cap \mathbb{B}$ is compact in (b), the boundedness of $z_1 - \Lambda_1$ implies its compactness and so the compactness of Λ_1. The latter ensures the existence of $\bar{z} \in \text{Min}\, F(\bar{x})$ with

$$\bar{z} \in \bigcap_{k=0}^{\infty} \Lambda_k \ \text{ for all } \ k \in \mathbb{N}.$$

This tells us by (9.9) that $\bar{z} \preceq z_k$ as $k \in \mathbb{N}$, which justifies the limiting monotonicity condition for F at $\bar{x}$ in case (b) and completes the proof. △

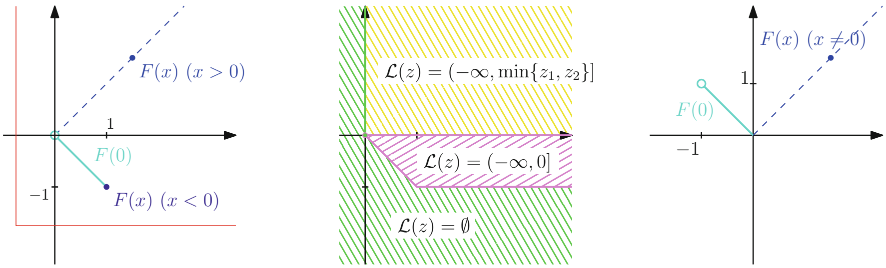
Observe that the *closedness* assumption on the set $\text{Min}\, F(\bar{x})$ is *essential* in Proposition 9.8 as in the case of $\Theta = \mathbb{R}^2_+$ and $F\colon \mathbb{R} \rightrightarrows \mathbb{R}^2$ given by

$$F(x) := \begin{cases} (x, x) & \text{if } \ x > 0,\\ \big\{(y, -y)\big|\ y \in (0, 1]\big\} & \text{if } \ x = 0,\\ (1, -1) & \text{if } \ x < 0 \end{cases}$$

for which $\text{Min}\, F(0)$ is not closed and the limiting monotonicity condition doesn't hold at $\bar{x} = 0$ although the other assumptions of Proposition 9.8(ii) are satisfied; see Fig. 9.3(a). It is easy to check that the limiting monotonicity condition (9.8) can be fulfilled *with no closedness* requirement imposed on the set $\text{Min}\, F(\bar{x})$ as for the mapping $F\colon \mathbb{R} \rightrightarrows \mathbb{R}^2$ given by

$$F(x) := \begin{cases} \big(|x|, |x|\big) & \text{if } \ x \neq 0,\\ \big\{(y, -y)\big|\ y \in (-1, 0]\big\} & \text{if } \ x = 0 \end{cases}$$

with the ordering cone $\Theta = \mathbb{R}^2_+$ and $\bar{x} = 0$; see Fig. 9.3(b).



(a) Min $F(0) = F(0)$ which is nonclosed, while the level-closedness and quasiboundedness from below are satisfied.

(b) Min $F(0) = F(0)$ which is nonclosed.

Fig. 9.3 Closedness of Min $F(\bar{x})$ and validity of limiting monotonicity

9.2.2 Variational Principle of Ekeland's Type

To establish the following set-valued version of Ekeland's variational principle, we need two more notions of minimizers for ordered mappings.

Definition 9.9 (Approximate Minimizers for Ordered Set-Valued Mappings). *Let $F\colon X \rightrightarrows Z$, where Z is partially ordered by a cone $\Theta \subset Z$. Then:*

(i) *Given $\varepsilon > 0$ and $\xi \in \Theta \setminus \{0\}$, we say that the pair $(\bar{x}, \bar{z}) \in \operatorname{gph} F$ is an* APPROXIMATE $\varepsilon\xi$-MINIMIZER *for F if*

$$z + \varepsilon\xi \not\preceq \bar{z} \ \text{ for all } \ z \in F(x) \ \text{ with } \ x \neq \bar{x}.$$

(ii) *Given $\varepsilon > 0$ and $\xi \in \Theta \setminus \{0\}$, we say that the pair $(\bar{x}, \bar{z}) \in \operatorname{gph} F$ is a* STRICT APPROXIMATE $\varepsilon\xi$-MINIMIZER *for F if there is a number $0 < \widetilde{\varepsilon} < \varepsilon$ such that $(\bar{x}, \bar{z})$ is an approximate $\widetilde{\varepsilon}\xi$-minimizer for this mapping.*

Here is the aforementioned far-going extension of Ekeland's variational principle to ordered set-valued mappings between arbitrary Banach spaces.

Theorem 9.10 (Ekeland-Type Variational Principle for Ordered Set-Valued Mappings). *Let $F\colon X \rightrightarrows Z$ be a set-valued mapping between Banach spaces, where Z is partially ordered by a proper, closed, and convex cone $\Theta \subset Z$ with $\Theta \setminus (-\Theta) \neq \emptyset$, i.e., Θ is not a linear subspace of Z. Suppose also that F is quasibounded from below, level-closed, and satisfies the limiting monotonicity condition on $\operatorname{dom} F$. Then for any $\varepsilon > 0$, $\lambda > 0$, $\xi \in \Theta \setminus (-\Theta)$ and $(x_0, z_0) \in \operatorname{gph} F$, there is $(\bar{x}, \bar{z}) \in \operatorname{gph} F$ such that*

$$\bar{z} - z_0 + \frac{\varepsilon}{\lambda}\|\bar{x} - x_0\|\xi \preceq 0, \quad \bar{z} \in \operatorname{Min} F(\bar{x}), \quad \text{and} \tag{9.10}$$

$$z - \bar{z} + \frac{\varepsilon}{\lambda}\|x - \bar{x}\|\xi \not\preceq 0 \ \text{ for all } \ (x, z) \in \operatorname{gph} F \setminus (\bar{x}, \bar{z}). \tag{9.11}$$

If furthermore (x_0, z_0) *is an approximate* $\varepsilon\xi$*-minimizer for* F*, then* $\bar{x}$ *can be chosen so that in addition to* (9.10) *and* (9.11)*, we have*

$$\|\bar{x} - x_0\| \le \lambda. \tag{9.12}$$

Proof. Note first that it is sufficient to prove the theorem in the case of $\varepsilon = \lambda = 1$. Indeed, the general case can be reduced to this special setting by applying the latter to the mapping $\widetilde{F}(x) := \varepsilon^{-1} F(x)$ on the Banach space X equipped with the equivalent norm $\lambda^{-1}\|\cdot\|$.

Having this in mind, introduce a set-valued mapping $T\colon X \times Z \rightrightarrows X$ by

$$T(x, z) := \big\{ y \in X \big|\ \exists\, v \in F(y) \ \text{ with } \ v - z + \|x - y\|\xi \preceq 0 \big\} \tag{9.13}$$

and observe that T enjoys the following properties:

• The sets $T(x, z)$ are *nonempty* for all $z \in F(x)$ by $x \in T(x, z)$.

• The sets $T(x, z)$ are *uniformly bounded* for all $z \in F(x)$ since the mapping F is quasibounded from below. Indeed, the latter property yields

$$T(x, z) \subset \big\{ y \in X \big|\ \|x - y\|\xi \in z - M - \Theta \big\}.$$

• We have the inclusion

$$T(y, v) \subset T(x, z) \text{ if } y \in T(x, z),\ \ v \in F(y),\ \ v - z + \|y - x\|\xi \preceq 0. \tag{9.14}$$

To check it, pick $u \in T(y, v)$ and by construction of T find $w \in F(u)$ satisfying

$$w - v + \|u - y\|\xi \preceq 0.$$

Summing the latter relationship with the one in (9.14) and employing

$$\big(\|x - u\| - \|u - y\| - \|y - x\|\big)\xi \preceq 0$$

that holds by the triangle inequality and the choice of $\xi \in \Theta$, we get

$$\begin{aligned} w - z + \|x - u\|\xi &= \big(w - v + \|u - y\|\xi\big) + \big(v - z + \|y - x\|\xi\big) \\ &\quad + \big(\|x - u\| - \|u - y\| - \|y - x\|\big)\xi \preceq 0, \end{aligned}$$

which therefore implies that $u \in T(x, z)$.

Let us now inductively construct a sequence of pairs $\{(x_k, z_k)\} \subset \operatorname{gph} F$ by the following *iterative procedure*: starting with (x_0, z_0) given in the theorem and having the k-iteration (x_k, z_k), we select the next one (x_{k+1}, z_{k+1}) by

$$\begin{cases} x_{k+1} \in T(x_k, z_k), \\ \|x_{k+1} - x_k\| \ge \sup\limits_{x \in T(x_k, z_k)} \|x - x_k\| - (k+1)^{-1}, \\ z_{k+1} \in F(x_{k+1}), \quad z_{k+1} - z_k + \|x_{k+1} - x_k\|\xi \preceq 0, \end{cases} \tag{9.15}$$

where $k \in \{0\} \cup \mathbb{N}$. It is clear from this construction and aforementioned properties of $T(x, z)$ that the iterative procedure (9.15) is *well defined*. Summing up the last preference relationship in (9.15) from $k = 0$ to n, we get

$$t_n \xi \in z_0 - z_{n+1} - \Theta \subset z_0 - M - \Theta \text{ with } t_n := \sum_{k=0}^{n} \|x_{k+1} - x_k\|. \tag{9.16}$$

Let us prove by passing to the limit as $n \to \infty$ in (9.16) that

$$\sum_{k=0}^{\infty} \|x_{k+1} - x_k\| < \infty. \tag{9.17}$$

Arguing by contradiction, suppose that (9.17) doesn't hold, i.e., the increasing sequence $\{t_n\}$ in (9.16) tends to ∞ as $n \to \infty$. By the first inclusion in (9.16) and boundedness of the set M in the quasiboundedness from below property of the mapping F, we find a bounded sequence $\{v_n\} \subset z_0 - M$ satisfying

$$t_n \xi - v_n \in -\Theta, \text{ i.e., } \xi - \frac{v_n}{t_n} \in -\Theta \text{ for all } n \in \mathbb{N}.$$

Passing to the limit in the latter inclusion and taking into account the closedness of Θ and the boundedness of $\{v_n\}$ and that $t_n \to \infty$ as $n \to \infty$, we arrive at $\xi \in -\Theta$. This contradicts the choice of $\xi \in \Theta \setminus (-\Theta)$ and thus justifies (9.17).

Further, it readily follows from (9.14) and (9.15) that for all $k \in \mathbb{N}$, we have $\operatorname{diam} T(x_{k+1}, z_{k+1}) \le \operatorname{diam} T(x_k, z_k)$ and

$$\operatorname{diam} T(x_k, z_k) \le 2 \sup_{x \in T(x_k, z_k)} \|x - x_k\| \le 2\big(\|x_{k+1} - x_k\| + (k+1)^{-1}\big).$$

Hence $\operatorname{diam} T(x_k, z_k) \downarrow 0$ as $k \to \infty$ due to (9.17), and hence we conclude by the completeness of X that the closures of $T(x_k, z_k)$ shrink to a *singleton*:

$$\bigcap_{k=0}^{\infty} \operatorname{cl} T(x_k, z_k) = \big\{\bar{x}\big\} \text{ with some } \bar{x} \in \operatorname{dom} F. \tag{9.18}$$

Since $\bar{x} \in \operatorname{cl} T(x_0, z_0)$, there is a sequence $\{u_n\} \subset T(x_0, z_0)$ with $u_n \to \bar{x}$. By the definition of T in (9.13), there are $v_n \in F(u_n)$ such that

$$v_n \preceq z_0 - \|x_0 - u_n\| \xi \preceq z_0,$$

, i.e., $u_n \in \mathcal{L}(z_0)$ for all $n \in \mathbb{N}$. Taking into account the closedness of $\mathcal{L}(z_0)$, we get from $u_n \to \bar{x}$ that $\bar{x} \in \mathcal{L}(z_0)$ and that $\bar{x} \in \operatorname{dom} F$ in (9.18).

Next we justify the existence of $\bar{z} \in \operatorname{Min} F(\bar{x})$ such that the pair $(\bar{x}, \bar{z})$ satisfies the major relationships in (9.10) and (9.11). Observe from the third line in (9.15) and from (9.18) that for all $k \in \mathbb{N}$, we have

$$x_k \to \bar{x} \text{ as } k \to \infty \text{ and } z_k \in F(x_k) \text{ with } z_{k+1} \preceq z_k.$$

This ensures, by the imposed limiting monotonicity condition (9.8) for the mapping F on its domain, the existence of $\bar{z} \in \operatorname{Min} F(\bar{x})$ such that $\bar{z} \preceq z_k$ for all $k \in \mathbb{N}$. Let us verify that the pair $(\bar{x}, \bar{z}) \in \operatorname{gph} F$ satisfies the desired relationships in (9.10) and (9.11).

In fact, the inclusion in (9.10) immediately follows from the choice of $\bar{z}$. To proceed further, fix $k \in \{0\} \cup \mathbb{N}$ and sum up the preference conditions in (9.15) from k to $(k+n-1)$ with that of $\bar{z} - z_k \preceq 0$. Taking into account the triangle inequality for the norm function, we get in this way that

$$\bar{z} - z_k + \|x_k - x_{k+n}\|\xi \preceq 0 \quad \text{for all} \quad k \in \{0\} \cup \mathbb{N} \quad \text{and} \quad n \in \mathbb{N}.$$

The passage to the limit above with $x_{k+n} \to \bar{x}$ as $n \to \infty$ gives us

$$\bar{z} - z_k + \|x_k - \bar{x}\|\xi \preceq 0 \quad \text{whenever} \quad k \in \{0\} \cup \mathbb{N}, \tag{9.19}$$

which justifies (9.10), in the case of $\varepsilon = \lambda = 1$ under consideration, for $k = 0$. To verify now (9.11), suppose the contrary and find a pair (x, z) satisfying

$$(x, z) \in \operatorname{gph} F \text{ with } (x, z) \neq (\bar{x}, \bar{z}) \text{ and } z - \bar{z} + \|x - \bar{x}\|\xi \preceq 0. \tag{9.20}$$

If $x = \bar{x}$ in (9.20), we get $z \neq \bar{z}$ and $z \preceq \bar{z}$, which contradict the choice of $\bar{z} \in \operatorname{Min} F(\bar{x})$. If $x \neq \bar{x}$, then we get by summing up the preference conditions in (9.19), (9.20) and combining the result with the triangle inequality that

$$z - z_k + \|x - x_k\|\xi \preceq 0, \ \text{ i.e., } x \in T(x_k, z_k) \subset \operatorname{cl} T(x_k, z_k)$$

for all $k \in \{0\} \cup \mathbb{N}$. This means that x from (9.20) belongs to the set intersection in (9.18). Thus $x = \bar{x}$ by (9.18), which fully justifies (9.11) as $\varepsilon = \lambda = 1$ and hence in the general case as well.

To complete the proof of the theorem, it remains to estimate $\|\bar{x} - x_0\|$ when (x_0, z_0) is chosen as an *approximate* $\varepsilon\xi$*-minimizer* for F. Arguing by contradiction, suppose that (9.12) doesn't hold, i.e., $\|\bar{x} - x_0\| > \lambda$. Since $\bar{x} \in T(x_0, z_0)$ and $0 \preceq \xi$, we have the preference relationships

$$\bar{z} - z_0 + \varepsilon\xi \preceq \bar{z} - z_0 + \frac{\varepsilon}{\lambda}\|\bar{x} - x_0\|\xi \preceq 0$$

which contradict the choice of (x_0, z_0) as an approximate $\varepsilon\xi$-minimizer for F and thus complete the proof of the theorem. △

The following straightforward consequence of Theorem 9.10 concerns single-valued mappings $F = f : X \to Z$ in which case both assumptions and conclusions of the theorem admit significant simplifications. Similarly to the case of Definition 9.4 in the case of single-valued mappings, we avoid mentioning the unique image point $\bar{z} = f(\bar{x})$ while referring to approximate minimizers from Definition 9.9 in this case.

Corollary 9.11 (Ekeland-Type Variational Principle for Ordered Single-Valued Mappings). *Let $f\colon X \to Z$ be a single-valued mapping between Banach spaces, where Z is partially ordered by a proper, closed, and convex cone $\Theta \subset Z$ with $\Theta \setminus (-\Theta) \neq \emptyset$. Assume that f is level-closed and quasibounded from below. Then given $\varepsilon > 0$, $\lambda > 0$, $\xi \in \Theta \setminus (-\Theta)$, and an approximate $\varepsilon\xi$-minimizer $x_0 \in X$ of f, there exists $\bar{x} \in X$ such that*

$$\|\bar{x} - x_0\| \le \lambda, \quad f(\bar{x}) + \frac{\varepsilon}{\lambda}\|\bar{x} - x_0\|\xi \preceq f(x_0),$$

and $\bar{x}$ is a global Pareto minimizer of the perturbed mapping

$$f(x) + \frac{\varepsilon}{\lambda}\|x - \bar{x}\|\xi.$$

Proof. It follows directly from Theorem 9.10 due to the validity of the limiting monotonicity condition for single-valued mappings and the corresponding expressions for the properties (9.10) and (9.11) in this case. △

9.2.3 Subdifferential Variational Principle for Mappings

The next result is the *subdifferential variational principle*, which is an extension to ordered set-valued mappings of the corresponding scalar result formulated in Exercise 2.39, which is based on [522, Theorem 2.28].

Theorem 9.12 (Subdifferential Variational Principle for Ordered Set-Valued Mappings). *Let $F\colon X \rightrightarrows Z$ be a set-valued mapping between Asplund spaces, which is epiclosed with respect to the ordering cone $\Theta \subset Z$ in addition to the assumptions of Theorem 9.10. Then for any $\varepsilon > 0$, $\lambda > 0$, $\xi \in \Theta \setminus (-\Theta)$ with $\|\xi\| = 1$ and for any strict approximate $\varepsilon\xi$-minimizer $(x_0, z_0) \in \operatorname{gph} F$ for the mapping F, there is $(\bar{x}, \bar{z}) \in \operatorname{gph} F$ such that $\|\bar{x} - x_0\| \le \lambda$ and*

$$\widehat{\partial}_\Theta F(\bar{x}, \bar{z}) \cap \frac{\varepsilon}{\lambda}\mathbb{B}^* \neq \emptyset. \tag{9.21}$$

Proof. Note first that we impose the requirement $\|\xi\| = 1$ in the formulation of the theorem to get a "nicer" subdifferential condition (9.21). As follows from the arguments below, condition (9.21) can be replaced, with no change in the proof, by the modified subdifferential condition

$$\widehat{\partial}_\Theta F(\bar{x}, \bar{z}) \cap \frac{\varepsilon}{\lambda}\|\xi\|\mathbb{B}^* \neq \emptyset$$

if ξ is selected arbitrarily from $\Theta \setminus (-\Theta)$.

To verify (9.21), take the pair $(x_0, z_0) \in \operatorname{gph} F$ from the formulation of the theorem and find a positive number $\widetilde{\varepsilon} < \varepsilon$ such that (x_0, z_0) is an approximate $\widetilde{\varepsilon}\xi$-minimizer to F. Put further

$$\widetilde{\lambda} := \frac{\varepsilon + \widetilde{\varepsilon}}{2\varepsilon}\lambda \;\text{ with }\; 0 < \widetilde{\lambda} < \lambda \tag{9.22}$$

and apply the results from Theorem 9.10 to the mapping F and its approximate $\widetilde{\varepsilon}\xi$-minimizer (x_0, z_0) with the chosen parameters $\widetilde{\varepsilon}$ and $\widetilde{\lambda}$. In this way we have $(\bar{u}, \bar{v}) \in \operatorname{gph} F$ satisfying the relationships

$$\bar{v} \in \operatorname{Min} F(\bar{u}), \quad \|x_0 - \bar{u}\| \le \widetilde{\lambda}, \quad \text{and} \tag{9.23}$$

$$z - \bar{v} + \frac{\widetilde{\varepsilon}}{\widetilde{\lambda}}\|x - \bar{u}\|\xi \not\preceq 0 \text{ for } (x, z) \in \operatorname{gph} F \setminus (\bar{u}, \bar{v}). \tag{9.24}$$

Consider a single-valued Lipschitz continuous mapping $g\colon X \to Z$ given by

$$g(x) := \bar{v} - \frac{\widetilde{\varepsilon}}{\widetilde{\lambda}}\|x - \bar{u}\|\xi \tag{9.25}$$

and define the two closed subsets of the (Asplund) product space $X \times Z$ by

$$\Omega_1 := \operatorname{epi}_\Theta F \;\text{ and }\; \Omega_2 := \operatorname{gph} g. \tag{9.26}$$

We claim that $(\bar{u}, \bar{v})$ is a locally *extremal point* of $\{\Omega_1, \Omega_2\}$ in the sense of Definition 2.1. Since the inclusion $(\bar{u}, \bar{v}) \in \Omega_1 \cap \Omega_2$ is obvious, it remains to check the existence of a sequence $\{a_k\} \subset X \times Z$ such that $a_k \to 0$ as $k \to \infty$ and $\Omega_1 \cap (\Omega_2 + a_k) = \emptyset$ for all $k \in \mathbb{N}$. To proceed, let us verify that

$$\Omega_1 \cap \big(\Omega_2 + (0, -k^{-1}\xi)\big) = \emptyset \;\text{ for all }\; k \in \mathbb{N}, \tag{9.27}$$

i.e., (2.1) holds with $a_k := (0, -k^{-1}\xi)$. Supposing the contrary and using (9.26) give us (x, v) such that

$$v = g(x) - k^{-1}\xi \;\text{ and }\; (x, v) \in \operatorname{epi}_\Theta F. \tag{9.28}$$

It follows from $(x, v) \in \operatorname{epi}_\Theta F$ that there are $z \in F(x)$ and $\theta \in \Theta$ with $v = z + \theta$. Substituting this into the equality in (9.28) and taking into account that $-\xi \preceq 0$ and $-\theta \preceq 0$, we arrive at

$$z = v - \theta = g(x) - k^{-1}\xi - \theta \preceq g(x),$$

which allows us to deduce from (9.24) by the structure of g in (9.25) that $(x, z) = (\bar{u}, \bar{v})$. The latter implies together with (9.27) and $v = z + \theta$ that

$$z = \bar{v} = g(\bar{u}) = g(x) = v + k^{-1}\xi = z + \theta + k^{-1}\xi,$$

and so $\theta + k^{-1}\xi = 0$. It yields $\xi = -k\theta \in -\Theta$ contradicting the choice of $\xi \in \Theta \setminus (-\Theta)$. The obtained contradiction ensures the fulfillment of (9.27) and hence the extremality of the set system (9.26) at the reference point $(\bar{u}, \bar{v})$.

Thus we can apply to the system $\{\Omega_1, \Omega_2, (\bar u, \bar v)\}$ the *approximate extremal principle* formulated in Corollary 2.5 in finite dimensions, which holds in any Asplund space not being anymore a consequence of the exact extremal principle; see Exercise 2.24 and [522, Theorem 2.20] for all the details. For convenience we impose the sum norm $\|(x,z)\| := \|x\| + \|z\|$ on the product space $X \times Z$ that generates the dual norm on $X^* \times Z^*$ by

$$\|(x^*, z^*)\| := \max\big\{\|x^*\|, \|z^*\|\big\} \ \text{ for } \ (x^*, z^*) \in X^* \times Z^*.$$

In this way the relationships of the approximate extremal principle allow us for any $\nu > 0$ to find $(x_i, z_i, x_i^*, z_i^*) \in X \times Z \times X^* \times Z^*$, $i = 1, 2$, satisfying

$$\begin{cases} (x_i, z_i) \in \Omega_i, \quad \|x_i - \bar u\| + \|z_i - \bar v\| \le \nu, \quad i = 1, 2, \\ (x_i^*, -z_i^*) \in \widehat N\big((x_i, z_i); \Omega_i\big), \quad i = 1, 2, \\ \dfrac{1}{2} - \nu \le \max\big\{\|x_i^*\|, \|z_i^*\|\big\} \le \dfrac{1}{2} + \nu, \quad i = 1, 2, \\ \max\big\{\|x_1^* + x_2^*\|, \|z_1^* + z_2^*\|\big\} \le \nu. \end{cases} \tag{9.29}$$

By using the structure of Ω_2 in (9.26) and the Lipschitz continuity of g in (9.25) with constant $\ell = \widetilde\varepsilon/\widetilde\lambda$, we derive from (9.29) and the regular coderivative estimate for Lipschitz-like mappings in Exercise 3.41 that

$$\|x_2^*\| \le \frac{\widetilde\varepsilon}{\widetilde\lambda}\|z_2^*\| \ \text{ and hence } \ z_2^* \ne 0$$

by (9.29) with $\nu > 0$ being sufficiently small. Employing the relationships in (9.29) again allows us to deduce that

$$\|z_1^*\| \ne 0 \ \text{ and } \ \frac{\|x_1^*\|}{\|z_1^*\|} < \frac{\varepsilon}{\lambda}; \tag{9.30}$$

see Exercise 9.37. Furthermore, from the second line in (9.29) with $i = 1$, we find $\widetilde z_1 \in F(x_1)$ satisfying the inclusions

$$(x_1, \widetilde z_1) \in \operatorname{gph} F, \ (x_1^*, -z_1^*) \in \widehat N\big((x_1, \widetilde z_1); \operatorname{epi}_\Theta F\big), \ -z_1^* \in \widehat N(0; \Theta). \tag{9.31}$$

Denoting finally $(\bar x, \bar z) := (x_1, \widetilde z_1)$, $x^* := x_1^*/\|z_1^*\|$, and $z^* := z_1^*/\|z_1^*\|$ and taking into account the regular subdifferential construction in Definition 9.6(i), we get the desired subdifferential condition (9.21) from the relations in (9.30) and (9.31). To complete the proof of the theorem, it remains to observe that the estimate $\|\bar x - x_0\| < \lambda$ follows from the second inequality in (9.23), the first line in (9.29) for $i = 1$, and the choice of $\widetilde\lambda$ in (9.22). △

9.3 Existence of Relative Pareto-Type Minimizers

This section is devoted to deriving verifiable conditions that ensure the *existence* of relative Pareto-type minimizers for general multiobjective problems defined in Subsection 9.1.2. The obtained results strongly rely on the *subdifferential Palais-Smale conditions* we introduce and discuss first.

9.3.1 Subdifferential Palais-Smale Conditions

Recall that the *classical Palais-Smale* condition for differentiable real-valued function $\varphi\colon X \to \mathbb{R}$ asserts that if a sequence $\{x_k\} \subset X$ is such that $\{\varphi(x_k)\}$ is bounded and $\|\nabla\varphi(x_k)\| \to 0$ as $k \to \infty$ for the corresponding derivative sequence, then $\{x_k\}$ contains a convergent subsequence. In the following definition, we present two extensions of this condition to ordered nonsmooth and set-valued mappings that make use of the regular and basic subdifferentials of such mappings defined above.

Definition 9.13 (Subdifferential Palais-Smale Conditions for Ordered Multifunctions). *Let $F\colon X \rightrightarrows Z$ be a set-valued mapping between Banach spaces with an ordered image space Z, and let $\widehat{\partial}_\Theta F(x,z)$ and $\partial_\Theta F(x,z)$ be the regular and basic subdifferentials of F at some point $(x,z) \in \operatorname{gph} F$, respectively, taken from Definition 9.6. We say that:*

(i) *The* REGULAR SUBDIFFERENTIAL PALAIS-SMALE CONDITION *holds for F provided that any sequence $\{x_k\} \subset X$ satisfying*

$$\text{there are } z_k \in F(x_k) \text{ and } x_k^* \in \widehat{\partial}_\Theta F(x_k, z_k) \text{ with } \|x_k^*\| \to 0 \text{ as } k \to \infty$$

contains a convergent subsequence if $\{z_k\}$ is quasibounded from below.

(ii) *The* BASIC SUBDIFFERENTIAL PALAIS-SMALE CONDITION *holds for F provided that any sequence $\{x_k\} \subset X$ satisfying*

$$\text{there are } z_k \in F(x_k) \text{ and } x_k^* \in \partial_\Theta F(x_k, z_k) \text{ with } \|x_k^*\| \to 0 \text{ as } k \to \infty$$

contains a convergent subsequence if $\{z_k\}$ is quasibounded from below.

It is clear that both subdifferential conditions in Definition 9.13 reduce to the classical Palais-Smale condition for smooth functions $F = \varphi\colon X \to \mathbb{R}$ since in this case $\widehat{\partial}\varphi(x_k) = \partial\varphi(x_k) = \{\nabla\varphi(x_k)\}$.

Note that in general the first condition in Definition 9.13 is less restrictive than the second one since we always have $\widehat{\partial}_\Theta F(x,z) \subset \partial_\Theta F(x,z)$ for any $(x,z) \in \operatorname{gph} F$. However, the latter condition has serious advantages in applications to structural problems (in particular, to those with various constraints) due to a broad spectrum of available *calculus* rules.

In the subsequent subsections of this section, we obtain existence theorems for relative minimizers first in unconstrained and then in constrained settings of multiobjective optimization. Besides using both variational principles from Section 9.2

and the Palais-Smale subdifferential conditions, we need the following modification of the limiting monotonicity property.

Definition 9.14 (Strong Limiting Monotonicity Condition). *Given a set-valued mapping $F\colon X \rightrightarrows Z$ and a point $\bar{x} \in \operatorname{dom} F$, it is said that F satisfies the* STRONG LIMITING MONOTONICITY CONDITION *at $\bar{x}$ if for any sequence of pairs $\{(x_k, z_k)\} \subset \operatorname{gph} F$ with $x_k \to \bar{x}$, we have for all $k \in \mathbb{N}$ that*

$$\big[z_k \preceq v_k,\ v_{k+1} \preceq v_k\big] \Longrightarrow \big[\exists\, \bar{z} \in \operatorname{Min} F(\bar{x}) \text{ with } \bar{z} \preceq v_k\big] \tag{9.32}$$

and also that $\bar{z} \preceq \bar{v}$ if $v_k \to \bar{v}$ as $k \to \infty$ and the ordering cone Θ is closed.

The strong limiting monotonicity (9.32) obviously implies the limiting monotonicity (9.8) but not vice versa. However, the sufficient conditions for the limiting monotonicity property given in Proposition 9.8 ensure the strong limiting monotonicity property as well; see Exercise 9.39.

9.3.2 Existence of Solutions to Unconstrained Problems

In this subsection we study the existence of relative minimizers (and consequently weak minimizers) of set-valued mappings $F\colon X \rightrightarrows Z$ between Asplund spaces. Although such a multiobjective optimization problem is considered in the unconstrained format, it implicitly includes the domain constraint $x \in \operatorname{dom} F$. Existence issues in some explicitly constrained multiobjective problems are addressed in the next subsection.

The following theorem is the main result of the whole section. Its proof is based on the *variational principles* developed in Section 9.2.

Theorem 9.15 (Existence of Intrinsic Relative Minimizers for Set-Valued Mappings). *Let $F\colon X \rightrightarrows Z$ be a mapping between Asplund spaces that is epiclosed and quasibounded from below and satisfies the strong limiting monotonicity condition from Definition 9.14 on $\operatorname{dom} F$. Assume furthermore that the regular subdifferential Palais-Smale condition from Definition 9.13(i) holds and that $\Theta \setminus (-\Theta) \neq \emptyset$, i.e., Θ is not a linear subspace of Z. Then F admits an intrinsic relative minimizer provided that $\operatorname{iri} \Theta \neq \emptyset$.*

Proof. To justify the existence of intrinsic relative minimizers for F, we first apply the Ekeland-type principle from Theorem 9.10 to generate a minimizing sequence $\{(x_k, z_k)\} \subset \operatorname{gph} F$ and then prove that the chosen sequence $\{x_k\}$ contains a subsequence converging to an intrinsic relative minimizer for F. The latter arguments are rather involved based on applying the above version of the subdifferential variational principle, the approximate extremal principle, and the limiting monotonicity condition. Details follow.

To begin with, pick an arbitrary pair $(x_0, z_0) \in \operatorname{gph} F$ and element $\xi \in \Theta \setminus (-\Theta)$ with $\|\xi\| = 1$ and then inductively generate a sequence $\{(x_k, z_k)\} \subset \operatorname{gph} F$ by using the obtained Ekeland-type variational principle for ordered set-valued mappings. To proceed, fixing $k \in \mathbb{N}$ and having the $(k-1)$-iteration (x_{k-1}, z_{k-1}), apply Theorem 9.10 with the parameters $\varepsilon := k^{-2}$ and $\lambda := k^{-1}$ to get the next iteration $(x_k, z_k) \in \operatorname{gph} F$ satisfying the relationships

$$z_k \in \operatorname{Min} F(x_k), \quad z_k \preceq z_{k-1}, \quad \text{and} \tag{9.33}$$

$$z - z_k + k^{-1}\|x - x_k\|\xi \not\preceq 0 \text{ for all } (x, z) \in \operatorname{gph} F \setminus (x_k, z_k). \tag{9.34}$$

Assume for the moment that $\{x_k\}$ contains a subsequence converging to some point $\bar{x} \in \operatorname{dom} F$; we show that it is the case a bit later. Without loss of generality, suppose that $x_k \to \bar{x}$ as $k \to \infty$ for the whole sequence $\{x_k\}$ and get from (9.33) and the limiting monotonicity condition (9.8) that

$$\text{there is } \bar{z} \in F(\bar{x}) \text{ with } \bar{z} \preceq z_k \text{ for all } k \in \mathbb{N}. \tag{9.35}$$

Let us prove that the pair $(\bar{x}, \bar{z})$ is an *intrinsic relative minimizer* for F. Indeed, taking an arbitrary pair $(x, z) \in \operatorname{gph} F$ with $(x, z) \neq (\bar{x}, \bar{z})$ and employing (9.34) and (9.35), we have by elementary transformations that

$$z - \bar{z} + k^{-1}\|x - x_k\|\xi \in z_k - \bar{z} + Z \setminus (-\Theta) \text{ for all } k \in \mathbb{N},$$

which easily implies the inclusion

$$z - \bar{z} + k^{-1}\|x - x_k\|\xi \in \Theta + Z \setminus (-\Theta).$$

The latter gives, by the convexity of the ordering cone Θ, that

$$z - \bar{z} + k^{-1}\|x - x_k\|\xi \in Z \setminus (-\Theta), \quad k \in \mathbb{N}. \tag{9.36}$$

Our aim is to show, by passing to the limit in (9.36) as $k \to \infty$, that

$$z - \bar{z} \in Z \setminus (-\operatorname{iri} \Theta) \text{ provided that } \operatorname{iri} \Theta \neq \emptyset. \tag{9.37}$$

Arguing by contradiction, suppose that (9.37) doesn't hold, i.e., $z - \bar{z} =: \theta \in -\operatorname{iri} \Theta$. Employing Definition 9.2(ii) of the intrinsic relative interior, we have that the conic hull $\operatorname{cone}(\Theta + \theta)$ is a linear subspace of Z. This allows us to find a positive number $\bar{t} \leq 1$ such that

$$t(-\xi - \theta) \in \Theta + \theta \text{ for all } t \in [0, \bar{t}], \text{ and thus}$$

$$\theta + \tau\xi \in -\Theta \text{ for all } \tau = \frac{t}{1+t} \in \left[0, \frac{\bar{t}}{1+\bar{t}}\right]. \tag{9.38}$$

Since $k^{-1}\|x - x_k\| \to 0$ as $k \to \infty$, it gives us for large $k \in \mathbb{N}$ that

$$k^{-1}\|x - x_k\| \in \left[0, \bar{t}/(1+\bar{t})\right].$$

Substituting it into (9.38) and observing that $\theta = z - \bar{z}$, we arrive at

$$z - \bar{z} + k^{-1}\|x - x_k\|\xi \in -\Theta,$$

which contradicts (9.36) and therefore verifies (9.37). Since the pair $(x, z) \in \text{gph}\, F$ was chosen arbitrarily, the conditions in (9.37) yield those in Definition 9.4(iv) and thus verify the intrinsic relative minimality of $(\bar{x}, \bar{z})$.

To complete the proof, it remains to justify the claim announced above: the chosen sequence $\{x_k\}$ contains a *convergent subsequence*. To prove this convergence, we inductively construct another sequence $\{\widetilde{x}_k\} \subset \text{dom}\, F$ such that $\|\widetilde{x}_k - x_k\| \to 0$ as $k \to \infty$ and that the *subdifferential Palais-Smale condition* from Definition 9.13(i) can be applied to this new sequence. To proceed, define for each $k \in \mathbb{N}$ a set-valued mapping $F_k\colon X \rightrightarrows Z$ by

$$F_k(x) := F(x) + g_k(x) \;\text{ with }\; g_k(x) := k^{-1}\|x - x_k\|\xi \tag{9.39}$$

and deduce from (9.34) that (x_k, z_k) is a *strict approximate* $k^{-2}\xi$*-minimizer* for F_k. It is easy to check that F_k is epiclosed and quasibounded from below. We claim now that F_k enjoys the *limiting monotonicity property* (9.8) provided that F has the strong limiting monotonicity property (9.32). Fix $\bar{u} \in \text{dom}\, F_k = \text{dom}\, F$ and for any sequence $\{(u_n, w_n)\} \subset \text{gph}\, F_k$ satisfying

$$u_n \to \bar{u} \;\text{ as }\; n \to \infty \;\text{ with }\; w_{n+1} \preceq w_n \tag{9.40}$$

for all $n \in \mathbb{N}$ and define a sequence of $\bar{w}_n \in F(u_n)$ by

$$\bar{w}_n := w_n - k^{-1}t_n\xi \;\text{ with }\; t_n := \|u_n - x_k\|, \quad n \in \mathbb{N}.$$

Passing to subsequences if necessary, we assume without loss of generality that the sequence $\{t_n\}$ *monotonically* converges to $\bar{t} := \|\bar{u} - x_k\|$ as $n \to \infty$. Consider the following two possible cases:

• If $\{t_n\}$ is *decreasing*, then $\{-k^{-1}t_n\xi\}$ is increasing. Denote $v_n := w_n - k^{-1}\bar{t}\xi$ and observe that the sequence $\{v_n\}$ is decreasing with

$$\bar{w}_n = w_n - k^{-1}t_n\xi \preceq v_n, \quad n \in \mathbb{N}.$$

Applying the strong limiting monotonicity property (9.32) of the mapping F to the sequences $\{(u_n, \bar{w}_n)\}$ and $\{v_n\}$, we find $\bar{w} \in \text{Min}\, F(\bar{u})$ with $\bar{w} \preceq v_n$ for all $n \in \mathbb{N}$. This clearly implies that

$$\bar{w} + k^{-1}\bar{t}\xi = \bar{w} + k^{-1}\|\bar{u} - x_k\|\xi \in \text{Min}\, F_k(\bar{u}) \text{ and } \bar{w} + k^{-1}\bar{t}\xi \preceq w_n$$

as $n \in \mathbb{N}$, i.e., F_k enjoys the limiting monotonicity property (9.8) at $\bar{u}$.

• If $\{t_n\}$ is *increasing*, the sequence $v_n = w_n - k^{-1}t_n\xi$ is decreasing. The strong limiting monotonicity property of F applied to the sequences $\{(u_n, \bar{w}_n)\}$ and $\{v_n\}$ ensures the existence of $\bar{w} \in \text{Min}\, F(\bar{u})$ with $\bar{w} \preceq v_n$ for all $n \in \mathbb{N}$; the latter means that

$$\bar{w} + k^{-1}t_n\xi \preceq w_n \quad \text{whenever} \quad n \in \mathbb{N}. \tag{9.41}$$

Let us show that in this case we have

$$\bar{w} + k^{-1}\bar{t}\xi \preceq w_n \quad \text{with} \quad \bar{t} = \|\bar{u} - x_k\|, \qquad n \in \mathbb{N}. \tag{9.42}$$

Indeed, it follows from (9.41), the assumed increase of the sequence $\{t_n\}$ and the decrease of the sequence $\{w_n\}$ in (9.40) that

$$\bar{w} + k^{-1}t_n\xi \preceq \bar{w} + k^{-1}t_{n+m}\xi = w_{n+m} \preceq w_n \quad \text{for all} \quad n, m \in \mathbb{N},$$

and thus $\bar{w} + k^{-1}t_{n+m}\xi \preceq w_n$ for every $m \in \mathbb{N}$ while n is fixed. Passing to the limit as $m \to \infty$ gives us (9.42) by the closedness of Θ. Combining it with (9.40) and (9.42) verifies the claimed limiting monotonicity of F_k at $\bar{u}$.

Next fix $k \in \mathbb{N}$ and apply the *subdifferential variational principle* from Theorem 9.12 to the mapping F_k in (9.39) and its strict approximate $\varepsilon\xi$-minimizer (x_k, z_k) with $\varepsilon := k^{-2}$ and $\lambda := k^{-1}$. Taking into account the structure of F_k and the regular subdifferential construction $\widehat{\partial}_\Theta F_k$, we find $(\widetilde{x}_k, \widetilde{z}_k, \widetilde{v}_k, \widetilde{x}_k^*, \widetilde{z}_k^*) \in X \times Z \times Z \times X^* \times Z^*$ satisfying the relationships:

$$\begin{cases} \widetilde{z}_k \in F(\widetilde{x}_k), \widetilde{v}_k = g_k(\widetilde{x}_k), (x_k, \widetilde{z}_k + \widetilde{v}_k) \in \text{gph}\, F_k, \|\widetilde{x}_k - x_k\| \le 1/k, \\ (\widetilde{x}_k^*, -\widetilde{z}_k^*) \in \widehat{N}\big((\widetilde{x}_k, \widetilde{z}_k + \widetilde{v}_k); \text{epi}_\Theta F_k\big), -\widetilde{z}_k^* \in \widehat{N}(0; \Theta), \|\widetilde{z}_k^*\| = 1 \end{cases} \tag{9.43}$$

with $\|\widetilde{x}_k^*\| \le k^{-1}$, $k \in \mathbb{N}$. Consider the Asplund space $X \times Z \times Z$ equipped with the sum norm on the product (and hence by the corresponding maximum on the dual product space) and form the two subsets of $X \times Z \times Z$ by

$$\Omega_1 := \big\{(x, z, v)\big|(x, z) \in \text{epi}_\Theta F\big\}, \; \Omega_2 := \big\{(x, z, v)\big|(x, v) \in \text{gph}\, g_k\big\} \tag{9.44}$$

with g_k taken from (9.39). It is easy to see that $(\widetilde{x}_k, \widetilde{z}_k, \widetilde{v}_k) \in \Omega_1 \cap \Omega_2$ and that both sets Ω_1 and Ω_2 are locally closed around this point by the epiclosedness of F and the Lipschitz continuity of g_k. Observe also that

$$(x, z, v) \in \Omega_1 \cap \Omega_2 \Longrightarrow z \in F(x) + \Theta, \;\; v = g_k(x),$$

and so $(x, z + v) \in \text{epi}_\Theta F_k$. We have from the second line of (9.43) that

$$\limsup_{\substack{(x,z,v)\to(\widetilde{x}_k,\widetilde{z}_k,\widetilde{v}_k)\\(x,z,v)\in\Omega_1\cap\Omega_2}}\frac{\big\langle(\widetilde{x}_k^*,-\widetilde{z}_k^*,-\widetilde{z}_k^*),(x,z,v)-(\widetilde{x}_k,\widetilde{z}_k,\widetilde{v}_k)\big\rangle}{\|(x,z,v)-(\widetilde{x}_k,\widetilde{z}_k,\widetilde{v}_k)\|}$$
$$\le\limsup_{\substack{(x,z)\to(\widetilde{x}_k,\widetilde{z}_k+\widetilde{v}_k)\\(x,z)\in\text{epi}\,F_k}}\frac{\big\langle(\widetilde{x}_k^*,-\widetilde{z}_k^*),(x,z)-(\widetilde{x}_k,\widetilde{z}_k+\widetilde{v}_k)\big\rangle}{\|(x,z)-(\widetilde{x}_k,\widetilde{z}_k+\widetilde{v}_k)\|}\le 0,$$

which readily implies the inclusion

$$(\widetilde{x}_k^*,-\widetilde{z}_k^*,-\widetilde{z}_k^*)\in\widehat{N}\big((\widetilde{x}_k,\widetilde{z}_k,\widetilde{v}_k);\Omega_1\cap\Omega_2\big),\quad k\in\mathbb{N}.$$

Applying to this inclusion the *fuzzy intersection rule* from Exercise 2.42 as a consequence of the *approximate extremal principle* and using the particular structure of the sets Ω_i in (9.44) give us $t\ge 0$, $(x_{ik},z_{ik},v_{ik})\in\Omega_i$ and $(x_{ik}^*,z_{ik}^*,v_{ik}^*)\in X^*\times Z^*\times Z^*$ for $i=1,2$ satisfying the conditions

$$\begin{aligned}&(x_1,z_1)\in\text{epi}_\Theta F,\ v_2=g_k(x_2),\ \|x_1-\widetilde{x}_k\|\le k^{-1},\\&(x_1^*,-z_1^*)\in\widehat{N}\big((x_1,z_1);\text{epi}_\Theta F\big),\ -z_1^*\in N(0;\Theta),\ x_2^*\in\widehat{D}^*g_k(x_2)(z_2^*),\\&\|t\widetilde{x}_k^*-x_1^*-x_2^*\|\le k^{-1},\ \|t\widetilde{z}_k^*-z_1^*\|\le k^{-1},\ \|t\widetilde{z}_k^*-z_2^*\|\le k^{-1},\\&1-k^{-1}\le\max\big\{t,\|(x_2^*,0,z_2^*)\|\big\}\le 1+k^{-1},\end{aligned}\tag{9.45}$$

where we drop the index "k" in the i-sequences above to simplify the notation.

Working with (9.45), we first observe that t must be *nonzero* therein for all $k\in\mathbb{N}$ sufficiently large. Arguing by contradiction, suppose that it is not the case, i.e., $t=0$. Then it follows from the third line of (9.45) that $\|z_2^*\|\le k^{-1}$. Remembering the Lipschitz continuity of g_k with modulus k^{-1} and employing the *coderivative estimate* for Lipschitzian mappings from Exercise 3.41 used above in the proof of Theorem 9.12, we get from the second line of (9.45) that

$$\|x_2^*\|\le k^{-1}\|z_2^*\|\tag{9.46}$$

and therefore $\|x_2^*\|\le k^{-2}$. This contradicts the *nontriviality* condition on $(x_2^*,0,z_2^*)$ in the last line of (9.45) and thus justifies that $t>0$.

To proceed further, we consider the following two possibilities of realizing the *maximum* of the expression $\big\{t,\|(x_2^*,0,z_2^*)\|\big\}$ in (9.45):

Case 1. If $\max\{t,\|(x_2^*,0,z_2^*)\|\}=t$, then the last line in (9.45) becomes $1-k^{-1}\le t\le 1+k^{-1}$. Substituting the upper and lower bounds of t from the above into the inequalities in the third line of (9.45) and taking into account the triangle inequality, estimate (9.46), and that $\|\widetilde{z}_k^*\|=1$ while $\|\widetilde{x}_k^*\|\le k^{-1}$ in (9.43), we arrive at the expressions

$$1-2k^{-1}\le\|z_i^*\|\le 1+2k^{-1}\ \text{ for }\ i=1,2\quad\text{and hence}$$

$$\frac{\|x_1^*\|}{\|z_1^*\|} \le \frac{\big(t\|\widetilde{x}_k^*\| + \|x_2^*\| + k^{-1}\big)}{\|z_1^*\|} \le \frac{\big((1+k^{-1})k^{-1} + k^{-1}(1+2k^{-1}) + k^{-1}\big)}{(1-2k^{-1})}$$
$$= \frac{3k^{-1}+3k^{-2}}{1-2k^{-1}}.$$

Case 2. Assuming next that

$$\max\big\{t,\ \|(x_2^*,0,z_2^*)\|\big\} = \|(x_2^*,0,z_2^*)\|$$

and taking into account that $\|(x_2^*,0,z_2^*)\| = \|z_2^*\|$ by (9.46) and the dual norm form on $X^*\times Z^*\times Z^*$, we get from the last line of (9.45) that

$$1-k^{-1} \le \|z_2^*\| \le 1+k^{-1}.$$

Substituting this into $\|t\widetilde{z}_k^* - z_2^*\| \le k^{-1}$ from (9.45) and using $\|\widetilde{z}_k^*\| = 1$ from (9.43), we obtain the lower and upper estimates for t:

$$t \ge \|z_2^*\| - k^{-1} \ge 1-2k^{-1} \quad\text{and}\quad t \le \|z_2^*\| + k^{-1} \le 1+2k^{-1}. \tag{9.47}$$

Then the third line of (9.45) and the lower estimate of t in (9.47) yield

$$\|z_1^*\| \ge t - k^{-1} \ge 1-3k^{-1}. \tag{9.48}$$

Let us finally estimate the ratio $\|x_1^*\|/\|z_1^*\|$ in this case. Employing the inequality $\|t\widetilde{x}_k^* - x_1^* - x_2^*\| \le k^{-1}$ from the third line of (9.45) together with $\|\widetilde{x}_k^*\| \le k^{-1}$ from (9.43) and (9.46), the upper bound for t in (9.47), and the lower bound of $\|z_1^*\|$ in (9.48), we get

$$\frac{\|x_1^*\|}{\|z_1^*\|} \le \frac{\big(t\|\widetilde{x}_k^*\| + \|x_2^*\| + k^{-1}\big)}{\|z_1^*\|} \le \frac{\big((1+2k^{-1})k^{-1} + k^{-1}(1+k^{-1}) + k^{-1}\big)}{1-3k^{-1}}$$
$$= \frac{3k^{-1}(1+k^{-1})}{1-3k^{-1}},$$

which ends our considerations in Case 2. Thus in both Case 1 and Case 2, we have similar (while different) estimates of the ratio $\|x_1^*\|/\|z_1^*\|$.

Continuing now the proof of the theorem simultaneously for both of the above cases of realizing the maximum in the last line of (9.45), denote

$$\widetilde{x}_1^* := \frac{x_1^*}{\|z_1^*\|} \quad\text{and}\quad \widetilde{z}_1^* := \frac{z_1^*}{\|z_1^*\|} \quad\text{with}\quad \|\widetilde{z}_1^*\| = 1 \tag{9.49}$$

and, by the first two lines in (9.45) concerning (x_1, z_1, x_1^*, z_1^*) and the regular subdifferential construction for F, obtain the inclusions

$$\widetilde{x}_1^* \in \widehat{\partial}_\Theta F(x_1,z_1) \quad\text{with}\quad (x_1,z_1)\in \operatorname{epi}_\Theta F. \tag{9.50}$$

Let us show that we can improve (9.50) by using *graph* vs. epigraph points, i.e., replacing $(x_1, z_1) \in \text{epi}_\Theta F$ by some $(x_1, \widetilde{z}_1) \in \text{gph}\, F$ to get

$$\widetilde{x}_1^* \in \widehat{\partial}_\Theta F(x_1, \widetilde{z}_1) \;\text{ with }\; (x_1, \widetilde{z}_1) \in \text{gph}\, F, \tag{9.51}$$

which is needed for the subsequent applications of the subdifferential Palais-Smale condition. To verify (9.51), rewrite (9.50) as

$$(\widetilde{x}_1^*, -\widetilde{z}_1^*) \in \widehat{N}\big((x_1, z_1); \text{epi}_\Theta F\big) \text{ with } -\widetilde{z}_1^* \in N(0; \Theta),\;\; \|\widetilde{z}_1^*\| = 1$$

and for any $\gamma > 0$ find from the definition of regular normals such $\eta > 0$ that

$$\big\langle(\widetilde{x}_1^*, -\widetilde{z}_1^*), (x, z) - (x_1, z_1)\big\rangle \le \gamma\|(x, z) - (x_1, z_1)\| \tag{9.52}$$

whenever $(x, z) \in \text{epi}_\Theta F$ with $x \in x_1 + \eta\mathbb{B}$ and $z \in z_1 + \eta\mathbb{B}$. Observing that the second inclusion in (9.50) reads as

$$z_1 = \widetilde{z}_1 + \theta \;\text{ for some }\; \widetilde{z}_1 \in F(x_1) \;\text{ and } \theta \in \Theta$$

and picking an arbitrary vector $(u, v) \in \text{epi}_\Theta F$ with $u \in x_1 + \eta\mathbb{B}$ and $v \in \widetilde{z}_1 + \eta\mathbb{B}$ with $\widetilde{v} := v + \theta$, we get $v - \widetilde{z}_1 = \widetilde{v} - z_1$ and $(u, \widetilde{v}) \in \text{epi}_\Theta F$ with $u \in x_1 + \eta\mathbb{B}$ and $\widetilde{v} \in z_1 + \eta\mathbb{B}$. It follows then from (9.52) that

$$\begin{aligned}\big\langle(\widetilde{x}_1^*, -\widetilde{z}_1^*), (u, v) - (x_1, \widetilde{z}_1)\big\rangle &= \big\langle(\widetilde{x}_1^*, -\widetilde{z}_1^*), (u, \widetilde{v}) - (x_1, z_1)\big\rangle \\ &\le \gamma\|(u, \widetilde{v}) - (x_1, z_1)\| = \gamma\|(u, v) - (x_1, \widetilde{z}_1)\|,\end{aligned}$$

which yields $(\widetilde{x}_1^*, -\widetilde{z}_1^*) \in \widehat{N}\big((x_1, \widetilde{z}_1); \text{epi}_\Theta F\big)$ with $(x_1, \widetilde{z}_1) \in \text{gph}\, F$. Taking into account that $-\widetilde{z}_1^* \in N(0; \Theta)$ with $\|\widetilde{z}_1^*\| = 1$, we arrive at (9.51) by the regular subdifferential construction for set-valued mappings.

Now add the index "k" to indicate the sequences $(x_{1k}, \widetilde{z}_{1k})$ and $(\widetilde{x}_{1k}^*, \widetilde{z}_{1k}^*)$, $k \in \mathbb{N}$, defined in (9.49) and (9.51). Using estimates (9.47) and (9.51) yields

$$(x_{1k}, \widetilde{z}_{1k}) \in \text{gph}\, F \text{ and } \widetilde{x}_{1k}^* \in \widehat{\partial}_\Theta F(x_{1k}, \widetilde{z}_{1k}) \text{ with } \|\widetilde{x}_{1k}^*\| \to 0 \tag{9.53}$$

as $k \to \infty$. Employing finally the *subdifferential Palais-Smale condition* of Definition 9.13(i), deduce from (9.53), the sequence $\{x_{1k}\}$ contains a convergent subsequence. Since it follows from (9.43) and (9.45) that

$$\|x_k - x_{1k}\| \le \|x_k - \widetilde{x}_k\| + \|\widetilde{x}_k - x_{1k}\| \le k^{-1} + k^{-1} \;\text{ for all }\; k \in \mathbb{N},$$

we conclude that the sequence $\{x_k\}$ constructed in (9.33) and (9.34) also contains a convergent subsequence and thus completes the proof. △

Next we present efficient consequences of Theorem 9.15 ensuring the existence of other types of relative minimizers from Definition 9.4 as well as weak minimizers for ordered set-valued mappings.

Corollary 9.16 (Existence of Primary Relative and Quasi-Relative Minimizers). *Suppose in addition to the assumptions of Theorem 9.15 that* $\text{ri}\,\Theta \neq \emptyset$. *Then there exist a relative minimizer and a quasi-relative minimizer for the ordered set-valued mapping* $F\colon X \rightrightarrows Z$ *under consideration.*

Proof. As mentioned above, all the relative minimizers in Definition 9.3 agree in this case, and so the claimed existence follows from Theorem 9.15. △

Corollary 9.17 (Existence of Weak Pareto Minimizers). *Suppose in addition to the assumptions of Theorem 9.15 that* $\text{int}\,\Theta \neq \emptyset$. *Then there exists a weak Pareto minimizer for the mapping* $F\colon X \rightrightarrows Z$ *under consideration.*

Proof. Theorem 9.15 guarantees the existence of an intrinsic relative minimizer for F provided that $\emptyset \neq \text{int}\,\Theta \subset \text{iri}\,\Theta$, which is surely a weak Pareto minimizer for the mapping F in this case. △

Note that outside of the case where $\text{ri}\,\Theta \neq \emptyset$ as in Corollary 9.16, we don't have conditions ensuring the existence of *quasi-relative* minimizers for ordered set-valued mappings, which is a challenging *open question*.

9.3.3 Existence Theorems Under Explicit Constraints

Let us now investigate the existence issues for multiobjective problems with *explicit geometric constraints* given by

$$\text{minimize}\quad F(x)\quad \text{subject to}\quad x \in \Omega \subset X. \tag{9.54}$$

Defining the *restriction* of F on Ω by $F_\Omega(x) := F(x)$ if $x \in \Omega$ and $F(x) := \emptyset$ otherwise, we can rewrite (9.54) in the unconstrained format

$$\text{minimize}\quad F_\Omega(x) = F(x) + \Delta(x;\Omega),\qquad x \in X, \tag{9.55}$$

via the indicator mapping $\Delta(x;\Omega) := 0$ if $x \in \Omega$ and $\Delta(x;\Omega) := \emptyset$ if $x \notin \Omega$.

To apply the existence results from Theorem 9.15 and its consequences to the unconstrained format (9.55) and express the obtained conditions in terms of the initial data of the constrained problem (9.54), we need to deal effectively with the sum mapping in (9.55). Since a major ingredient of Theorem 9.15 is the *subdifferential Palais-Smale condition*, subdifferential sum rules are required to proceed in this way for problems with geometric constraints, while further rules of subdifferential calculus are needed to treat other types of constraints describing Ω in some structural form (via inequalities, equalities, operator constraints, equilibrium constraints, etc.). From this viewpoint, the *basic* version of the Palais-Smale condition from Definition 9.13(ii) is more convenient and easier to deal with than its regular counterpart from part (i) of that definition, although the latter one is generally less restrictive. The reason is that our basic generalized differential constructions (normals, subgradients, coderivatives) satisfy comprehensive pointbased *calculus rules* in contrast to the regular ones. Due to this, it is more standard to elaborate the basic Palais-Smale condition in constrained multiobjective optimization, and we leave it as exercises to the reader; see them below with some hints.

Our aim here is to deal with the more challenging *regular* subdifferential Palais-Smale condition in Theorem 9.15 by using some specific results of the regular subdifferential calculus, which allow us to treat effectively problems with geometric constraints. For definiteness, we confine ourselves to the case of single-valued objectives $f\colon X \to Z$ in multiobjective problems of type (9.54).

Given a single-valued mapping $f\colon X \to Z$ between Banach spaces with the ordering cone Θ of Z, observe directly from Definition 9.6(i) that its regular subdifferential can be represented as

$$\widehat{\partial}_\Theta f(\bar{x}) = \bigcup_{\substack{-z^*\in N(0;\Theta)\\ \|z^*\|=1}} \widehat{\partial}_\Theta f(\bar{x})(z^*) \tag{9.56}$$

with $\widehat{\partial}_\Theta f(\bar{x})(z^*) := \widehat{D}^*\mathcal{E}_{F,\Theta}(\bar{x}, f(\bar{x}))(z^*)$. It is not hard to check the following special sum rule for the regular subdifferential:

$$\widehat{\partial}_\Theta(f+\Delta)(\bar{x})(z^*) \subset \bigcap_{v\in\widehat{\partial}_\Theta(-f)(\bar{x})(z^*)} \big[\widehat{N}(\bar{x};\Omega) - v\big] \tag{9.57}$$

provided that $\widehat{\partial}_\Theta(-f)(\bar{x})(z^*) \neq \emptyset$ and that there is a neighborhood U of $\bar{x}$ as well as nonnegative numbers ℓ and γ such that

$$\|f(u) - f(\bar{x})\| \le \ell\|u - \bar{x}\| + \gamma\big|\langle z^*, f(u) - f(\bar{x})\rangle\big| \text{ for all } u \in U. \tag{9.58}$$

Note that condition (9.58) automatically holds if either $Z = \mathbb{R}$ or the mapping f is *upper Lipschitzian* at $\bar{x}$, i.e., $\gamma = 0$ in (9.58) and is surely fulfilled when f is locally Lipschitzian around this points.

The next theorem ensures the existence of *intrinsic relative minimizers* for the constrained problem (9.54) as well as the existence of relative Pareto and weak Pareto minimizers for (9.54) under additional assumptions.

Theorem 9.18 (Existence of Relative and Weak Pareto Minimizers for Constrained Multiobjective Problems). *Let $f\colon X \to Z$ and $\Theta \subset Z$ satisfy the general assumptions of Theorem 9.15, and let $\Omega \subset X$ be closed. Suppose in addition that $\widehat{\partial}_\Theta(-f)(x)(z^*) \neq \emptyset$, (9.58) holds for any $x \in \Omega$ and $z^* \in -N(0;\Theta)$ with $\|z^*\| = 1$ and that every sequence $\{x_k\} \subset \Omega$ with*

$$\exists\, x_k^* \in \bigcap_{v\in\widehat{\partial}_\Theta(-f)(x_k)(z_k^*)} \Big[\widehat{N}(x_k;\Omega) - v\Big] \text{ such that } -z_k^* \in N(0;\Theta), \tag{9.59}$$

$\|z_k^*\| = 1$, and $\|x_k^*\| \to 0$ as $k \to \infty$ contains a convergent subsequence. Then problem (9.54) admits an intrinsic relative minimizer if $\operatorname{iri}\Theta \neq \emptyset$. *Furthermore,*

this problem admits a primary relative Pareto minimizer if $\operatorname{ri}\Theta \neq \emptyset$ *and also a weak Pareto minimizer if* $\operatorname{int}\Theta \neq \emptyset$.

Proof. Considering the unconstrained format (9.55) of problem (9.54), it is easy to see that the restriction mapping f_Ω satisfies all the assumptions of Theorem 9.15 except the regular subdifferential Palais-Smale condition, which should be verified. To do it, take sequences $\{x_k\}$, $\{x_k^*\}$ from Definition 9.13 for $F = f_\Omega$ and by (9.56) find $\{z_k^*\}$ such that

$$x_k^* \in \widehat{\partial}_\Theta\big[f + \Delta(\cdot;\Omega)\big](x_k)(z_k^*),\ \ -z_k^* \in N(0;\Theta),\ \ \|z_k^*\| = 1 \tag{9.60}$$

with $\|x_k^*\| \to 0$. This yields $\{x_k\} \subset \Omega$. Employing further the subdifferential sum rule (9.57) in (9.60) tells us that

$$x_k^* \in \bigcap_{v \in \widehat{\partial}_\Theta(-f)(x_k)(z_k^*)} \big[\widehat{N}(x_k;\Omega) - v\big] \text{ with } \|x_k^*\| \to 0 \text{ as } k \to \infty,$$

where $-z_k^* \in N(0;\Theta)$ and $\|z_k^*\| = 1$, i.e., the triple $\{x_k, x_k^*, z_k^*\}$ satisfies (9.59). Thus the sequence $\{x_k\} \subset \Omega$ contains a convergent subsequence, which verifies the claimed Palais-Smale condition for f_Ω and hence ensures the existence of intrinsic relative minimizers for (9.54) provided that $\operatorname{iri}\Theta \neq \emptyset$. The existence of primary relative Pareto minimizers and weak Pareto minimizers for (9.54) provided that $\operatorname{ri}\Theta \neq \emptyset$ and $\operatorname{int}\Theta \neq \emptyset$, respectively, is justified similarly to the proofs of Corollaries 9.16 and 9.17. △

The major assumptions of Theorem 9.18 are automatically fulfilled and/or significantly simplified if the cost mapping f is Fréchet differentiable on Ω.

Corollary 9.19 (Existence of Relative and Weak Pareto Minimizers for Constrained Problems with Fréchet Differentiable Objectives). *Let* $f: X \to Z$ *and* $\Theta \subset Z$ *satisfy the general assumptions of Theorem 9.15, let* $\Omega \subset X$ *be closed, and let* f *be Fréchet differentiable on* Ω. *Assume also that every sequence* $\{x_k\} \subset \Omega$ *such that*

$$\exists\, x_k^* \in \nabla f(x_k)^* z_k^* + \widehat{N}(x_k;\Omega) \text{ with} - z_k^* \in N(0;\Theta), \tag{9.61}$$

$\|z_k^*\| = 1$, *and* $\|x_k^*\| \to 0$ *as* $k \to \infty$ *contains a convergent subsequence. Then problem (9.54) admits an intrinsic relative minimizer provided that* $\operatorname{iri}\Theta \neq \emptyset$. *Furthermore, (9.54) admits a primary relative Pareto minimizer if* $\operatorname{ri}\Theta \neq \emptyset$ *as well as a weak Pareto minimizer if* $\operatorname{int}\Theta \neq \emptyset$.

Proof. It easily follows from the definitions that

$$\widehat{\partial}_\Theta(-f)(x)(z^*) = \big\{-\nabla f(x)^* z^*\big\} \neq \emptyset \ \text{ whenever } \ -z^* \in N(0;\Theta).$$

Furthermore, we can directly check that the Fréchet differentiability of f implies property (9.58) on Ω. Thus all the assumptions of Theorem 9.18 are satisfied, and condition (9.59) reduces to (9.61) in this setting. △

9.4 Optimality Conditions for Multiobjective Problems

In this section we establish *necessary optimality conditions* for all the types of *local* minimizers for multiobjective problems defined in Section 9.1 in the *Asplund* space setting. These necessary optimality conditions for all the solution types will be defined in the unified way based on the *extremal principle*.

9.4.1 Fermat Rules in Set-Valued Optimization

We begin with the unconstrained problem of minimizing $F\colon X \rightrightarrows Z$, where Z is ordered by the cone Θ satisfying the standing assumptions listed above but which may not be pointed. Our conditions below include the SNC/PSNC properties defined in (2.41) and (3.65), respectively, that automatically hold in finite dimensions. Note that the SNC property of a convex cone $C \subset Z$ at the origin can be equivalently written as

$$\big[z_k^* \stackrel{w^*}{\to} 0,\ z_k^* \in C^+,\ k \in \mathbb{N}\big] \Longrightarrow \|z_k^*\| \to 0 \ \text{ as } \ k \to \infty,$$

where C^+ stands for the positive polar cone to C given by

$$C^+ := \big\{z^* \in Z^* \big|\ \langle z^*, z\rangle \geq 0 \ \text{ for all } \ z \in C\big\}.$$

In what follows we use a remarkable fact that the SNC property of a convex cone $C \subset Z$ with $\operatorname{ri} C \neq \emptyset$ at the origin is equivalent to the *finite codimensionality* of the space $\operatorname{cl}(C - C)$; see Exercise 2.28(iii). Recall that the symbol D^* below signifies the basic/normal coderivative of a set-valued mapping.

Theorem 9.20 (Fermat Rules for Local Solutions to Multiobjective Problems). *Let $F\colon X \rightrightarrows Z$ be a set-valued mapping between Asplund spaces such that its graph is locally closed around the reference point $(\bar{x}, \bar{z}) \in \operatorname{gph} F$, while the image space Z is partially ordered by a closed, convex, and proper cone $\Theta \subset Z$. Then the (Fermat type)* CODERIVATIVE *condition*

$$0 \in D^*F(\bar{x}, \bar{z})(z^*) \text{ with some } -z^* \in N(0; \Theta),\ \ \|z^*\| = 1 \tag{9.62}$$

is necessary for optimality of $(\bar{x}, \bar{z})$ to F in each of the following cases:

- $(\bar{x}, \bar{z})$ *is a local* PARETO MINIMIZER/EFFICIENT SOLUTION *provided that $\Theta \setminus (-\Theta) \neq \emptyset$ and that either Θ is SNC at $0 \in Z$ or F^{-1} is PSNC at $(\bar{z}, \bar{x})$.*

- $(\bar{x}, \bar{z})$ *is a local* QUASI-RELATIVE MINIMIZER *provided that either Θ is SNC at $0 \in Z$ or F^{-1} is PSNC at $(\bar{z}, \bar{x})$.*

- $(\bar{x}, \bar{z})$ *is a local* INTRINSIC RELATIVE MINIMIZER *provided that either Θ is SNC at $0 \in Z$ or F^{-1} is PSNC at $(\bar{z}, \bar{x})$.*

- $(\bar{x}, \bar{z})$ *is a local* PRIMARY RELATIVE MINIMIZER *provided that either the closed* *subspace* $\mathrm{cl}(\Theta - \Theta)$ *is finite-codimensional in* Z *or* F^{-1} *is PSNC at* $(\bar{z}, \bar{x})$.

• $(\bar{x}, \bar{z})$ *is a local* WEAK PARETO MINIMIZER.

Furthermore, we have the SUBDIFFERENTIAL *necessary optimality condition*

$$0 \in \partial_\Theta F(\bar{x}, \bar{z}) \tag{9.63}$$

in each of the listed cases of (*efficient, quasi-relative, intrinsic relative, primary relative, weak*) *local minimizers* $(\bar{x}, \bar{z})$ *provided that the epigraph vs. graph of* F *is closed around* $(\bar{x}, \bar{z})$ *and that the PSNC property of* F^{-1} *at* $(\bar{z}, \bar{x})$ *in the assumptions above is replaced by the PSNC property of the inverse mapping* $\mathcal{E}_{F,\Theta}^{-1}$ *to the associated epigraphical multifunction at this point.*

Proof. Arguing in the unified way, take any local minimizer $(\bar{x}, \bar{z}) \in \mathrm{gph}\, F$ for F considered in theorem and reduce it to a *local extremal point* of some system of sets in the (Asplund) product space $X \times Z$. Namely, define the sets

$$\Omega_1 := \mathrm{gph}\, F, \quad \Omega_2 := X \times (\bar{z} - \Theta), \tag{9.64}$$

which are locally closed around $(\bar{x}, \bar{z})$ due to the closedness assumptions imposed on F and Θ. We obviously have $(\bar{x}, \bar{z}) \in \Omega_1 \cap \Omega_2$. To verify the local extremality of $(\bar{x}, \bar{z})$ for $\{\Omega_1, \Omega_2\}$, let us show that there is a sequence $\{c_k\} \subset Z$ with $c_k \to 0$ as $k \to \infty$ such that

$$\Omega_1 \cap \big(\Omega_2 + (0, c_k)\big) \cap (U \times Z) = \emptyset, \quad k \in \mathbb{N}, \tag{9.65}$$

where U is a neighborhood of $\bar{x}$ from its local minimality property. This gives us the required extremality relation (2.1) with $a_k := (0, c_k) \in X \times Z$.

We construct an appropriate sequence $\{c_k\} \subset Z$ in (9.65) by putting $c_k := c/k$ as $k \in \mathbb{N}$, where $0 \neq c \in Z$ is selected in the following way for each type of the local minimizers considered in the theorem. This can be done by using the definitions of the corresponding minimizers with taking into account the additional assumption $\Theta \setminus (-\Theta) \neq \emptyset$ imposed in the case of (Pareto) efficient solutions:

• $c \in -\big(\Theta \setminus (-\Theta)\big)$ if $(\bar{x}, \bar{z})$ is a local *Pareto minimizer*;

• $c \in -\mathrm{qri}\, \Theta$ if $(\bar{x}, \bar{z})$ is a local *quasi-relative minimizer*;

• $c \in -\mathrm{iri}\, \Theta$ if $(\bar{x}, \bar{z})$ is a local *intrinsic relative minimizer*;

• $c \in -\mathrm{ri}\, \Theta$ if $(\bar{x}, \bar{z})$ is a local *primary relative minimizer*;

• $c \in -\mathrm{int}\, \Theta$ if $(\bar{x}, \bar{z})$ is a local *weak Pareto minimizer*.

Arguing by contradiction, suppose that (9.65) doesn't hold, i.e.,

$$\text{there is } (x, z) \in U \times Z \text{ with } (x, z) \in \Omega_1 \cap \big(\Omega_2 + (0, c_k)\big). \tag{9.66}$$

Then by the construction of sets (9.64), we find some $(x, z) \in X \times Z$ such that

$$x \in U, \ z \in F(x), \ \text{ and } \ z \in \bar{z} - \Theta + c_k, \quad k \in \mathbb{N}. \tag{9.67}$$

In the case of Pareto minimizers, the latter tells us by the choice of $\{c_k\}$ that

$$\bar z-\Theta+c_k\subset\bar z-\Theta-(\Theta\setminus(-\Theta))\subset\bar z-(\Theta\setminus\{0\})\tag{9.68}$$

as $k\in\mathbb{N}$. In all the cases of the relative minimizers, as well as for weak efficient solutions to F, we have by the choice of $\{c_k\}$ that

$$\bar z-\Theta+c_k\subset\bar z-\Theta-\widetilde\Theta=\bar z-\widetilde\Theta,\quad k\in N,\tag{9.69}$$

where $\widetilde\Theta$ stands for either $\operatorname{qri}\Theta$, or $\operatorname{iri}\Theta$, or $\operatorname{ri}\Theta$, or $\operatorname{int}\Theta$ in the corresponding cases of local minimizers; see Exercise 9.44 for relative minimizers while observing that it is obvious for weak minimal ones. Combining the relationships in (9.66)–(9.69), we get $z\in(\bar z-\widetilde\Theta)\cap F(U)$ for relative and weak minimizers and $z\in(\bar z-(\Theta\setminus\{0\}))\cap F(U)$ for local efficient solutions to F. It surely contradicts the definitions of these minimizers and thus justifies by (9.65) the local extremality of $(\bar x,\bar z)$ for $\{\Omega_1,\Omega_2\}$ in all the cases considered.

Equip now the space $X\times Z$ with the usual sum norm $\|(x,z)\|:=\|x\|+\|z\|$. Then applying the (approximate) extremal principle to the system of closed sets $\{\Omega_1,\Omega_2\}$ in (9.64) from the Asplund space and taking into account the particular structures of Ω_1,Ω_2 as well as the maximum dual norm on $X^*\times Z^*$, for any sequence $\varepsilon_k\downarrow 0$, we find $\{(x_{ik},z_{ik})\}\subset X\times Z$ and $\{(x^*_{ik},z^*_{ik})\}\subset X^*\times Z^*$, $i=1,2$, satisfying the relationships:

$$(x_{1k},z_{1k})\in\operatorname{gph}F,\ \ (x_{2k},z_{2k})\in X\times(\bar z-\Theta),\ \ \|(x_{ik},z_{ik})-(\bar x,\bar z)\|\le\varepsilon_k,$$

$$(x^*_{1k},-z^*_{1k})\in\widehat N\big((x_{1k},z_{1k});\operatorname{gph}F\big),\ \ 0=x^*_{2k}\in\widehat N(x_{2k};X),\ \ z^*_{2k}\in\widehat N\big(\bar z-z_{2k};\Theta\big),$$

$$\begin{cases}\max\big\{\|x^*_{1k}\|,\|z^*_{1k}+z^*_{2k}\|\big\}\le\varepsilon_k,\\[1ex] 1-\varepsilon_k\le\max\big\{\|x^*_{1k}\|,\|z^*_{1k}\|\big\}+\|z^*_{2k}\|\le 1+\varepsilon_k.\end{cases}\tag{9.70}$$

It follows from the second condition in (9.70) that the sequences $\{(x^*_{ik},z^*_{ik})\}$, $i=1,2$, are bounded in the (dual to Asplund) space $X^*\times Z^*$, and hence they contain weak* converging subsequences. Using the first condition in (9.70), we get without loss of generality that

$$\|x^*_{1k}\|\to 0,\ \ z^*_{1k}\xrightarrow{w^*}z^*,\ \text{ and } z^*_{2k}\xrightarrow{w^*}-z^*\ \text{as } k\to\infty,\tag{9.71}$$

where the weak* limit $z^*\in Z^*$ satisfies the inclusions

$$(0,-z^*)\in N\big((\bar x,\bar z);\operatorname{gph}F\big)\ \text{ and }\ -z^*\in N\big(0;\Theta\big)\tag{9.72}$$

obtained by passing to the limit in the relationships above (9.70) as $k\to\infty$.

Next we show that $z^* \neq 0$ in (9.72) if either Θ is SNC at the origin or F^{-1} is PSNC at $(\bar{z}, \bar{x})$ for all the types of the local minimizers under consideration. Assume by the contrary that $z^* = 0$ and deduce then from (9.71) that

$$z_{1k}^* \xrightarrow{w^*} 0 \text{ and } z_{2k}^* \xrightarrow{w^*} 0 \text{ as } k \to \infty. \tag{9.73}$$

If Θ is SNC at the origin, then the second expression in (9.73) yields $\|z_{2k}^*\| \to 0$ and therefore $\|z_{1k}^*\| \to 0$ as $k \to \infty$ by the first relationship in (9.70). Combining the latter with (9.71), we thus contradict the nontriviality/second expression in (9.70). Suppose now that F^{-1} is PSNC at $(\bar{z}, \bar{x})$. Using the convergence of regular normals in (9.71), we conclude from the imposed PSNC property that $\|z_{1k}^*\| \to 0$ as $k \to \infty$. This gives us $\|z_{2k}^*\| \to 0$ as $k \to \infty$ and also contradicts the second expression in (9.70). Hence $z^* \neq 0$ in (9.72), which yields the coderivative condition (9.62) by normalization and the coderivative definition. Thus we arrive at the conclusions of the theorem regarding the coderivative necessary condition (9.62) for the cases of Pareto minimizers, quasi-relative minimizers, and intrinsic relative minimizers.

The case of primary relative minimizers requires the assumption $\operatorname{ri} \Theta \neq \emptyset$. The latter allows us to fully characterize the SNC property of Θ. Indeed, the aforementioned result of Exercise 2.28(iii) tells us that this property is equivalent to the imposed assumption that the subspace $\operatorname{cl}(\Theta - \Theta)$ is finite-codimensional in Z, and thus we are done in this case. If finally $(\bar{x}, \bar{z})$ is a weak Pareto minimizer for F, then $\operatorname{int} \Theta \neq \emptyset$. In this case the convex ordering cone Θ is automatically SNC; see Exercise 2.29, and thus the coderivative result (9.62) *unconditionally* holds for weak Pareto minimizers.

It remains to justify the *subdifferential necessary condition* (9.63) for all the local minimizers under consideration. Using the epigraphical multifunction $\mathcal{E}_{F,\Theta}\colon X \rightrightarrows Z$ associated with F, define the set-valued optimization problem:

$$\text{minimize } \mathcal{E}_{F,\Theta}(x) = F(x) + \Theta, \quad x \in X. \tag{9.74}$$

It is clear that every local optimal solution to (9.74) in each of the aforementioned senses is a local optimal solution in the corresponding sense to the mapping F. For our purposes we need to verify the opposite implication. Let us first show that it holds for all the relative and weak Pareto local minimizers. The latter follows from the fact that the corresponding localized minimality notions for F from Definition 9.4 yield

$$\big(\bar{z} - \widetilde{\Theta}\big) \cap \big(F(U) + \Theta\big) = \emptyset \tag{9.75}$$

for (9.74), where $\widetilde{\Theta}$ stands, respectively, for each of $\operatorname{qri} \Theta$, $\operatorname{iri} \Theta$, $\operatorname{ri} \Theta$, and $\operatorname{int} \Theta$. Indeed, the negation of (9.75) tells us that $z \in (\bar{z} - \widetilde{\Theta}) \cap (F(U) + \Theta)$, and hence

$$\text{there are } u \in U, \quad v \in F(u), \text{ and } \theta \in \Theta \text{ such that } z = v + \theta \in \bar{z} - \widetilde{\Theta}.$$

This gives us the relationships

$$v = z - \theta \in \bar{z} - \theta - \widetilde{\Theta} \subset \bar{z} - \Theta - \widetilde{\Theta} = \bar{z} - \widetilde{\Theta}$$

for all the cases of $\widetilde{\Theta}$ under consideration, where the latter inclusion is trivial for $\widetilde{\Theta} = \text{int}\,\Theta$ while follows from Exercise 9.44 for the cases of relative minimizers. Hence we get $v \in (\bar{z} - \widetilde{\Theta}) \cap F(U)$, which contradicts the localized minimality relationships in Definition 9.4 and so verifies the claim. Applying now the coderivative condition (9.62) to (9.74) and using the basic subdifferential definition, justify the subdifferential optimality condition (9.63) for weak and all the relative minimizers.

To complete the proof of the theorem, we need to justify the subdifferential condition (9.63) for the case of Pareto/efficient local minimizers under the general assumptions made, which don't include the pointedness of Θ. Let us proceed similarly to the proof of the first/coderivative part of the theorem with the replacement of $\Omega_1 = \text{gph}\,F$ by $\widetilde{\Omega}_1 := \text{epi}_\Theta F$. It is sufficient to verify the extremality property (9.65) with the set $\widetilde{\Omega}_1$ therein. Arguing by contradiction, suppose that the latter doesn't hold, i.e.,

$$\text{there is } (x, z) \in U \times Z \ \text{ such that } \ (x, z) \in \widetilde{\Omega}_1 \cap \big(\Omega_2 + (0, c_k)\big). \tag{9.76}$$

Then by the constructions of the sets $\widetilde{\Omega}_1$ and Ω_2 and the definition of the vector epigraphs, we find from (9.76) some $(x, z, \theta) \in X \times Z \times \Theta$ satisfying

$$x \in U, \ z \in F(x) + \theta, \ \text{ and } \ z \in \bar{z} - \Theta + c_k, \quad k \in \mathbb{N}.$$

This implies, by the convexity property of the ordering cone Θ, that

$$x \in U, \ z - \theta \in F(U), \ \text{ and } \ z - \theta \in \bar{z} - \theta - \Theta + c_k \subset \bar{z} - \Theta + c_k \ \text{ as } \ k \in \mathbb{N}.$$

Similarly to (9.68) we get from the latter that

$$z - \theta \in \big(\bar{z} - (\Theta \setminus \{0\})\big) \cap F(U),$$

which obviously contradicts the local Pareto minimality and thus verifies that $(\bar{x}, \bar{z})$ is a local extremal point of the set system $\{\widetilde{\Omega}_1, \Omega_2\}$. Using finally the same arguments as in the above proof of the coderivative optimality condition (9.62), we arrive at the subdifferential optimality condition (9.63) for Pareto minimizers and thus complete the proof of the theorem. △

The following remark with the example therein reveals a significant difference between scalar minimization of extended-real-valued functions and multiobjective optimization of our study.

Remark 9.21 (Failure of Fermat Rule for Multiobjective Problems via Regular Subgradients). The scalar counterpart of Theorem 9.20 for local minimizers of $\varphi\colon X \to \overline{\mathbb{R}}$ is $0 \in \partial\varphi(\bar{x})$, which follows from the more selective regular subdifferential Fermat rule $0 \in \widehat{\partial}\varphi(\bar{x})$ in general Banach space; see Proposition 1.30(i). It is interesting to observe that multiobjective counterpart of the latter necessary optimality condition *fails* even for simple mappings between finite-dimensional spaces. To illustrate this, consider a set-valued mapping $F\colon \mathbb{R} \rightrightarrows \mathbb{R}^2$ defined by

$$F(x) \equiv \Xi \text{ with } \Xi := \big\{z = (z_1, z_2) \in \mathbb{R}^2 \big| \text{ either } 2z_1 + z_2 \geq 0 \text{ or } z_1 + 2z_2 \geq 0\big\}.$$

We see that $(\bar{x}, \bar{z}) = (0, 0) \in \mathbb{R} \times \mathbb{R}^2$ is a Pareto minimizer for F with

$$\widehat{N}\big((0, 0); \mathbb{R} \times (\Xi + \mathbb{R}^2_+)\big) = \widehat{N}\big((0, 0); \mathbb{R} \times \Xi\big) = \widehat{N}(0, \mathbb{R}) \times \widehat{N}(0; \Xi) = \emptyset.$$

Thus $\widehat{\partial}_\Theta F(0, 0) = \emptyset$, which demonstrates the failure of the multiobjective version of the regular subdifferential Fermat rule.

9.4.2 Optimality Conditions in Constrained Settings

In the final subsection of this section (and of the whole chapter), we return to the multiobjective optimization problem (9.54) with explicit geometric constraints and derive necessary optimality conditions for all the types of (local) minimizers from Definition 9.4 applied to the equivalent unconstrained format (9.55) written via $F_\Omega(x) = F(x) + \Delta(x; \Omega)$. The results obtained are expressed in terms of our basic generalized differentiable constructions and associated SNC/PSNC properties, which both enjoy *full calculi* in the framework of Asplund spaces; see [522, Chapter 3] and exercises for Chapter 3 above. Note that in this way we can treat multiobjective problems with other types of structural constraints (functional, operator, equilibrium, etc.) while leaving this as exercises for the reader formulated at the end of this chapter.

Along with the basic/normal coderivative D^*, in what follows we use also the *mixed coderivative* construction D^*_M for set-valued mappings, which is defined in (1.65) and also enjoys the same kind of pointbased full calculus as its normal counterpart; see [522, Chapter 3] and the corresponding commentaries and exercises in Chapter 3 above. The mixed coderivative is employed below for formulating refined qualification conditions and their specifications for remarkable classes of mappings under consideration.

Using the mixed coderivative, we define the *singular subdifferential* of a mapping $F\colon X \rightrightarrows Z$ with values in a partially ordered space Z by

$$\partial^\infty_\Theta F(\bar{x}, \bar{z}) := D^*_M \mathcal{E}_{F,\Theta}(\bar{x}, \bar{z})(0) \text{ at } (\bar{x}, \bar{z}) \in \text{epi}_\Theta F. \tag{9.77}$$

The next theorem presents necessary optimality conditions for local minimizers of all the types in Definition 9.4 for constrained problems (9.54).

Theorem 9.22 (Necessary Conditions for Relative Pareto Minimizers of Constrained Multiobjective Problems). *Let $F\colon X \rightrightarrows Z$ be a mapping between Asplund spaces with the image space Z partially ordered by a closed, convex, and* *proper cone Θ. Suppose that $\Omega \subset X$ is locally closed around the reference local minimizer $(\bar{x}, \bar{z})$ for (9.54). The following assertions hold:*

(i) *Assume that the graph of F is locally closed around $(\bar{x}, \bar{z})$, that the mixed qualification condition*

$$D_M^* F(\bar{x}, \bar{z})(0) \cap \big(- N(\bar{x}; \Omega)\big) = \{0\} \tag{9.78}$$

is satisfied and that either F is PSNC at $(\bar{x}, \bar{z})$ or Ω is SNC at $\bar{x}$; both (9.78) and the PSNC property hold automatically if F is Lipschitz-like around $(\bar{x}, \bar{z})$. Then there exists $-z^* \in N(0; \Theta)$ with $\|z^*\| = 1$ such that

$$0 \in D^* F(\bar{x}, \bar{z})(z^*) + N(\bar{x}; \Omega) \tag{9.79}$$

in each of the following cases of local minimizers for (9.54);

- $(\bar{x}, \bar{z})$ *is a local* PARETO MINIMIZER/EFFICIENT SOLUTION *provided that $\Theta \setminus (-\Theta) \neq \emptyset$ and that either Θ is SNC at $0 \in Z$ or F_Ω^{-1} is PSNC at $(\bar{z}, \bar{x})$.*
- $(\bar{x}, \bar{z})$ *is a local* QUASI-RELATIVE MINIMIZER *provided that either Θ is SNC at $0 \in Z$ or F_Ω^{-1} is PSNC at $(\bar{z}, \bar{x})$.*
- $(\bar{x}, \bar{z})$ *is a local* INTRINSIC RELATIVE MINIMIZER *provided that either Θ is SNC at $0 \in Z$ or F_Ω^{-1} is PSNC at $(\bar{z}, \bar{x})$.*
- $(\bar{x}, \bar{z})$ *is a local* PRIMARY RELATIVE MINIMIZER *provided that either the subspace* $\mathrm{cl}(\Theta - \Theta)$ *is finite-codimensional in Z or F_Ω^{-1} is PSNC at $(\bar{z}, \bar{x})$.*
- $(\bar{x}, \bar{z})$ *is a local* WEAK PARETO MINIMIZER.

(ii) *Assume that F is epiclosed around $(\bar{x}, \bar{z})$ and that the singular subdifferential qualification condition*

$$\partial_\Theta^\infty F(\bar{x}, \bar{z}) \cap \big(- N(\bar{x}; \Omega)\big) = \{0\} \tag{9.80}$$

is satisfied. Then we have the subdifferential necessary optimality condition

$$0 \in \partial_\Theta F(\bar{x}, \bar{z}) + N(\bar{x}; \Omega) \tag{9.81}$$

for all the local minimizers considered in assertion (i) *provided that the assumptions on F in* (i) *are replaced by the corresponding assumptions imposed on its epigraphical multifunction $\mathcal{E}_{F,\Theta}$.*

Proof. To verify (i), represent (9.54) in the equivalent multiobjective format (9.55) and apply to the latter Theorem 9.20(i) for all the types of local minimizers. In this way we find $z^* \in -N(0; \Theta)$ with $\|z^*\| = 1$ such that

$$0 \in D^* F_\Omega(\bar{x}, \bar{z})(z^*) = D^*\big(F + \Delta(\cdot; \Omega)\big)(\bar{x}, \bar{z})(z^*). \tag{9.82}$$

Employing in (9.82) the coderivative sum rule from Exercise 3.59(iii) yields

$$D^*\big(F + \Delta(\cdot; \Omega)\big)(\bar{x}, \bar{z})(z^*) \subset D^* F(\bar{x}, \bar{z})(z^*) + N(\bar{x}; \Omega)$$

under (9.78) and the imposed SNC/PSNC requirements. Substituting the latter into (9.82) and taking into account the corresponding assumptions of Theorem 9.20(i), we justify (i) for all the local minimizers under consideration. Note that the fulfillment of the qualification condition (9.78) and the PSNC property claimed in this assertion follows from Theorem 3.3 and the results of [522, Chapter 4] in Asplund spaces discussed in Sections 3.4 and 3.5.

To verify now assertion (ii), we get from Theorem 9.20(ii) that

$$0 \in \partial_\Theta F_\Omega(\bar{x},\bar{z}) \;\text{ and hence }\; 0 \in D^*\mathcal{E}_{F_\Omega,\Theta}(\bar{x},\bar{z})(z^*)$$

with some $z^* \in -N(0;\Theta)$, $\|z^*\| = 1$ for all the types of local minimizers. Applying the aforementioned coderivative sum rule from Exercise 3.59(iii) to

$$\mathcal{E}_{F_\Omega,\Theta}(x) = \mathcal{E}_{F,\Theta}(x) + \Delta(x;\Omega), \qquad x \in X,$$

and taking into account the definitions of the basic and singular subdifferentials of F, we get under the assumptions made in (ii) that

$$0 \in D^*\big(\mathcal{E}_{F,\Theta} + \Delta(\cdot;\Omega)\big)(\bar{x},\bar{z})(z^*) \subset \partial_\Theta F(\bar{x},\bar{z}) + N(\bar{x};\Omega),$$

which justifies (9.81) and thus completes the proof of the theorem. △

Finally, we present a consequence of Theorem 9.22, which doesn't refer to the mapping F_Ω while using the SNC/PSNC properties of the initial mapping F and its inverse in the infinite-dimensional setting.

Corollary 9.23 (Necessary Optimality Conditions for Constrained Multiobjective Problems via PSNC Calculus). *Let the qualification condition* (9.78) *in Theorem* 9.22(i) *be replaced by*

$$D^*F(\bar{x},\bar{z})(0) \cap \big(-N(\bar{x};\Omega)\big) = \{0\}, \tag{9.83}$$

and let the PSNC assumption on F_Ω^{-1} be replaced by

- *either F^{-1} is PSNC at $(\bar{z},\bar{x})$ and Ω is SNC at $\bar{x}$,*
- *or F is SNC at $(\bar{x},\bar{z})$.*

Then condition (9.79) *with some $z^* \in -N(0;\Theta)$ and $\|z^*\| = 1$ is necessary for optimality in all the cases of local minimizers under consideration.*

Proof. To justify this statement, we need to check that the qualification condition (9.83) and either one of the alternative assumptions made in the corollary imply that F_Ω^{-1} is PSNC at $(\bar{z},\bar{x})$. To proceed, observe that the PSNC property of F_Ω^{-1} at $(\bar{z},\bar{x})$ is equivalent to the PSNC property at this point of the set $\operatorname{gph} F_\Omega \subset X \times Z$ with respect to Z; see Exercise 3.69. Since $\operatorname{gph} F_\Omega = \operatorname{gph} F \cap (\Omega \times Z)$, we apply the intersection rule for the PSNC property from this exercise to the sets $\Omega_1 := \operatorname{gph} F$ and $\Omega_2 := \Omega \times Z$. This gives us the required result due to the specific structures of Ω_1 and Ω_2. △

9.5 Exercises for Chapter 9

Exercise 9.24 (Properties of Quasi-Relative and Intrinsic Relative Interiors). Let $\emptyset \neq \Theta \subset X$ be a closed and convex subset of a Banach space X.

(i) Verify that $\text{qri}\,\Theta \neq \emptyset$ if the space X is separable and give an example where it fails in nonseparable spaces. *Hint:* Compare it with [104].

(ii) Show that $\text{iri}\,\Theta = \text{ri}\,\Theta$ is the conic hull of $\Theta - \bar{z}$ which is a linear subspace of Z.

(iii) Show that the inclusions in (9.5) hold as equalities if $\text{ri}\,\Theta \neq \emptyset$, while otherwise both inclusions may be strict.

(iv) Establish sufficient conditions ensuring that $\text{iri}\,\Theta \neq \emptyset$ in infinite dimensions.

Exercise 9.25 (Relationships Between Relative Minimal Points of Sets). Given a subset Ξ of a Hilbert space Z partially ordered by a closed and convex cone Θ with $\text{ri}\,\Theta = \emptyset$, construct examples showing that:

(i) $\bar{z} \in \Xi$ is an intrinsic minimal point of the set Ξ but not a quasi-relative minimal point of this set.

(ii) The set Ξ admits an intrinsic relative minimal point but not a primary relative minimal point.

(iii) Both sets $\text{ri}\,\Xi$ and $\text{iri}\,\Xi$ are empty.

Exercise 9.26 (Range of Dual Vectors for Subdifferentials of Ordered Set-Valued Mappings). Show that the range condition $-z^* \in N(0; \Theta)$ in Definition 9.6 follows from each of the inclusions $x^* \in \widehat{D}^*\mathcal{E}_{F,\Theta}(\bar{x}, \bar{z})(z^*)$ and $x^* \in D^*\mathcal{E}_{F,\Theta}(\bar{x}, \bar{z})(z^*)$.

Exercise 9.27 (Subdifferential Calculus for Ordered Set-Valued Mappings). Based on the coderivative calculus, which is presented in Chapter 3 for multifunctions between finite-dimensional spaces and in [522, Chapter 3] for multifunctions between Asplund spaces, derive major calculus rules for the basic subdifferential $\partial_\Theta F(\cdot)$ of ordered set-valued mappings in these settings.

Exercise 9.28 (Relationships Between Level-Closedness and Epiclosedness of Mappings with Ordered Values). Let $F\colon X \rightrightarrows Z$ be a mapping between Banach spaces, where Z is ordered by a closed and convex cone Θ with $\text{int}\,\Theta \neq \emptyset$. Assuming that the F is compact-valued and level-closed, show that it is epiclosed.

Hint: Proceed by using the definitions of the properties involved.

Exercise 9.29 (Normality Property). Let $\emptyset \neq \Theta \subset Z$ be a closed, convex, and pointed cone in a Banach space Z.

(i) Show that the cone Θ enjoys the normality property if it has a bounded base; in particular, when Z is finite-dimensional.

(ii) Give an example when the normality property fails in Hilbert spaces.

Exercise 9.30 (Compact-Based Property of Cones). Verify that the compact-based property of a cone $\Theta \subset Z$ is equivalent to the normality property of this cone and the compactness of the set $\Theta \cap \mathbb{B}$. *Hint:* Compare it and the statements of Exercise 9.29 with the corresponding results in [300].

Exercise 9.31 (Generalized Order Optimality). Given a mapping $f\colon X \to Z$ between normed spaces and a set $\Theta \subset Z$ containing $0 \in Z$, we say that a point $\bar{x} \in X$ is *locally* (f, Θ)*-optimal* if there exist a neighborhood U of $\bar{x}$ and a sequence $\{z_k\} \subset Z$ with $\|z_k\| \to 0$ as $k \to \infty$ such that

$$f(x) - f(\bar{x}) \notin \Theta - z_k \quad \text{for all} \quad x \in U \quad \text{and} \quad k \in \mathbb{N}. \tag{9.84}$$

(i) Let Θ be a convex cone in (9.84). Show that the introduced notion of generalized order optimality covers: (a) *Slater optimality* where $\text{ri}\,\Theta \neq \emptyset$ and there is no $x \in U$ with $f(x) - f(\bar{x}) \in$

ri Θ; (b) *weak Pareto optimality* where int $\Theta \neq \emptyset$ and there is no $x \in U$ with $f(x) - f(\bar{x}) \in \text{int}\,\Theta$; **(c)** *Pareto optimality* where there is no $x \in U$ such that $f(x) - f(\bar{x}) \in \Theta$ and $f(\bar{x}) - f(x) \notin \Theta$.

(ii) Let $\bar{x}$ be a local optimal solution to the *minimax problem*:

$$\text{minimize} \;\; \varphi(x) := \max\big\{\langle z^*, f(x)\rangle \big|\, z^* \in \Lambda\big\}, \quad x \in X,$$

where $f\colon X \to Z$, where Λ is weak* sequentially compact subset of Z^* such that there is $z_0 \in Z$ with $\langle z^*, z_0\rangle > 0$ for all $z^* \in \Lambda$, and where $\varphi(\bar{x}) = 0$ for simplicity. Show that $\bar{x}$ is locally (f, Θ)-optimal in the sense of (9.84) with

$$\Theta := \big\{z \in Z \big|\, \langle z^*, z\rangle \le 0 \;\text{ whenever }\; z^* \in \Lambda\big\}.$$

Hint: Take $z_k := z_0/k$ for all $k \in \mathbb{N}$.

(iii) Extend the notion of generalized order optimality (9.84) to set-valued costs and compare it with the notions of Pareto-type minimizers from Definition 9.4.

Exercise 9.32 (Closed Preference Relations). Given a subset $Q \subset Z^2$ for the normed space Z, we say that z_1 is *preferred* to z_2 and write $z_1 \prec z_2$ if $(z_1, z_2) \in Q$. Suppose that Q doesn't contain the diagonal (z, z) and define the *level set*

$$\mathcal{L}(z) := \big\{u \in Z \big|\, u \prec z\big\}, \quad z \in Z. \tag{9.85}$$

We say that the preference $\prec$ is *locally satiated* around $\bar{z}$ if $z \in \text{cl}\,\mathcal{L}(z)$ for all z near $\bar{z}$ and $\prec$ is *almost transitive* on Z if $v \prec z$ whenever $v \in \text{cl}\,\mathcal{L}(u)$ with $u \prec z$. If both these properties are satisfied, then the preference $\prec$ is called *closed* around $\bar{z}$.

(i) Considering the *generalized Pareto* preference:

$$z_1 \prec z_2 \;\text{ if and only if }\; z_1 - z_2 \in \Theta \;\text{ and }\; z_1 \neq z_2$$

generated by a closed cone $\Theta \subset Z$ show that this preference is almost transitive if and only if the cone Θ is convex and pointed.

(ii) Let $\prec$ be a preference on $\mathbb{R}^m$, $m \ge 3$, defined by the *lexicographical order*, i.e., $u \prec v$ if there is an integer $j \in \{0, \ldots, m-1\}$ such that $u_i = v_i$ for $i = 1, \ldots, j$ and $u_{j+1} < v_{j+1}$ for the corresponding components of the vectors $u, v \in \mathbb{R}^m$. Show that this preference is locally satiated but not almost transitive on $\mathbb{R}^m$.

Hint: Compare it with [523, Subsection 5.3.1].

Exercise 9.33 (Limiting Monotonicity and Its Weak Counterpart). Consider a weak version of the limiting monotonicity condition from Definition 9.7 with the replacement of the minimum set Min $F(\bar{x})$ therein by the collections wMin $F(\bar{x})$ of the weak Pareto efficient points.

(i) Establish sufficient conditions for weak limiting monotonicity of the type of Proposition 9.8 with the closedness assumption on wMin $F(\bar{x})$.

(ii) Give an example of a mapping in $\mathbb{R}^2$ with $\Theta = \mathbb{R}^2_+$ where the set wMin $F(\bar{x})$ is closed and the weak limiting monotonicity property holds, while it is not the case for Min $F(\bar{x})$ and the limiting monotonicity condition.

Hint: Compare it with [55, Theorem 3.4 and Remark 3.5].

Exercise 9.34 (Limiting Monotonicity and Domination Property).

(i) Give an example where the limiting monotonicity condition (9.8) holds but the domination property (9.7) fails.

(ii) Give an example where the weak monotonicity condition from Exercise 9.33 holds but the weak version of domination (with the replacement of Min $F(\bar{x})$ in (9.7) by the weak minimum set wMin $F(\bar{x})$) fails.

Exercise 9.35 (Ekeland-Type Variational Principle for Sets in Product Spaces). Let Ξ be a nonempty set in the product of Banach spaces $X \times Z$, where Z is partially ordered by a proper, closed, and convex cone $\Theta \subset Z$ with $\Theta \setminus (-\Theta) \neq \emptyset$.

(i) Derive the corresponding version of Ekeland's variational principle for sets Ξ by specifying Theorem 9.10 for the associated set-valued mapping F_Ξ defined by

$$F_\Xi(x) := \{z \in Z \big| \ (x, z) \in \Xi\} \ \text{ with } \ \text{gph}\, F_\Xi = \Xi.$$

(ii) Establish relationships between the result of (i) and the so-called *authentic minimal point theorem* for sets $\Xi \subset X \times Z$ obtained in [300, Theorem 3.10.7].

Hint: Compare it with [56, Corollary 3.6 and Remark 3.7].

Exercise 9.36 (Ekeland-Type Variational Principle via Weak Minimizers). Establish a version of Theorem 9.10 for weak minimizers provided that $\text{int}\,\Theta \neq \emptyset$ with replacing the set Min $F(\bar{x})$ by wMin $F(\bar{x})$ and the preference relation (9.1) by

$$z_1 \prec z_2 \ \text{ if and only if } \ z_2 - z_1 \in \text{int}\,\Theta.$$

Hint: Proceed as in the proof of Theorem 9.10 by using the weak limiting monotonicity condition from Exercise 9.33.

Exercise 9.37 (Estimates in the Subdifferential Variational Principle). Derive the relationships in (9.30) from the approximate extremal principle (9.29) due to the set structures in (9.26).

Exercise 9.38 (Subdifferential Variational Principle for Weak Approximate Minimizers). Establish a counterpart of Theorem 9.12 for weak versions of approximate minimizers from Definition 9.9. *Hint:* Proceed similarly to the proof of Theorem 9.12 with the usage of the weak Ekeland-type variational principle from Exercise 9.36 instead of Theorem 9.10.

Exercise 9.39 (Strong Limiting Monotonicity). Let $F\colon X \rightrightarrows Z$ with a partially ordered space Z, and let $\bar{x} \in \text{dom}\, F$.

(i) Show that all the conditions listed in Proposition 9.8 ensure the strong limiting monotonicity of F at $\bar{x}$. *Hint:* Proceed as in the proof of Proposition 9.8.

(ii) Give an example of a mapping that enjoys the limiting monotonicity property (9.8) but not the strong limiting monotonicity one.

(iii) Formulate a version of strong limiting monotonicity for weak minimizers and establish sufficient conditions for it.

Exercise 9.40 (Existence of Quasi-Relative and Pareto Minimizers). Consider the setting of Theorem 9.15.

(i) Identify those parts in the proof of Theorem 9.15 which don't work in the case of quasi-relative and Pareto minimizers.

(ii) Find additional assumptions under which the proof procedure of Theorem 9.15 can be modified to establish the existence of quasi-relative minimizers.

(iii) Find additional assumptions under which the proof procedure of Theorem 9.15 can be modified to establish the existence of Pareto minimizers.

Exercise 9.41 (Existence of Relative and Weak Pareto Minimizers for Constrained Multiobjective Problems via Basic Subdifferential Calculus). Using basic subdifferential calculus for model (9.55) in the finite-dimensional and Asplund space frameworks (see Chapters 2–4 above and [522, Chapter 3]), deduce from Theorem 9.15 and its corollaries efficient results ensuring the existence of relative and weak Pareto minimizers in the following constrained settings:

(i) Problems with only geometric constraints (9.54).

(ii) Problems with inequality and equality constraints described by

$$\Omega := \{x \in X \big| \ \varphi_i(x) \le 0, \ i = 1, \ldots, m; \ \varphi_i(x) = 0, \ i = m+1, \ldots, m+r\}$$

via some Lipschitz continuous functions φ_i, $i = 1, \ldots, m + r$.

(iii) Problems with operator constraints $G(x) \cap S \neq \emptyset$ for $G\colon X \rightrightarrows Y$ and $S \subset Y$.

(iv) Problems with equilibrium constraints described generally as $0 \in G(x) + Q(x)$ via certain set-valued mappings; cf. [523, Chapter 5]. *Hint:* Follow the procedure developed in [53] for the case of weak Pareto minimizers.

Exercise 9.42 (Vector Subdifferential Representations for Lipschitz Continuous Mappings). Let $f: X \to Z$ be a single-valued mapping between Banach spaces that is locally Lipschitzian around a given point $\bar{x}$.

(i) Specify representation (9.56) of the regular subdifferential via scalarization.

(ii) Find conditions under which a similar representation holds for the basic vector subdifferential from Definition 9.6(ii). *Hint:* Compare with [522, Subsection 3.1.3] concerning the normal coderivative case.

Exercise 9.43 (Special Sum Rules for Regular Subgradients of Scalar and Vector Functions). Let f be either an extended-real-valued function or a single-valued mapping between Banach spaces.

(i) Derive a scalar counterpart of the sum rule (9.57). *Hint:* Compare with [547].

(ii) Verify the sum rule (9.57) as formulated in the vector case.

Exercise 9.44 (Properties of Relative Minimizers). Prove the equality in (9.69) for all the three types of relative minimizers. *Hint:* Compare with [103, Lemma 3.1].

Exercise 9.45 (Necessary Conditions for Relative Pareto Minimizers in Problems with Structured Constraints). Derive counterparts of Theorem 9.22 and Corollary 9.23 for multiobjective problems with structured constraints listed in Exercise 9.41. *Hint:* Use the generalized differential and SNC/PSNC calculi developed in [522, Chapter 3] and discussed in the previous chapters of the book.

Exercise 9.46 (Super Minimizers in Multiobjective Optimization). Given $F: X \rightrightarrows Z$ and $\Omega \subset X$, consider the constrained optimization problem (9.54), where "minimization" is understood with respect to the generalized Pareto preference relation $\preceq$ defined in (9.1) via a closed and convex cone $\Theta \subset Z$. We say that a pair $(\bar{x}, \bar{z}) \in \text{gph}\, F$ with $\bar{x} \in \Omega$ is a local *super minimizer* of problem (9.54) if there exist a neighborhood U of $\bar{x}$ and a number $M > 0$ such that

$$\|z - \bar{z}\| \le M\|v\| \text{ if } x \in \Omega \cap U,\ z \in F(x),\ v \in Z \text{ with } z - \bar{z} \preceq v. \tag{9.86}$$

(i) Compare this notion with those in Definition 9.4.

(ii) Show that even for $X = \mathbb{R}$, $Z = \mathbb{R}^2$, and $\Theta = \mathbb{R}^2_+$, the necessary optimality condition (9.79) for weak and other types of local minimizers obtained in Corollary 9.23 is not necessary for local super minimizers in (9.54).

(iii) Using the techniques of variational analysis and generalized differentiation similar to those employed in Section 9.4, show that under the validity of the same assumptions as in Theorem 9.22(i), a given super minimizer $(\bar{x}, \bar{z})$ of (9.54) satisfies the following coderivative optimality conditions: there exists $-z^* \in N(0; \Theta)$ with $\|z^*\| \le M$ for the constraint M from (9.86) such that

$$0 \in D^*F(\bar{x}, \bar{z})(z^* - v^*) + N(\bar{x}; \Omega)$$

whenever $v^* \in \mathbb{B}^* \subset Z^*$. Derive a subdifferential counterpart of this condition similarly to Theorem 9.22(ii). *Hint:* Compare it with [54].

Exercise 9.47 (Extremal Systems of Multifunctions). Let $S_i: M_i \rightrightarrows X$, $i = 1, \ldots, m$, be set-valued mappings from metric spaces (M_i, d_i) into a normed space X. We say that $\bar{x}$ is a *local extremal point* of the system $\{S_1, \ldots, S_m\}$ at $(\bar{s}_1, \ldots, \bar{s}_m)$ provided that $\bar{x} \in S_1(\bar{s}_1) \cap \ldots \cap S_m(\bar{s}_m)$ and there exists a neighborhood U of $\bar{x}$ such that for every $\varepsilon > 0$ there are $s_i \in \text{dom}\, S_i$ satisfying the conditions

$$d(s_i, \bar{s}_i) \le \varepsilon, \quad \text{dist}\big(\bar{x}; S_i(s_i)\big) \le \varepsilon, \quad i = 1, \ldots, m,$$
$$S_1(s_1) \cap \ldots \cap S_m(s_m) \cap U = \emptyset.$$

(i) Consider the vector minimization problem:

$$\text{minimize } f(x) \text{ subject to } x \in \Omega \subset X \tag{9.87}$$

with respect to the closed preference $\prec$, where $f\colon X \to Z$ is a mapping between normed spaces. Show that $(\bar{x}, f(\bar{x}))$ is a local extremal point at $(f(\bar{x}), 0)$ for the system of multifunctions $S_i\colon M_i \rightrightarrows X \times Z$, $i = 1, 2$, defined by

$$\begin{aligned} &S_1(s_1) := \Omega \times \text{cl}\,\mathcal{L}(s_1) \text{ with } M_1 := \mathcal{L}(f(\bar{x})) \cup \{f(\bar{x})\},\\ &S_2(s_2) = S_2 := \big\{(x, f(x))\big|\ x \in X\big\} \text{ with } M_2 := \{0\} \end{aligned}$$

via the level set (9.85) associated with the preference $\prec$.

(ii) Let $(\bar{x}, \bar{y}) \in \Omega \times \Theta$ be a *saddle point* of a payoff function $\varphi\colon X \times Y \to \mathbb{R}$ over subsets $\Omega \subset X$ and $\Theta \subset Y$ of normed spaces, i.e.,

$$\varphi(x, \bar{y}) \le \varphi(\bar{x}, \bar{y}) \le \varphi(\bar{x}, y) \text{ whenever } (x, y) \in \Omega \times \Theta.$$

Define a set-valued mapping $S_1\colon [\varphi(\bar{x}, \bar{y}), \infty) \times (-\infty, \varphi(\bar{x}, \bar{y})] \rightrightarrows \Omega \times \mathbb{R} \times \Theta \times \mathbb{R}$ and a set $S_2 \subset \Omega \times \mathbb{R} \times \Theta \times \mathbb{R}$ by

$$S_1(\alpha, \beta) := \Omega \times [\alpha, \infty) \times \Theta \times (-\infty, \beta], \quad S_2 := \text{hypo}\,\varphi(\cdot, \bar{y}) \times \text{epi}\,\varphi(\bar{x}, \cdot)$$

and show that the point $\big(\bar{x}, \varphi(\bar{x}, \bar{y}), \bar{y}, \varphi(\bar{x}, \bar{y})\big)$ is locally extremal for the system of multifunctions $\{S_1, S_2\}$ at $\big(\varphi(\bar{x}, \bar{y}), \varphi(\bar{x}, \bar{y})\big)$.

Hint: Proceed by the definitions and compare it with [523, Subsection 5.3.3].

Exercise 9.48 (Extremal Principles for Systems of Multifunctions). Let $\bar{x} \in S_1(\bar{s}_1) \cap \ldots \cap S_m(\bar{s}_m)$ be a locally extremal point at $(\bar{s}_1, \ldots, \bar{s}_m)$ for closed-valued multifunctions $S_i\colon M_i \rightrightarrows X$ from metric spaces (M_i, d_i) into an Asplund space X.

(i) Prove that for every $\varepsilon > 0$ there are $s_i \in \text{dom}\, S_i$, $x_i \in S_i(s_i)$, and $x_i^* \in X^*$, $i = 1, \ldots, m$, satisfying the relationships of the *approximate extremal principle*:

$$\begin{aligned} &d(s_i, \bar{s}_i) \le \varepsilon, \quad \|x_i - \bar{x}\| \le \varepsilon, \quad x_i^* \in \widehat{N}\big(x_i; S_i(s_i)\big) + \varepsilon\mathbb{B}^*,\\ &x_1^* + \ldots + x_m^* = 0, \qquad \|x_1^*\| + \ldots + \|x_m^*\| = 1. \end{aligned}$$

Hint: Use Ekeland's variational principle and the approximate extremal principle for systems of sets; compare it with the proof of [523, Theorem 5.38].

(ii) Find verifiable conditions ensuring the validity of the *exact extremal principle* for systems of multifunctions in terms of limiting normals. *Hint:* Compare it with [523, Proposition 5.70 and Theorem 5.72].

(iii) Apply the extremal principles from (i) and (ii) to derive necessary optimality conditions for vector optimization problems of type (9.87) with respect to closed preferences. *Hint:* Use the reduction to the extremal system in Exercise 9.47 and compare it with [523, Theorem 5.73].

Exercise 9.49 (Necessary and Sufficient Conditions for Generalized Order Optimality). Let $f\colon X \to Z$ be a mapping between Banach spaces, and let $\Omega \subset X$ and $\Theta \subset Z$ be such sets that $\bar{x} \in \Omega$ and $0 \in \Theta$. Consider the generalized epigraph

$$\mathcal{E}(f, \Omega, \Theta) := \{(x, z) \in X \times Z\big|\ f(x) - z \in \Theta,\ x \in \Omega\big\}$$

and suppose that it is locally closed around $(\bar{x}, \bar{z})$ with $\bar{z} := f(\bar{x})$.

(i) Assume that $\bar{x}$ is a locally (f, Θ)-optimal point subject to the constraint $x \in \Omega$, that the space X is Asplund, and that the space Z is finite-dimensional. Prove that there is $z^* \in Z^*$ satisfying the conditions

$$(0, -z^*) \in N\big((\bar{x}, \bar{z}); \mathcal{E}(f, \Omega, \Theta)\big),\ z^* \ne 0, \tag{9.88}$$

which always imply that $z^* \in N(0;\Theta)$; it also yields $0 \in D^*_N f_\Omega(\bar{x})(z^*)$ provided that f is continuous around $\bar{x}$ relative to Ω and that Ω and Θ are locally closed around $\bar{x}$ and 0, respectively. If in addition f is Lipschitz continuous around $\bar{x}$ relative to Ω, then show that (9.88) is equivalent to

$$0 \in \partial\langle z^*, f_\Omega\rangle(\bar{x}), \quad z^* \in N(0;\Theta) \setminus \{0\}, \tag{9.89}$$

where f_Ω stands for the restriction of f on Ω.

(ii) Suppose in addition to the assumptions in (i) in the case of the continuity of f relative to Ω that either Θ is SNC at the origin, or f_Ω^{-1} is PSNC at $(\bar{z},\bar{x})$. Then show that there is $z^* \in Z^*$ satisfying

$$0 \neq z^* \in N(0;\Theta) \cap \ker D^*_N f_\Omega(\bar{x}),$$

which is equivalent to (9.89) and also to (9.88) provided that f is Lipschitz continuous around $\bar{x}$ relative to Ω and that the restriction f_Ω is *strongly coderivatively normal* at this point in the sense that $D^*_N f_\Omega(\bar{x},\bar{z}) = D^*_M f_\Omega(\bar{x},\bar{z})$.

Hint: To verify (i), take $\bar{z} = 0$ for simplicity and apply the exact extremal principle from Exercise 2.31 to the closed set system

$$\Omega_1 := \mathcal{E}(f,\Omega,\Theta) \;\text{ and }\; \Omega_2 := \operatorname{cl} U \times \{0\} \;\text{ at }\; (\bar{x},0) \in X \times Z, \tag{9.90}$$

where U is a neighborhood of the local optimality of $\bar{x}$ in (9.84) relative to Ω. Justifying (ii) requires more involved elaborations that include the usage of a product version of the extremal principle and PSNC preservation rules for set intersections; compare with the proof of [523, Theorem 5.59].

(iii) Give examples in finite-dimensional spaces showing that necessary optimality conditions from (i) and (ii) are not sufficient for the generalized order optimality.

(iv) Assume in the general Banach space setting that Ω is locally convex around $\bar{x}$, that Θ is a convex cone with $\operatorname{int}\Theta \neq \emptyset$, and that f is locally Θ-convex on Ω in the sense that there is a convex neighborhood U of $\bar{x}$ such that

$$f\big(\lambda x + (1-\lambda)u\big) \in \lambda f(x) + (1-\lambda) f(u) - \Theta \;\text{ for all }\; x,u \in \Omega \cap U.$$

Show that in this case the conditions in (9.88) are sufficient for (f,Θ)-optimality of $\bar{x}$ subject to the constraint $x \in \Omega$. What about the sufficiency of the other necessary optimality conditions in (i) and (ii)? *Hint:* Compare it with [718, Theorem 4.5].

Exercise 9.50 (Sufficient Optimality Conditions for Global Weak Pareto Maximizers in Multiobjective Problems). Given a closed and convex cone $\Theta \subset Z$ with $\operatorname{int}\Theta \neq \emptyset$, consider the following set-valued *maximization* problem:

$$\Theta - \text{maximize } \; F(x) \;\text{ subject to }\; x \in \Omega, \tag{9.91}$$

where values of $F: X \rightrightarrows Z$ are partially ordered by

$$z_1 \prec z_2 \;\text{ if and only if }\; z_2 - z_1 \in \operatorname{int}\Theta.$$

We say that a feasible pair $(\bar{x},\bar{z})$ is a *global weak Pareto maximizer* of (9.91) if there is no $z \in F(x)$ with $x \in \Omega$ such that

$$F(\Omega) \cap (\bar{z} + \operatorname{int}\Theta) = \emptyset.$$

(i) Find appropriate assumptions on the data of (9.91) so that the conditions

$$0 \notin \partial_\Theta F(\bar{u},\bar{v}) + N(\bar{u};\Omega), \;\; \partial_\Theta F(\bar{u},\bar{v}) + N(\bar{u};\Omega) \subset N(\bar{u};\Omega)$$

for all $(\bar{u}, \bar{v}) \in \operatorname{gph} F$ with $\bar{u} \in \Omega$ and $\bar{v} \in \bar{z} - \operatorname{bd}\Theta$ are *sufficient* for the global weak Pareto maximality of $(\bar{x}, \bar{z})$. *Hint:* Follow the scheme of [58] with applying the approximate *extremal principle* in Asplund spaces.

(ii) Specify and improve the conditions in (i) in the cases of single-valued and real-valued objectives in (9.91).

(iii) Clarify whether the conditions from (i) are sufficient for (global or local) *Pareto maximizers* and investigate the possibility to replace the basic subdifferential $\partial_\Theta F$ therein by the *regular* one.

Exercise 9.51 (Multiobjective Optimization with Equilibrium Constraints). Let $F\colon X \times Y \rightrightarrows Z$, $G\colon X \times Y \rightrightarrows W$, and $Q\colon X \times Y \rightrightarrows W$ be set-valued mappings between Banach spaces with some ordering on the space Z. Consider the following parametric multiobjective optimization problem:

$$\text{minimize } \; F(x, y) \; \text{ subject to } \; 0 \in G(x, y) + Q(x, y), \tag{9.92}$$

where the "minimization" in (9.92) is understood in the sense of some ordering or equilibrium relations and where the constraints therein can be treated as generalized equilibrium constraints with both base $G(x, y)$ and field $Q(x, y)$ mappings being set-valued; cf. (6.73). Problems of this type arise, e.g., in modeling *set-valued variational inequalities*: given $G\colon X \times Y \rightrightarrows Y^*$ and $\Xi \subset Y$, find $y \in \Xi$ such that

$$\text{there is } \; y^* \in G(x, y) \; \text{ with } \; \langle y^*, u - y\rangle \geq 0 \; \text{ for all } \; u \in \Xi.$$

Another source of the multivalued constraints in (9.92) is given by the KKT systems

$$0 \in \partial_y \varphi(x, y) + N\big(y; \Xi(x)\big), \quad (x, y) \in X \times Y, \tag{9.93}$$

arising as necessary (and sufficient in the convex case) optimality conditions for the parametric lower-level problems of the type:

$$\text{minimize } \; \varphi(x, y) \; \text{ subject to } \; y \in \Xi(x) \subset Y,$$

in bilevel programming, where the base $G(x, y) := \partial_y \varphi(x, y)$ is set-valued provided that the cost function φ is nondifferentiable with respect to the decision variable y. Note that in the case where the cost F on the upper level in (9.92) is vector-valued or set-valued, problems of this type describe *multiobjective bilevel programs* in contrast to usual ones with scalar costs. Observe also that, since equilibrium relations in (9.92) may appear in both costs (upper level) and constraints (lower level) and can be viewed as Pareto-type as well as Nash-type equilibria, models (9.92) are often labeled as *equilibrium problems with equilibrium constraints* (EPECs).

(i) Derive necessary optimality conditions for problems (9.92) in Asplund spaces, where the optimization is considered with respect to Pareto-type notions studied in this chapter. *Hint:* Use the extremal principle and compare the results with those obtained in [51] in the case of the generalized order optimality, which is defined in Exercise 9.31 for problems with single-valued costs.

(ii) Specify the results of (i) for equilibrium constraints given in the variational form (9.93) and also when G and Q are represented in the composite subdifferential forms as in (3.41) and (3.48). *Hint:* Use the second-order subdifferentials and the corresponding calculus rules similarly to Section 3.3. Compare this with [526], where the analysis is done for the case of generalized order optimality in (9.92).

(iii) Consider EPEC models involving *noncooperative* (Cournot-Nash) equilibria on either upper or lower level in (9.92) and deduce necessary optimality conditions for them from those obtained in the general scheme of (i). *Hint:* Compare it with the results of [560] dealing with weak Pareto optimality on the upper level and Cournot-Nash equilibrium on the lower one with applications to oligopolistic markets.

9.6 Commentaries to Chapter 9

Problems of *vector optimization*, with *single-valued* vector objectives, have been under consideration in optimization theory and applications for a long time. Original motivations mainly came from economics, engineering, etc., but then the vector optimization theory has been developed for its own sake with a variety of approaches and results; see, e.g., the books [147, 244, 300, 301, 385, 389, 474, 478, 507, 523, 629] and the references therein. Largely related while different from vector optimization problems are *vector variational inequalities*, various models of *equilibria*, and *EPECs*; see [10, 17, 65, 85, 88, 89, 146, 188, 219, 272, 291, 292, 317, 323, 356, 360, 500, 424, 491, 492, 498, 520, 565, 560, 598, 622, 628, 735, 740, 755] among other publications. Note that, besides pure theoretical developments in vector optimization and related topics, there are efficient algorithms to solve such problems numerically; see, e.g., [9, 97, 145, 159, 275, 304, 379, 424, 597], where the reader can find additional bibliographies.

Problems of *set-valued optimization*, which objectives are given by set-valued mappings with ordered values, have come to consideration in optimization theory much later. Among the first models and results in set-valued optimization, we mentioned those by Oettli [618] and his Ph.D. students Tagawa [702], by Corley [179], El Abdoini and Thibault [251], and Kuroiwa [446]. After that, various aspects of set-valued optimization and its applications were studied in a great many publications; see, e.g., [184, 237, 362, 386, 324, 325, 343, 415, 646, 701, 788] to list just a few in addition to the sources discussed below and the references therein.

The recent monograph by Khan, Tammer, and Zălinescu [409] provides a comprehensive, systematic study of set-valued optimization problems from various viewpoints, together with related topics and some applications. The extensive bibliography of [409] refers the reader to the additional material. Although other approaches to multiobjective optimizations are also discussed in [409], the main attention is paid to the primal-space approach (involving tangent cone and derivative approximations of sets and mappings) and scalarization techniques largely developed by the authors.

The major emphasis of Chapter 9 and the subsequent Chapter 10 is a *dual-space* approach of variational analysis, which is based on *extremal principles* (not at all related to scalarization) and utilizes normal cone and coderivative constructions for sets and mappings that may be *dual to none*. This approach was developed in the author's book [523] for problems of (single-objective) vector optimization; see also the references and commentaries therein. Its extension to set-valued optimization (with new results for vector optimization problems as well) presented in Chapter 9 is mainly based on the paper by Bao and Mordukhovich [55]; see in addition the related publications of these and other researchers cited below.

Section 9.1. Among the major motivations to introduce and study in [55], the notions of *relative Pareto minimizers* for general problems of set-valued optimization were addressing multiobjective problems with *nonsolid* ordering cones. This takes into account the fact that the nonempty interior requirement conventionally imposed on ordering cones in multiobjective optimization has been realized as restrictive for both optimization theory and applications, particularly in infinite dimensions, where it fails (together with the nonempty relative interior condition) in many important settings. Note to this end that the *quasi-relative interior* of the ordering cone used in Definition 9.4(v) of quasi-relative minimizers is nonempty for any closed and convex cone in a separable Banach space; this was proved by Borwein and Lewis [104] and then was employed in infinite-dimensional analysis and optimization in [106, 242, 103, 116, 117, 119, 120, 188, 303, 492, 493, 758] along with other publications.

Several attempts to avoid the nonempty interior and relative interior assumptions were undertaken in, e.g., [236, 242, 317, 430, 506, 507, 523, 586] for some classes of vector optimization problems in finite and infinite dimensions. The new techniques developed in [55] and in this chapter are different from those used in the aforementioned publications and address not only *necessary optimality conditions* in set-valued optimization but also the *existence* of relative Pareto minimizers.

The *subdifferential* notions for *set-valued* (in particular, vector-valued) mappings with *ordered* values from Definition 9.6 first appeared in another paper by Bao and Mordukhovich [52]. These constructions were introduced in the same geometric pattern as the corresponding subdifferentials for extended-real-valued functions in Chapter 1 while clearly admitting similar analytic representations. Due to their definitions via coderivatives, these "vector" subdifferentials inherit properties and calculus similar to those for scalar functions. Other notions of subdifferentials for vector/set-valued mappings associated with various types of efficiency in multiobjective optimization can be found in [317, 409, 447, 692, 701] and the references therein.

Section 9.2. In this section we present two *variational principles* for *ordered set-valued mappings* that play, together with the underlying extremal principle for closed sets, a crucial role in deriving existence theorems and necessary optimality conditions for relative Pareto minimizers. The first one is an appropriate version of the seminal *Ekeland's variational principle* in the case of multifunctions with (partially) ordered values. Its formulation and proof extend the previous result from [52], where the notions of *approximate $\varepsilon\xi$-minimizers* and its strict counterpart appeared first. A significant part in the proof of Theorem 9.10, which has no analogs in the classical Ekeland principle for extended-real-valued functions and its vector-valued counterparts, is verifying the existence of $\bar{z} \in \operatorname{Min} F(\bar{x})$ satisfying the conditions in (9.10) and (9.11). The last condition (9.12) of the theorem is based on the definition of approximate $\varepsilon\xi$-minimizers. The *limiting monotonicity* condition and its modifications presented above were also introduced in [52] and improved in [55].

Various extensions of Ekeland's variational principle to vector-valued and set-valued mappings have been a subject of many publications; see, e.g., [10, 75, 300, 319, 409, 415, 651] and their bibliographies to mention just a few. Our motivations in [52, 55] came strictly from the main issues of these papers to obtain appropriate existence theorems and necessary optimality conditions in set-valued optimization. The proofs developed in [52, 55] occurred to be instrumental to derive variational principles and related results for set-valued mappings on *quasimetric spaces* [170] with cone-valued ordering *variable structures*; see the papers by Bao, Mordukhovich, and Soubeyran [61, 62, 63] for more details. Such requirements unavoidably arise in applications to several models in *behavioral sciences* (psychology, economics, human behavior, etc.) treated from the viewpoint of Soubeyran's *variational rationality* approach [695]. These models were comprehensively studied in [61, 62, 63] by using variational principles and other techniques of variational analysis.

Note to this end that, although problems of vector optimization with *variable preferences* appeared in the literature in the 1970s (see Yu [754]), recent years have witnessed a growing interest to such problems from both viewpoints of optimization theory and application. We refer the reader to the excellent book by Eichfelder [245] and related papers [246, 247] for a thorough study and applications of variable structures in vector optimization by using scalarization techniques. The dual-space variational approach of [60] allowed us to obtain general necessary conditions for nondominated solutions to such problems by using the extremal principle; see also [269] for related developments.

Coming back to the material of Section 9.2, observe that *subdifferential variational principle* from Theorem 9.12 extends to set-valued mappings with ordered values of the (lower) subdifferential variational principle by Mordukhovich and Wang [587] established under the same name for extended-real-valued functions; see Exercise 2.39. Similarly to the scalar case, the proof of Theorem 9.12 is based on the application of the extremal principle for set systems together with the new version of Ekeland's variational principle for ordered set-valued mappings. The subdifferential variational principle from Theorem 9.12 improves the previous one from [52] established under essential more restrictive assumptions.

Section 9.3. The main result of this section is Theorem 9.15, which justifies the *existence* of *intrinsic* relative Pareto minimizers under the validity of the *regular* subdifferential Palais-Smale condition from Definition 9.13(i) that is taken from [55]. The previous significantly weaker result in this direction given in [52] verifies the existence of *weak* Pareto minimizers under the subdifferential Palais-Smale condition from Definition 9.13(ii) formulated in terms of the larger basic subdifferential for ordered mappings and other more restrictive assumptions in comparison with those in Theorem 9.15. The reader can see that the involved proof of Theorem 9.15 employs both variational principles for ordered set-valued mappings from Section 9.2 together with the underlying extremal principle for closed subsets of Asplund spaces.

Although the *basic* subdifferential Palais-Smale condition from Definition 9.13(ii) is generally more restrictive than its regular counterpart, it has advantages in applications to *constrained* set-valued optimization problems due to much better calculus rules available for the basic subdifferential. Implementations of this approach to the existence of solutions for some classes in constrained multiobjective optimization (including those with equilibrium constraints) are given in yet another paper by Bao and Mordukhovich [53]. On the other hand, Theorem 9.18 and its corollary presented above (which are taken from [55]) justify the existence of relative Pareto minimizers in multiobjective problems with explicit geometric constraints by using a specific sum rule for the regular subdifferential of ordered vector-valued mappings.

Section 9.4. Following [55], we develop in this section a *unified dual-space* approach to deriving *necessary optimality conditions* for all the types of Pareto, weak Pareto and relative Pareto minimizers of multiobjective problems given in both unconstrained and constrained formats in the Asplund space setting. The results obtained in this vein by using the underlying *extremal principle* are expressed in the pointbased terms of *coderivatives* and *subdifferentials* of ordered set-valued mappings in exactly the same way for all the Pareto-type minimizers under consideration, with the only distinctness between them in imposing different *SNC/PSNC* assumptions at the minimizers in question. Observe that such assumptions are not needed (hold automatically) for weak Pareto minimizers in the general setting and for other types of minimizers of multiobjective problems in finite-dimensional spaces.

We distinguish between necessary optimality conditions for problems of "minimizing" set-valued mappings $F\colon X \rightrightarrows Z$ (i.e., given in the *unconstrained* format, with the implicit constraint $x \in \operatorname{dom} F$) and problems with explicit constraints of type (9.54) and their specifications. Optimality conditions for problems of the first type are called *Fermat rules* and are expressed via the basic coderivative of F and–under a slightly different PSNC assumption–in terms of the basic subdifferential of F; see Theorem 9.20 taken from [55]. The coderivative version (9.62) of this result was obtained by Zheng and Ng [773] for efficient/Pareto optimal solutions under the "dual compactness" requirement on the pointed ordering cone Θ that is more restrictive than our SNC property, while no alternative assumptions on F were made in [773]. We refer the reader to the paper by Ha [321] for a survey and further results on the coderivative Fermat rules in multiobjective optimization with considering also various *properly efficient* solutions (Benson, Henig, etc.). Observe that, in contrast to scalar optimization, the Fermat rule version in terms of the regular subdifferential fails for multiobjective problems; see Remark 9.21.

The necessary conditions for set-valued optimization problems with *explicit constraints* are derived from those for implicit one by employing well-developed *calculus* for our basic constructions. They are known as *Lagrange multiplier rules* by the analogy with scalar problems. In Theorem 9.22 taken from [55], we present such conditions for problems with geometric constraints $x \in \Omega$, but the available calculus rules allow us to proceed with more structural constraints of functional, operator, equilibrium, and other types; see [51, 52, 53, 54, 59] for some implementations. As in the case of unconstrained problems, we distinguished in Theorem 9.22 coderivative and subdifferential necessary conditions for all Pareto-type minimizers under consideration. Note that the coderivative condition (9.79) is expressed in terms of the *normal* coderivative $D^*F = D^*_N F$, while the qualification condition (9.78) is formulated via the smaller *mixed* coderivative. Similarly, the subdifferential Lagrange multiplier rule (9.81) is formulated via the basic subdifferential of F, while the corresponding qualification condition (9.80) uses the smaller singular subdifferential of the ordered mapping F introduced in [52].

To the best of our knowledge, first results of the Lagrange multiplier rule via coderivatives for weak Pareto minimizers in constrained set-valued optimization were obtained by El Abdoini and Thibault [251] under some interiority assumptions. Improved coderivative conditions were later derived by Zheng and Ng [774] for Pareto efficient solutions under certain "dual compactness" assumption on the ordering cone Θ, which yields the SNC property of Θ in Theorem 9.22. An interesting approach and coderivative conditions were developed by Ha [318] for strongly efficient solutions of multiobjective problems by using scalarization and subdifferential estimates for marginal functions discussed above in Chapter 4. Further results in this direction can be found in

[59, 64, 320] and the references therein for various notions of extended Pareto-type optimality. We particularly mention impressive developments by Bao and Tammer [64] who combined the scalarization technique from the paper by Gerth (Tammer) and Weidner [279] with the basic tools of generalized differentiation presented in this book to establish new versions of the Lagrange multiplier rule for efficient and proper efficient solutions to set-valued optimization problems while providing in this way valuable applications to models of risk management.

Section 9.5. As in the case of the previous chapters, the exercises presented in this section are of different levels of difficulties. Some of them can be derived from definitions and well-known results, while some are essentially more involved and even unsolved; see below. Hints and references to the original sources are given when needed and available. The following comments to some of the exercises, in addition to those made in Sections 9.1–9.4, seem to be useful.

The notion of *generalized order optimality* discussed in Exercise 9.31 goes back to the early work by Kruger and Mordukhovich [430, 441, 506, 507] while being directly related by the notion of set extremality without using any scalarization. In [523, Subsections 5.3.1 and 5.3.2] and the commentaries to them, the reader can find more information on this notion and the results available by that time, including those formulated in Exercise 9.31 together with necessary conditions for generalized order optimality in problems of vector optimization. The necessary optimality conditions from Exercise 9.49(i,ii) are due to the author [523, Theorem 5.59], while their sufficiency and examples mentioned in Exercise 9.49(iii,iv) are due to Tuyen and Yen [718]. We also refer the reader to [51, 59, 526, 717, 718] for other results in this direction. A proper extension of this notion to set-valued optimization with establishing existence theorems and optimality conditions is a *challenging issue*.

Closed preferences discussed in Exercise 9.32 were introduced by Mordukhovich, Treiman, and Zhu [586] who defined in that paper the *extremality notion* for systems of *multifunctions* from Exercise 9.47 and derived the versions of the *extended extremal principles* for such systems formulated in Exercise 9.48; see [523, Subsections 5.3.1 and 5.3.3] for more details. Necessary optimality conditions for vector optimization problems with respect to closed preferences were given in [523, Subsection 5.3.4]. The reader can find further results in this direction in [57, 76, 468, 525, 592] and the references therein. Applications of these notions and results to problems of *set-valued* optimization have *not been developed* yet.

A major *open question* in the existence theory for relative Pareto optimality concerns finding appropriate conditions ensuring the existence of *quasi-relative* Pareto minimizers for multiobjective problems when $\text{ri}\,\Theta = \emptyset$; see Exercise 9.40. We strongly believe that it can be done in the framework of Theorem 9.15. Note that this theorem also doesn't contain existence statements for the usual Pareto/efficient solutions as well as for their properly efficient counterparts.

The notion of *super minimizers* (or super efficiency) discussed in Exercise 9.46 was introduced by Borwein and Zhuang [115] for problems of (single-valued) vector optimization and then has been studied in many publications; see, e.g., [312, 321, 357, 409] and the references therein. The extension of this notion to problems of set-valued optimization and the necessary optimality conditions from Exercise 9.46 are taken from the author's joint paper with Bao [54].

Similarly to the case of scalar problems, *sufficient conditions* for vector/set-valued "minimization" are known under some convexity and the like; see, e.g., [251, 409, 718]. The results discussed in Exercise 9.50, which are taken from Bao and Mordukhovich [58], go in differential direction. They present sufficient conditions for *global weak* Pareto solutions to "maximization" problems without any convexity assumptions. We are not familiar with any other results of this type for (vector or set-valued) multiobjective problems, but in the scalar case, certain analogs of such results were obtained by Hiriart-Urruty and Ledyaev [351] and Dutta [241] under some convexity assumptions. Note that the proof of the main theorem in [58] (*sufficient* conditions) is based on the *extremal principle*, which provides *necessary* conditions for set extremality. A challenging *open question* remains about the possibilities to establish counterparts of the results in [58] for other (not just global weak) Pareto-type maximizers of multiobjective problems.

Observe finally that deriving existence theorems and subdifferential optimality conditions for multiobjective problems with *structural* costs and/or constraints largely depends on *subdifferential calculus* for mappings with *ordered values* in both finite and infinite dimensions that are still due to be developed for the basic subdifferential in full generality and for its regular analog in particular settings.

Chapter 10
Set-Valued Optimization and Economics

The concluding chapter of the book is devoted to applications of advanced constructions and techniques of variational analysis to economic modeling. As our basic framework, we consider the fundamental model of *welfare economics*, which has been broadly studied in the economic and mathematical literature including the author's book [523, Chapter 8]; see more discussions in Section 10.6. Here we develop a new approach to this model from the viewpoint of *set-valued optimization*. However, the mainstream developments in multiobjective optimization in the vein presented, e.g., in Chapter 9 cannot be directly applied to welfare and related economic models. In particular, to obtain the most adequate versions of the so-called *second fundamental theorem of welfare economics* (or *marginal price equilibria*), we need to derive necessary optimality conditions for *new types of minimizers* in set-valued optimization, which are inspired by the corresponding notions of *Pareto optimal allocations* adequate for economic modeling. Thus this chapter establishes deep two-sided relationships between economic modeling and set-valued optimization.

10.1 Economic Modeling via Set-Valued Optimization

First we formulate the fundamental model of welfare economics with appropriate notions of Pareto optimal allocations therein and then reduce this economic model to a special problem of set-valued optimization with the corresponding notions of local minimizers.

10.1.1 Models of Welfare Economics

Given a normed *commodity space* E, consider the *economy*

$$\mathcal{E} = (C_1, \ldots, C_n, S_1, \ldots, S_m, W) \tag{10.1}$$

B. S. Mordukhovich, *Variational Analysis and Applications*, Springer Monographs in Mathematics, https://doi.org/10.1007/978-3-319-92775-6_10

involving $m \in \mathbb{N}$ firms with their *production sets* $S_j \subset E$ $(j = 1, \dots, m)$, $n \in \mathbb{N}$ customers with their *consumption sets* $C_i \subset E$ $(i = 1, \dots, n)$, and the *net demand constraint set* W representing constraints related to the initial inventory of commodities in $\mathcal{E}$. Without loss of generality, suppose that all the sets in (10.1) are *locally closed* around the reference points.

Denote production strategies by $y = (y_1, \dots, y_m) \in S_1 \times \dots \times S_m$ and consumption plans by $z = (z_1, \dots, z_n) \in C_1 \times \dots \times C_n$ and say that the pair (y, z) is an *admissible state* of the economy $\mathcal{E}$. Further, associate with each consumer his/her *preference set* $P_i(z)$ that consists of elements in C_i preferred to z_i by this consumer at the consumption plan z. Observe that we cannot assume the local closedness of the preference sets $P_i(z)$, since it contradicts their meaning as extensions of the "$<$" notion to the general setting under consideration. The corresponding preference mappings $P_i : Z \rightrightarrows E$ are *set-valued* with $Z := E^n$. By definition we have $z_i \notin P_i(z)$ for every $i = 1, \dots, n$ and naturally suppose that $P_i(z) \neq \emptyset$ at least for some $i \in \{1, \dots, n\}$. Put for convenience $\operatorname{cl} P_i(z) := \{z_i\}$ if $P_i(z) = \emptyset$.

The market/budget constraints in the economy $\mathcal{E}$ are given as follows.

Definition 10.1 (Feasible Allocations). *An admissible state (y, z) of the economy $\mathcal{E}$ in (10.1) is a* FEASIBLE ALLOCATION *of $\mathcal{E}$ if*

$$w := \sum_{i=1}^{n} z_i - \sum_{j=1}^{m} y_j \in W. \tag{10.2}$$

In the classical case of welfare economics, the set W consists of a single element $W = \{\omega\}$, where ω signifies the initial *aggregate endowment* of scarce resources. In this case, constraint (10.2) reduces to the "markets clear" condition. Another conventional setting of (10.2) is $W = \omega - E_+$, where E_+ is the closed positive cone of the partially ordered commodity space; this corresponds to the "implicit free disposal" of commodities. In the general setting of (10.2), we can interpret W as an *uncertainty region* reflecting *incomplete information* on the initial aggregate endowment value.

Our aim is to study the following notions of Pareto-type *optimal allocations* of the economy $\mathcal{E}$ and support them by certain *price equilibria*.

Definition 10.2 (Pareto-Type Optimal Allocations). *Let $(\bar{y}, \bar{z}) \in E^m \times E^n$ be a feasible allocation of the economy $\mathcal{E}$. We say that:*

(i) *The pair $(\bar{y}, \bar{z})$ is a* LOCAL WEAK PARETO OPTIMAL ALLOCATION *of $\mathcal{E}$ if there is a neighborhood $\mathcal{O} \subset E^m \times E^n$ of $(\bar{y}, \bar{z})$ such that for every feasible allocation $(y, z) \in \mathcal{O}$, we have $z_i \notin P_i(\bar{z})$ for some $i \in \{1, \dots, n\}$.*

(ii) *The pair $(\bar{y}, \bar{z})$ is a* LOCAL PARETO/EFFICIENT OPTIMAL ALLOCATION *of $\mathcal{E}$ if there is a neighborhood $\mathcal{O}$ of $(\bar{y}, \bar{z})$ such that for every feasible allocation $(y, z) \in \mathcal{O}$, we have either $z_i \notin \operatorname{cl} P_i(\bar{z})$ for some index $i \in \{1, \dots, n\}$ or $z_i \notin P_i(\bar{z})$ for all indices $i \in \{1, \dots, n\}$.*

(iii) *The pair* $(\bar{y}, \bar{z})$ *is a* LOCAL STRICT PARETO OPTIMAL ALLOCATION *of* $\mathcal{E}$ *if there is a neighborhood* $\mathcal{O}$ *of* $(\bar{y}, \bar{z})$ *such that for every feasible allocation* $(y, z) \in \mathcal{O}$ *with* $z \neq \bar{z}$*, we have* $z_i \notin \operatorname{cl} P_i(\bar{z})$ *for some* $i \in \{1, \ldots, n\}$.

(iv) *The pair* $(\bar{y}, \bar{z})$ *is a* LOCAL STRONG PARETO OPTIMAL ALLOCATION *of* $\mathcal{E}$ *if there is a neighborhood* $\mathcal{O}$ *of* $(\bar{y}, \bar{z})$ *such that for every feasible allocation* $(y, z) \in \mathcal{O}$ *with* $(y, z) \neq (\bar{y}, \bar{z})$ *we have* $z_i \notin \operatorname{cl} P_i(\bar{z})$ *for some* $i \in \{1, \ldots, n\}$.

(v) *We replace "local" by "*GLOBAL*" in* (i)–(iv) *if* $\mathcal{O} = E^m \times E^n$.

It is clear from the definitions that (iv)⇒(iii)⇒(ii)⇒(i) but not vice versa; the same implications hold for the global version in (v). Note that the notions of (both local and global) *weak Pareto* and *Pareto* optimal allocations are conventional in welfare economics. They correspond to the similar Pareto-type concepts (weakly efficient and efficient solutions) for standard problems of vector optimization in the case of preferences given by utility functions. The notions of *strong Pareto* and *strict Pareto* optimal allocations are less conventional while they have also appeared in models of welfare economics; see more discussions in the text and exercises below and as well as in Section 10.6.

10.1.2 Constrained Set-Valued Optimization

Consider now a problem of *set-valued optimization* with *geometric constraints*:

$$\text{minimize} \quad F(x) \quad \text{subject to} \quad x \in \Omega, \tag{10.3}$$

where the cost mapping $F\colon X \rightrightarrows Z$ is set-valued between Banach spaces, Ω is a subset in X, and "minimization" in (10.3) is understood with respect to some preference relation on Z. We define this preference via a given *preference mapping* $L\colon Z \rightrightarrows Z$ as follows:

$$u \in Z \quad \text{is preferred to} \quad z \quad \text{if and only if} \quad u \in L(z). \tag{10.4}$$

Note that the above preference (10.4) can be equivalently written as the one $\prec$ for which $L\colon Z \rightrightarrows Z$ is the *level-set* mapping

$$L(z) := \big\{ u \in Z \big|\ u \prec z \big\}. \tag{10.5}$$

Next we introduce the notions of *fully localized* optimal solutions to the set-valued optimization problem (10.3) with respect to preference (10.4).

Definition 10.3 (Fully Localized Optimal Solutions to Constrained Multiobjective Problems). *Let* $(\bar{x}, \bar{z}) \in \operatorname{gph} F$ *with* $\bar{x} \in \Omega$. *We say that:*

(i) $(\bar{x}, \bar{z})$ *is a* FULLY LOCALIZED WEAK MINIMIZER *for* (10.3) *if there exist neighborhoods* U *of* $\bar{x}$ *and* V *of* $\bar{z}$ *such that there is no element* $z \in F(\Omega \cap U) \cap V$*, which is preferred to* $\bar{z}$*, i.e.,*

$$F(\Omega \cap U) \cap L(\bar{z}) \cap V = \emptyset. \tag{10.6}$$

(ii) $(\bar{x}, \bar{z})$ *is a* FULLY LOCALIZED MINIMIZER *for* (10.3) *if there exist neighborhoods* U *of* $\bar{x}$ *and* V *of* $\bar{z}$ *such that there is no element* $z \in F(\Omega \cap U) \cap V$ *with* $z \neq \bar{z}$ *and* $z \in \operatorname{cl} L(\bar{z})$*, i.e.,*

$$F(\Omega \cap U) \cap \operatorname{cl} L(\bar{z}) \cap V = \{\bar{z}\}. \tag{10.7}$$

(iii) $(\bar{x}, \bar{z})$ *is a* FULLY LOCALIZED STRONG MINIMIZER *for* (10.3) *if there exist neighborhoods* U *of* $\bar{x}$ *and* V *of* $\bar{z}$ *such that there is no element* $(x, z) \in \operatorname{gph} F \cap (U \times V)$ *with* $(x, z) \neq (\bar{x}, \bar{z})$ *satisfying* $x \in \Omega$ *and* $z \in \operatorname{cl} L(\bar{z})$*, i.e.,*

$$\operatorname{gph} F \cap \big(\Omega \times \operatorname{cl} L(\bar{z})\big) \cap (U \times V) = \{(\bar{x}, \bar{z})\}. \tag{10.8}$$

It is easy to see that (iii)⇒(ii)⇒(i) in Definition 10.3. If $\Omega = X$, we speak about the corresponding fully localized minimizers for the mapping F.

The underlying feature of all the notions in Definition 10.3 is that they reflect the *image localization* of minimizers in constructions (10.6)–(10.8). It provides new information even in the case of single-valued objectives $F = f : X \to Z$ and, in contrast to the optimality notions from Chapter 9, allows us to study *local* Pareto-type optimal allocations of welfare economies introduced in Definition 10.2; see more details below.

10.1.3 Optimal Allocations as Fully Localized Minimizers

Here we associate the model of welfare economics described in Subsection 10.1.1 with a special problem of *set-valued* optimization involving a *level-set* preference relation and *geometric* constraints, which is constructed upon the initial data of the economy $\mathcal{E}$. Then we establish the equivalence between the notions of *local* Pareto-type optimal allocations for $\mathcal{E}$ and *fully localized* optimal solutions to the constructed problem of multiobjective optimization.

Given the economy $\mathcal{E}$, consider the following set-valued optimization problem in form (10.3) with $X = E^{m+1}$, $x = (y, w)$, and $Z := E^n$ given by:

$$\begin{cases} \text{minimize} \;\; F(x) := \Big\{ z \in Z \,\Big|\, w = \displaystyle\sum_{i=1}^{n} z_i - \sum_{j=1}^{m} y_j \Big\} \\ \text{subject to} \;\; x \in \Omega := \displaystyle\prod_{j=1}^{m} S_j \times W \subset X, \end{cases} \tag{10.9}$$

where "minimization" is understood with respect to the preference/level-set mapping $L : Z \rightrightarrows Z$ defined by

$$L(z) := \prod_{i=1}^{n} P_i(z), \quad z \in Z, \tag{10.10}$$

via the preference mappings $P_i : Z \rightrightarrows E$ of the economy $\mathcal{E}$.

Theorem 10.4 (Equivalence Between Local Pareto-Type Optimal Allocations in Welfare Economics and Fully Localized Minimizers in Set-Valued Optimization). *Let $(\bar{y}, \bar{z})$ be a feasible allocation of the welfare economy $\mathcal{E}$ in (10.1) with the preference sets $P_i(z)$, and let $\bar{x} := (\bar{y}, \bar{w})$ with $\bar{w} := \sum_{i=1}^n z_i - \sum_{j=1}^m y_j$. Then we have the following equivalence relationships:*

(i) *$(\bar{y}, \bar{z})$ is a local* WEAK PARETO *optimal allocation of $\mathcal{E}$ if and only if $(\bar{x}, \bar{z})$ is a fully localized* WEAK MINIMIZER *for the multiobjective optimization problem (10.9) with respect to the preference $L\colon Z \rightrightarrows Z$ defined in (10.10).*

(ii) *$(\bar{y}, \bar{z})$ is a local* STRICT PARETO *optimal allocation of $\mathcal{E}$ if and only if $(\bar{x}, \bar{z})$ is a fully localized* MINIMIZER *for (10.9) with respect to $L\colon Z \rightrightarrows Z$.*

(iii) *$(\bar{y}, \bar{z})$ is a local* STRONG PARETO *optimal allocation of $\mathcal{E}$ if and only if it is a fully localized* STRONG MINIMIZER *for (10.9) with respect to $L\colon Z \rightrightarrows Z$.*

Proof. Let us first justify (i). Assuming that $(\bar{y}, \bar{z})$ is a local *weak Pareto* optimal allocation of $\mathcal{E}$ yields by Definition 10.2(i) and by the structure of $L(\cdot)$ the existence of a neighborhood $\mathcal{O} = \mathcal{O}_y \times \mathcal{O}_z$ of $(\bar{y}, \bar{z})$ such that

$$z \notin L(\bar{z}) \ \text{ for all feasible allocations } \ (y, z) \in \mathcal{O}. \tag{10.11}$$

Choosing $V := \mathcal{O}_z$ and $U := \mathcal{O}_y \times \mathcal{O}_w$, where $\mathcal{O}_w := \big\{w = \sum_{i=1}^n z_i - \sum_{j=1}^m y_j \big| \ (y, z) \in \mathcal{O}\big\}$ is a neighborhood of $\bar{w}$, we claim that

$$F(\Omega \cap U) \cap L(\bar{z}) \cap V = \emptyset. \tag{10.12}$$

Indeed, the violation of (10.12) means that there is $z \in F(\Omega \cap U) \cap L(\bar{z}) \cap V$. Taking into account the constructions of F and Ω in (10.9), we find $y \in \prod_{j=1}^m S_j \cap \mathcal{O}_y$ satisfying $w = \displaystyle\sum_{i=1}^n z_i - \sum_{j=1}^m y_j \in W$. This implies that $(y, z) \in \mathcal{O}$ is a feasible allocation of $\mathcal{E}$ with $z \in L(\bar{z})$. The latter clearly contradicts (10.11) and thus verifies (10.12), which means that $(\bar{x}, \bar{z})$ is a fully localized weak minimizer for the set-valued optimization problem (10.9).

Conversely, let $(\bar{x}, \bar{z})$ be a fully localized *weak minimizer* for problem (10.9) with $\bar{x} = (\bar{y}, \bar{w})$. Definition 10.3(i) gives us neighborhoods $U = \mathcal{O}_y \times \mathcal{O}_w$ of $(\bar{y}, \bar{w})$ and V of $\bar{z}$ such that (10.12) holds and that the set $\big\{w = \sum_{i=1}^n z_i - \sum_{j=1}^m y_j \big| \ (y, z) \in \mathcal{O}_y \times V\big\}$ is contained in $\mathcal{O}_w$. For any feasible allocation (y, z) from the neighborhood $\mathcal{O} := \mathcal{O}_y \times V$ of $(\bar{y}, \bar{z})$, we get by Definition 10.1 of feasible allocations and the above choice of the neighborhoods that $(\bar{y}, \bar{w}) \in \Omega \cap U$, $\bar{z} \in F(\bar{y}, \bar{w}) \subset F(\Omega \cap U)$, and

$$z_i \notin P_i(\bar{z}) \ \text{ for some } \ i \in \big\{1, \ldots, n\big\}. \tag{10.13}$$

Indeed, the violation of (10.13) reads that $z_i \in P_i(\bar{z})$ for all $i \in \{1, \ldots, n\}$, which implies by (10.10) that $z \in L(\bar{z})$ and thus $z \in F(\Omega \cap U) \cap L(\bar{z}) \cap V$. The latter surely contradicts (10.12) and hence gives us (10.13) while justifying that $(\bar{y}, \bar{z})$ is a local weak Pareto optimal allocation of $\mathcal{E}$.

Next we verify assertion (ii) of the theorem. Take a local *strict Pareto* optimal allocation of $\mathcal{E}$ and find by Definition 10.2(iii) a neighborhood $\mathcal{O} = \mathcal{O}_y \times \mathcal{O}_z$ of $(\bar{y}, \bar{z})$ such that

$$z \notin \operatorname{cl} L(\bar{z}) \text{ for all feasible allocations } (y, z) \in \mathcal{O} \text{ with } z \neq \bar{z}. \tag{10.14}$$

We claim that (10.7) is satisfied with $U = \mathcal{O}_y \times \mathcal{O}_w$ and $V = \mathcal{O}_z$, where

$$\mathcal{O}_w := \Big\{ w = \sum_{i=1}^n z_i - \sum_{j=1}^m y_j \in E \ \Big| \ (y, z) \in \mathcal{O} \Big\} \tag{10.15}$$

is obviously a neighborhood of $\bar{w}$. Suppose on the contrary that (10.7) is violated and get a pair (x, z) with $z \neq \bar{z}$ that belongs to the set on the left-hand side of (10.7). Taking into account the structures of F and Ω in (10.9), we have $x = (y, w) \in \Omega \cap U$ and $z \in F(x) \cap V$ such that

$$(y, z) \in \mathcal{O}, \ \ y \in \prod_{j=1}^m S_j \times \mathcal{O}_y, \ \text{ and } \ w = \sum_{i=1}^n z_i - \sum_{j=1}^m y_j \in W \tag{10.16}$$

with $z_i \in \operatorname{cl} P_i(\bar{z})$. Since the consumption sets C_i are closed and since $P_i(\bar{z}) \subset C_i$ as $i = 1, \ldots, n$, we get that $(y, z) \in \mathcal{O}$ is a feasible allocation of $\mathcal{E}$. This contradicts (10.14) and thus justifies the claim.

To verify the converse implication in (ii), take an arbitrary fully localized *minimizer* $(\bar{x}, \bar{z})$ for (10.9) with $\bar{x} = (\bar{y}, \bar{w})$ and find neighborhoods $U = \mathcal{O}_y \times \mathcal{O}_w$ of $(\bar{y}, \bar{w})$ and V of $\bar{z}$ such that (10.7) holds with the preference set $L(\bar{z})$ defined in (10.10). We claim that

$$z_i \notin \operatorname{cl} P_i(\bar{z}) \ \text{ for some } \ i \in \{1, \ldots, n\} \tag{10.17}$$

whenever $(y, z) \in \mathcal{O}_y \times V$ is a feasible allocation of $\mathcal{E}$ with $z \neq \bar{z}$. Indeed, the violation of (10.17) and the structure of $L(\cdot)$ yield $z \in \operatorname{cl} L(\bar{z})$. Due to

$$\Big\{ w = \sum_{i=1}^n z_i - \sum_{j=1}^m y_j \Big| \ (y, z) \in \mathcal{O}_y \times V \Big\} \subset \mathcal{O}_w$$

valid for sufficiently small neighborhoods $\mathcal{O}_y$ and V, we have $z \in F(\Omega \cap U) \cap V$. This contradicts (10.7) and thus justifies assertion (ii).

Let us finally verify assertion (iii). Take any local *strong Pareto* optimal allocation $(\bar{y}, \bar{z})$ and find by Definition 10.2(iv) a neighborhood $\mathcal{O} = \mathcal{O}_y \times \mathcal{O}_z$ of $(\bar{y}, \bar{z})$ such that (10.17) holds for every feasible allocation $(y, z) \in \mathcal{O}$ of $\mathcal{E}$ with $(y, z) \neq (\bar{y}, \bar{z})$. Then similarly to the proof of (ii) we claim that the strong minimality condition (10.8) is satisfied with $U = \mathcal{O}_y \times \mathcal{O}_w$ and $V = \mathcal{O}_z$, where $\mathcal{O}_w$ is a neighborhood of $\bar{w}$ from (10.15). Indeed, supposing that (10.8) is violated allows us to find some pair $(x, z) \neq (\bar{x}, \bar{z})$ that belongs to the set on the left-hand side of (10.8). Taking into account the structures of F and Ω in (10.9), we have $x = (y, w) \in \Omega \cap U$ and $z \in F(x) \cap V$ satisfying (10.16). Since $(y, z) \neq (\bar{y}, \bar{z})$, it contradicts (10.17) and thus verifies the claim while arguing as in the proof of assertion (ii).

To finish the proof of (iii), it remains to justify the converse implication therein. Picking an arbitrary fully localized *strong minimizer* $(\bar{x}, \bar{z})$ for (10.9) with $\bar{x} = (\bar{y}, \bar{w})$, find a neighborhood $U = \mathcal{O}_y \times \mathcal{O}_w$ of $(\bar{y}, \bar{w})$ and a neighborhood V of $\bar{z}$ such that condition (10.8) holds with $L(\bar{z})$ from (10.10). Our goal is to show that $(\bar{y}, \bar{z})$ is a local strong Pareto optimal allocation of $\mathcal{E}$, i.e., condition (10.17) is satisfied for any feasible allocation $(y, z) \neq (\bar{y}, \bar{z})$ in some neighborhood $\mathcal{O}$ of $(\bar{y}, \bar{z})$. Indeed, the violation of the latter means that

$$z_i \in \operatorname{cl} P_i(\bar{z}) \quad \text{for all} \quad i = 1, \ldots, n. \tag{10.18}$$

Since (y, z) is a feasible allocation of $\mathcal{E}$ and since $x = (y, w)$ for $w \in W$ from (10.2), we get $(x, z) \in \operatorname{gph} F$ with F defined in (10.9). Furthermore, it follows from (10.18) and the constructions of Ω in (10.9) and of $L(\cdot)$ in (10.10) that

$$(y, z) \in \prod_{j=1}^{m} S_j \times W \quad \text{and} \quad z \in \prod_{i=1}^{n} \operatorname{cl} P_i(\bar{z}) = \operatorname{cl} L(\bar{z}).$$

Combining the latter with $x = (y, w) \in \operatorname{gph} F$ and taking into account the above choice of the neighborhoods, we arrive at the relationships

$$(\bar{x}, \bar{z}) \neq (x, z) \in \operatorname{gph} F \cap \big(\Omega \times \operatorname{cl} L(\bar{z})\big) \cap (U \times V),$$

which clearly contradict (10.8) and thus justify (10.17) for all the feasible allocations $(y, z) \neq (\bar{y}, \bar{z})$ from the neighborhood $\mathcal{O}$ of $(\bar{y}, \bar{z})$. This verifies (iii) and thus completes the proof of the theorem. $\triangle$

Note that Theorem 10.4 doesn't reveal a local notion of optimal solutions in multiobjective optimization equivalent to local Pareto optimal allocations in the welfare economics. Hence the study of the latter economic concept at the *local* level remains an *open question* within the (fully localized) multiobjective optimization approach developed below. Nevertheless, we are able to accomplish this on the *global* level as shown in Section 10.4.

10.2 Optimality Conditions with Full Localization

This section concerns constrained set-valued optimization problems of type (10.3) with respect to the level-set preference relations (10.4) and studies them for their own sake. Besides the different types of preferences in comparison with the multiobjective problems investigated in Chapter 9, the major distinction between optimization problems considered here and in the previous chapter is the *fully localized* nature of minimizers studied below. This is certainly of its independent interest while being largely motivated by the subsequent applications to welfare economics due to the material of Section 10.1.

The main goal of this section is to establish necessary optimality conditions for all the three types of fully localized minimizers of problem (10.3) introduced in Definition 10.3. To proceed, we present first some required pieces of variational analysis and generalized differentiation, which have not been considered previously in this book.

10.2.1 Exact Extremal Principle in Product Spaces

Dealing with constrained set-valued optimization problems of type (10.3) and having in mind their specifications of type (10.9) needed for the subsequent applications to welfare economics, we face an unavoidable *product structure* of the cost mapping F and the constraint set Ω in (10.3). This requires considering the following *product versions* of the basic *partial sequential normal compactness* (PSNC) property. Since our applications concern closed sets in products of Asplund spaces, we confine ourselves just to this setting without mentioning it explicitly in the formulations.

Definition 10.5 (PSNC and Strong PSNC Properties in Product Spaces). *Given a set $\Omega \subset X$ in the product space $X = \prod_{i=1}^n X_i$ and given a point $\bar{x} \in \Omega$, we say that:*

(i) Ω *is* PSNC *at $\bar{x} \in \Omega$ with respect to $\{X_i|\ i \in I\}$ as $I \subset \{1, \ldots, n\}$ (or simply with respect to the indices I) if for any sequences of $(x_k, x_k^*) \in X \times X^*$ with $x_k = (x_{1k}, \ldots, x_{nk})$ and $x_k^* = (x_{1k}^*, \ldots, x_{nk}^*)$ satisfying*

$$x_k \stackrel{\Omega}{\to} \bar{x} \ \text{ and } \ x_k^* \in \widehat{N}(x_k; \Omega) \ \text{ for all } \ k \in \mathbb{N} \tag{10.19}$$

we have the implication

$$\big[x_{ik}^* \stackrel{w^*}{\to} 0,\ i \in I,\ \|x_{ik}^*\| \to 0,\ i \in \{1, \ldots, n\} \setminus I\big] \Longrightarrow \|x_{ik}^*\| \to 0,\ i \in I.$$

(ii) Ω *is* STRONG PSNC *at $\bar{x}$ with respect to $\{X_i|\ i \in I\}$ as $I \subset \{1, \ldots, n\}$ if for any sequences (x_k, x_k^*) satisfying (10.19) we have*

$$\big[x_{ik}^* \stackrel{w^*}{\to} 0,\ i \in \{1, \ldots, n\}\big] \Longrightarrow \|x_{ik}^*\| \to 0,\ i \in I.$$

Observe that in the extreme case of $I = \{1, \ldots, n\}$ both PSNC and strong PSNC properties introduced don't depend on the product structure and reduce to the SNC property of Ω at $\bar{x}$; see (2.41). Note also that the general product version of PSNC from Definition 10.5(i) surely agrees with its specifications for mappings (3.65) as well as for products of two sets in Exercise 3.69. The reader can find various effective conditions ensuring the validity of such properties and also their calculus/preservation rules in the corresponding exercise and commentary sections of this book as well as in [522, 523].

The main variational instrument to deal with deriving necessary optimality conditions for fully localized solutions to the multiobjective problem (10.3) with a product structure of constraints is the following *product version* of the (exact) *extremal principle* formulated at the locally extremal point in question. Since the *local closedness* operation is an issue in applications to welfare economics, we don't take this property for granted as before and always formulate it explicitly when needed in the rest of this chapter.

Lemma 10.6 (Product Extremal Principle). *Let $\bar{x}$ be a local extremal point of the set system $\{\Omega_1, \Omega_2\}$, where both Ω_1 and Ω_2 are locally closed around $\bar{x}$ in the product $\prod_{i=1}^n X_i$ of the Asplund spaces X_i, $i = 1, \ldots, n$. Take two index sets $I, J \subset \{1, \ldots, n\}$ with $I \cup J = \{1, \ldots, n\}$ and suppose that either one of the PSNC conditions below is satisfied for Ω_1 and Ω_2:*

- *The set Ω_1 is PSNC at $\bar{x}$ with respect to I while the set Ω_2 is strongly PSNC at $\bar{x}$ with respect to J.*
- *The set Ω_1 is strongly PSNC at $\bar{x}$ with respect to I while the set Ω_2 is PSNC at $\bar{x}$ with respect to J.*

Then there is a dual element $x^* \in X^*$ such that

$$0 \neq x^* \in N(\bar{x}; \Omega_1) \cap \big(-N(\bar{x}; \Omega_2)\big). \tag{10.20}$$

This lemma extends [523, Lemma 5.58] to the case of finitely many spaces in the product without the requirement that $I \cap J = \emptyset$ as in [523], while the proof of the updated version proceeds in the same way; see Exercise 10.27.

10.2.2 Asymptotic Closedness of Sets

Another crucial ingredient in deriving necessary optimality conditions for fully localized minimizers is the *asymptotic closedness property* of appropriate sets that allows us to reduce the notions of fully localized minimizers considered for (10.3) to *locally extremal points* of the corresponding systems of sets. Here is this property in the abstract framework of linear topological spaces.

Definition 10.7 (Asymptotic Closedness Property). *A set $\Xi \subset Z$ is* ASYMPTOTICALLY CLOSED *at $\bar{z} \in \mathrm{cl}\, \Xi$ if there exists a neighborhood V of $\bar{z}$ such that for any $\varepsilon > 0$ we can find $c \in \varepsilon\mathbb{B}$ satisfying*

$$(\mathrm{cl}\, \Xi + c) \cap V \subset \Xi \backslash \big\{\bar{z}\big\}. \tag{10.21}$$

This property (which is fully independent of the local closedness of a set) holds in many fairly general settings and is satisfied under natural assumptions in models of welfare economics; see Fig. 10.1 and Section 10.3 below.

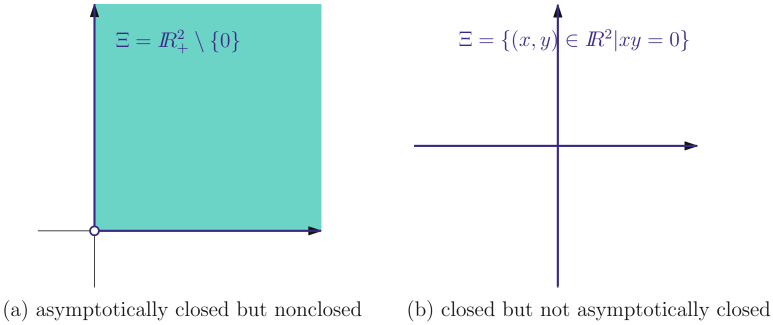


Fig. 10.1 Asymptotic closedness vs. closedness of sets.

Let us list some sufficient conditions ensuring the asymptotic closedness property for remarkable classes of sets while leaving their verifications as exercises to the reader; see Section 10.5:

(i) Every proper, convex, and solid subcone $\Xi \subset Z$ and its nonconvex complement $Z \backslash \Xi$ have the asymptotic closeness property at $0 \in Z$.

(ii) Every closed, convex, and pointed cone $\Xi \subset Z$ with $\Xi \setminus (-\Xi) \neq \emptyset$ has the asymptotic closedness property at $0 \in Z$.

(iii) The epigraph of an extended-real-valued function $\varphi\colon X \to \overline{\mathbb{R}}$ has the asymptotic closedness property at $(\bar{x}, \varphi(\bar{x}))$ if φ is l.s.c. around $\bar{x}$.

Note that the condition $\Xi \setminus (-\Xi) \neq \emptyset$ in (iii) says that the cone Ξ is *not a linear subspace* of Z; this is more general than the standard pointedness requirement $\Xi \cap (-\Xi) = \{0\}$, which means that Ξ *doesn't contain* a linear subspace. As mentioned in Exercise 9.32(i), the generalized Pareto optimality (9.84) induced by a closed and convex cone Θ doesn't correspond to a closed preference relation (defined therein) unless the ordering cone Θ is pointed; the latter is not required by the asymptotic closedness property.

The next result needed in what follows ensures the asymptotic closedness property of product sets via this property of their selected components.

Proposition 10.8 (Asymptotic Closedness Property of Product Sets). *Let $\bar{z} \in \text{cl} \prod_{i=1}^n \Xi_i \subset \prod_{i=1}^n Z_i$ in the normed space setting, let $I \subset \{1, \ldots, n\}$ be a nonempty index set, and let $J := \{1, \ldots, n\} \backslash I$. Suppose that the sets Ξ_i are asymptotically closed at $\bar{z}_i \in \text{cl}\, \Xi_i$ for $i \in I$ while the other sets Ξ_j are locally closed around $\bar{z}_j$ for $j \in J$. Then the product set $\Xi := \prod_{i=1}^n \Xi_i$ enjoys the asymptotic closedness property at $\bar{z}$.*

Proof. Without loss of generality, assume that $I = \{1, \ldots, m\}$ with some $0 < m \leq n$. Since for each $i \in I$ the set Ξ_i is asymptotically closed at $\bar{z}_i$, there is a neighborhood U_i of $\bar{z}_i$ such that, whenever $\varepsilon > 0$, we have

$$(\text{cl}\, \Xi_i + c_i) \cap U_i \subset \Xi_i \backslash \{\bar{z}_i\} \ \text{ with some } \ c_i \in \varepsilon \mathbb{B}_{Z_i}, \ i \in I.$$

On the other hand, by the assumed local closedness of Ξ_j around $\bar{z}_j$, for each $j \in J$ we find a neighborhood U_j of $\bar{z}_j$ such that

$$\mathrm{cl}\,\Xi_j \cap U_j \subset \Omega_j, \quad j \in J.$$

It is obvious that the set $U := \prod_{i\in I} U_i \times \prod_{j\in J} U_j$ is a neighborhood of $\bar{z}$ in the product space $Z := \prod_{i=1}^n Z_i$ equipped with the maximum norm. Furthermore, for any number $\varepsilon > 0$ there is $c := (c_1, \ldots, c_m, 0, \ldots, 0) \in \varepsilon\mathbb{B}_Z$ satisfying

$$\big(\mathrm{cl}\,\Xi + c\big) \cap U = \Big(\prod_{i\in I}(\mathrm{cl}\,\Xi_i + c_i) \cap U_i\Big) \times \Big(\prod_{j\in J}(\mathrm{cl}\,\Xi_j \cap U_j)\Big)$$

$$\subset \Big(\prod_{i\in I}\big(\Xi_i\backslash\{\bar{z}_i\}\big)\Big) \times \Big(\prod_{j\in J}\Xi_j\Big) \subset \Big(\prod_{i=1}^n \Xi_i\Big)\backslash\{\bar{z}\} = \Xi\backslash\{\bar{z}\},$$

where the last inclusion holds due to $I \neq \emptyset$. This gives us (10.21) and thus justifies the asymptotic closedness property of the product Ξ at $\bar{z}$. △

10.2.3 Necessary Conditions for Localized Minimizers

Now we are ready to derive necessary conditions for all the types of fully localized minimizers in Definition 10.3 for the multiobjective problem (10.3).

Theorem 10.9 (Necessary Conditions for Fully Localized Minimizers in Constrained Set-Valued Optimization). *Let $F\colon X \rightrightarrows Z$ be a set-valued mapping between Asplund spaces with the graph* gph F *locally closed around some point* $(\bar{x}, \bar{z}) \in \mathrm{gph}\,F$, *let $\Omega \subset X$ be locally closed around $\bar{x}$, and let $L\colon Z \rightrightarrows Z$ be a preference mapping on Z locally satiated at $\bar{z}$ in the sense formulated in Exercise 9.32. Impose also* EITHER ONE *of the following SNC assumptions on the initial data (F, Ω, L) in* (10.3), (10.4)*:*

(a) gph F *is SNC at* $(\bar{x}, \bar{z})$*;*
(b) Ω *is SNC at $\bar{x}$ and* cl $L(\bar{z})$ *is SNC at $\bar{z}$;*
(c) F *is PSNC at* $(\bar{x}, \bar{z})$ *and* cl $L(\bar{z})$ *is SNC at $\bar{z}$;*
(d) Ω *is SNC at $\bar{x}$ and F^{-1} is PSNC at* $(\bar{z}, \bar{x})$.

Then there is a pair $(0, 0) \neq (x^*, z^*) \in X^* \times Z^*$ *satisfying the necessary optimality conditions*

$$x^* \in D^*F(\bar{x}, \bar{z})(z^*) \cap \big(-N(\bar{x}; \Omega)\big) \;\; \text{and} \;\; z^* \in N\big(\bar{z}; \mathrm{cl}\,L(\bar{z})\big) \tag{10.22}$$

in each of the following cases:

• $(\bar{x}, \bar{z})$ *is a* FULLY LOCALIZED WEAK MINIMIZER *for* (10.3), (10.4) *provided that the set $L(\bar{z})$ is asymptotically closed at $\bar{z} \in \mathrm{cl}\,L(\bar{z})$;*

• $(\bar{x}, \bar{z})$ *is a* FULLY LOCALIZED MINIMIZER *for* (10.3), (10.4) *provided that the set* cl $L(\bar{z})$ *is asymptotically closed at $\bar{z}$;*

• $(\bar{x}, \bar{z})$ *is a* FULLY LOCALIZED STRONG MINIMIZER *for* (10.3), (10.4) *provided that either* $\mathrm{cl}\, L(\bar{z})$ *or* Ω *is asymptotically closed at* $\bar{z}$ *or* $\bar{x}$*, respectively.*

Proof. Arguing in a unifying way, take any fully localized minimizer $(\bar{x}, \bar{z})$ for problem (10.3), (10.4) considered in the theorem and reduce it to a local extremal point of some system of sets in the product space $X \times Z$. Indeed, define the sets $\Omega_1, \Omega_2 \subset X \times Z$ by

$$\Omega_1 := \mathrm{gph}\, F \quad \text{and} \quad \Omega_2 := \Omega \times \mathrm{cl}\, L(\bar{z}) \tag{10.23}$$

and observe that they are locally closed around $(\bar{x}, \bar{z})$ in the Asplund space $X \times Z$ under the local closedness assumptions made. Let us check that $(\bar{x}, \bar{z})$ is a locally extremal point of the set system $\{\Omega_1, \Omega_2\}$ in each case of the fully localized minimizers under consideration. Observe first that $(\bar{x}, \bar{z}) \in \Omega_1 \cap \Omega_2$ due to the imposed local satiation property of the preference mapping. Let us show next that there is a sequence $\{a_k\} \subset X \times Z$ with $a_k \to 0$ as $k \to \infty$ with

$$\Omega_1 \cap (\Omega_2 + a_k) \cap \mathcal{O} = \emptyset \;\; \text{for all} \;\; k \in \mathbb{N}, \tag{10.24}$$

where $\mathcal{O}$ is a neighborhood of $(\bar{x}, \bar{z})$ specified later. We choose an appropriate sequence $\{a0_k\}$ in (10.24) in the following way for each type of the fully localized minimizers considered in the theorem:

• Let $(\bar{x}, \bar{z})$ be a fully localized *weak minimizer* for problem (10.3), (10.4). Since $L(\bar{z})$ is assumed to be asymptotically closed at $\bar{z}$ in this case, there exist a neighborhood $\widetilde{V}$ of $\bar{z}$ and a sequence $\{c_k\} \subset Z$ with $c_k \to 0$ such that

$$\big(\mathrm{cl}\, L(\bar{z}) + c_k\big) \cap \widetilde{V} \subset L(\bar{z}) \setminus \{\bar{z}\} = L(\bar{z}). \tag{10.25}$$

Put $a_k := (0, c_k) \in X \times Z$ for all $k \in \mathbb{N}$ and $\mathcal{O} := U \times (V \cap \widetilde{V})$, where U is a neighborhood of $\bar{x}$ and V is a neighborhood of $\bar{z}$ from definition (10.6) of the localized weak minimality of $(\bar{x}, \bar{z})$. Then we get from (10.25) and the structures of Ω_1, Ω_2 in (10.23) that

$$\begin{aligned} &\Omega_1 \cap (\Omega_2 + a_k) \cap \mathcal{O} \\ &= \mathrm{gph}\, F \cap \Big(\Omega \times (\mathrm{cl}\, L(\bar{z}) + c_k) \cap \widetilde{V}\Big) \cap (U \times V) \\ &\subset \mathrm{gph}\, F \cap \Big(\Omega \times (L(\bar{z}) \setminus \{\bar{z}\})\Big) \cap (U \times V) = \emptyset, \end{aligned} \tag{10.26}$$

where the last equality is due to (10.6). This justifies the extremality condition (10.24) in the case of fully localized weak minimizers.

• Let $(\bar{x}, \bar{z})$ is a fully localized *minimizer* for problem (10.3), (10.4). In this case we use the same arguments as for weak minimizers above replacing now the set $L(\bar{z})$ in (10.25) and in the last line of (10.26) by its closure $\mathrm{cl}\, L(\bar{z})$. This can be done, since the set $\mathrm{cl}\, L(\bar{z})$ is assumed to be asymptotically closed at $\bar{z}$. Thus we get (10.24) in the case of fully localized minimizers.

• Let $(\bar{x}, \bar{z})$ is a fully localized *strong minimizer* for problem (10.3), (10.4). Applying Proposition 10.8 to the product set $\Omega_2 = \Omega \times \mathrm{cl}\, L(\bar{z})$, we get from the imposed assumptions on the sets Ω and $\mathrm{cl}\, L(\bar{z})$ in this case that Ω_2 is asymptotically closed at $(\bar{x}, \bar{z})$. Hence there are a neighborhood $\mathcal{O}$ of $(\bar{x}, \bar{z})$ (without loss of generality we suppose that $\mathcal{O} \subset U \times V$) and a sequence $\{a_k\} \subset X \times Z$ with $a_k \to 0$ as $k \to \infty$ such that

$$(\Omega_2 + a_k) \cap \mathcal{O} \subset \Omega_2 \setminus \big\{(\bar{x}, \bar{z})\big\}.$$

The latter readily implies that

$$\Omega_1 \cap (\Omega_2 + a_k) \cap \mathcal{O} = \Omega_1 \cap \big((\Omega_2 + a_k) \cap \mathcal{O}\big) \cap \mathcal{O}$$

$$\subset \Omega_1 \cap \big(\Omega_2 \backslash \{(\bar{x}, \bar{z})\}\big) \cap (U \times V)$$

$$= \mathrm{gph}\, F \cap \big((\Omega \times \mathrm{cl}\, L(\bar{z})) \backslash \{(\bar{x}, \bar{z})\}\big) \cap (U \times V) = \emptyset,$$

where the last equality holds by (10.8). This justifies the extremality condition (10.24) in the strong minimum case and shows that $(\bar{x}, \bar{z})$ is a locally extremal point of the set system $\{\Omega_1, \Omega_2\}$ in all the cases under consideration.

Now we can apply the *product extremal principle* of Lemma 10.6 to the system $\{\Omega_1, \Omega_2\}$ from (10.23) at the local extremal point $(\bar{x}, \bar{z})$ in the product space $X \times Z$. Observe that each of the SNC/PSNC conditions (a)–(d) imposed in the theorem ensures the fulfillment of the PSNC requirements imposed on the sets Ω_1 and Ω_2 in Lemma 10.6. Indeed, denoting $X_1 := X$ and $X_2 := Z$ therein, we have the following relationships:

• Ω_1 is strongly PSNC at $(\bar{x}, \bar{z})$ with respect to $I = \{1, 2\}$ if (a) holds;
• Ω_2 is strongly PSNC at $(\bar{x}, \bar{z})$ with respect to $J = \{1, 2\}$ if (b) holds;
• Ω_1 is PSNC at $(\bar{x}, \bar{z})$ with respect to $I = \{1\}$ and Ω_2 is strongly PSNC at $(\bar{x}, \bar{z})$ with respect to $J = \{2\}$ if (c) holds;
• Ω_1 is PSNC at $(\bar{x}, \bar{z})$ with respect to $I = \{2\}$ and Ω_2 is strongly PSNC at $(\bar{x}, \bar{z})$ with respect to $J = \{1\}$ if (d) holds.

Thus the optimality conditions in (10.22) follow from (10.20) due to the coderivative definition and the normal cone product rule from Proposition 1.4. This justifies the existence of $(x^*, z^*) \in X^* \times Z^*$ with $(x^*, z^*) \neq 0$ satisfying

$$(-x^*, z^*) \in N\big((\bar{x}, \bar{z}); \Omega_2\big) = N(\bar{x}; \Omega) \times N\big(\bar{z}; \mathrm{cl}\, L(\bar{z})\big),$$

$$(x^*, -z^*) \in N\big((\bar{x}, \bar{z}); \mathrm{gph}\, F\big), \quad \text{i.e.,} \quad x^* \in D^*F(\bar{x}, \bar{z})(z^*),$$

which therefore completes the proof of the theorem. △

As we know from Exercises 3.42 and 3.48 based on [522, Theorems 4.10 and 4.18], the PSNC assumption on F in (c) and on F^{-1} in (d) imposed in Theorem 10.9 holds automatically if F is *Lipschitz-like* around $(\bar{x}, \bar{z})$ and if F^{-1} is *metrically regular* around $(\bar{z}, \bar{x})$, respectively. Note also that it is not hard to formulate and justify modifications of both assumptions and conclusions in Theorem 10.9 in the case where the spaces X and Z therein are given as products of finitely many Asplund spaces; see Exercise 10.29.

10.3 Local Extended Second Welfare Theorems

This subsection presents fargoing extensions of the so-called *second fundamental theorem of welfare economics*, which ensures the existence of *marginal prices* supporting local Pareto-type optimal allocations; see Section 10.6 for more discussions, historical comments, and references. Here we establish extended versions of the second welfare theorem for *local strong, strict*, and *weak* Pareto optimal allocations of the nonconvex economy $\mathcal{E}$ from (10.1) by using their reduction to the appropriate fully localized minimizers of the constrained set-valued optimization problem (10.3), (10.4) and the necessary optimality conditions for such minimizers obtained in Theorem 10.9 under the asymptotic closedness property of the corresponding sets.

10.3.1 Results in General Commodity Spaces

First we establish general versions of the extended second welfare theorem without ordering structures of commodities in the economy $\mathcal{E}$.

Theorem 10.10 (Extended Second Welfare Theorems for Local Pareto-Type Optimal Allocations). *Let $(\bar{y}, \bar{z})$ be a local optimal allocation of economy (10.1) in the senses listed below with respect to the preference sets $P_i(z)$ under the local satiation requirement:*

$$\bar{z}_i \in \operatorname{cl} P_i(\bar{z}) \ \textit{for all} \ i = 1, \ldots, n. \tag{10.27}$$

Assume that the commodity space E is Asplund and that one of the sets

$$\operatorname{cl} P_i(\bar{z}),\ i = 1,\ \ldots, n, \quad S_j,\ j = 1, \ldots, m, \ \ \textit{and} \ \ W \tag{10.28}$$

is SNC at $\bar{z}_i$, $\bar{y}_j$, and $\bar{w} = \sum_{i=1}^n \bar{z}_i - \sum_{j=1}^m \bar{y}_j$, respectively. Then there exists a nonzero marginal price $p^* \in E^*$ satisfying the conditions

$$\begin{cases} -p^* \in N\big(\bar{x}_i; \operatorname{cl} P_i(\bar{z})\big), & i = 1, \ldots, n, \\ p^* \in N(\bar{y}_j; S_j), & j = 1, \ldots, m, \\ p^* \in N(\bar{w}; W) \end{cases} \tag{10.29}$$

in each of the following cases of local optimal allocations of the economy $\mathcal{E}$:

• $(\bar{y}, \bar{z})$ *is a* LOCAL WEAK PARETO OPTIMAL ALLOCATION *provided that the sets $P_i(\bar{z})$, $i = 1, \ldots, n$ are asymptotically closed at $\bar{z}_i$;*

• $(\bar{y}, \bar{z})$ *is a* LOCAL STRICT PARETO OPTIMAL ALLOCATION *provided that there is* $i \in \{1, \ldots, n\}$ *such that the set* $\text{cl}\, P_i(\bar{z})$ *is asymptotically closed at* $\bar{z}_i$*;*

• $(\bar{y}, \bar{z})$ *is a* LOCAL STRONG PARETO OPTIMAL ALLOCATION *provided that* ONE *of sets in* (10.28) *is asymptotically closed at the corresponding point.*

Proof. Observe first that the satiation property of the preference mapping in Theorem 10.9 reduces, for the preference $L(\cdot)$ defined in (10.10), to the local satiation requirement (10.27) on $P_i(\cdot)$ imposed in this theorem. Using further the equivalence relationships of Theorem 10.4 between the local Pareto-type optimal allocations of the economy $\mathcal{E}$ and the fully localized minimizers for the set-valued optimization problem (10.9) with respect to the preference therein and then specifying Proposition 10.8 for the asymptotic closedness property of the corresponding product sets in the welfare economy, the three statements of the theorem reduce to the following ones:

• $(\bar{x}, \bar{z})$ is a *fully localized weak minimizer* for (10.9) with respect to (10.10) provided that the set $L(\bar{z})$ is asymptotically closed at $\bar{z}$;

• $(\bar{x}, \bar{z})$ is a *fully localized Pareto minimizer* for (10.9) with respect to (10.10) provided that the set $\text{cl}\, L(\bar{z})$ is asymptotically closed at $\bar{z}$;

• $(\bar{x}, \bar{z})$ is a *fully localized strong Pareto minimizer* for (10.9) with respect to (10.10) provided that the set

$$\Omega \times \text{cl}\, L(\bar{z}) := \prod_{j=1}^{m} S_j \times W \times \prod_{i=1}^{n} \text{cl}\, P_i(\bar{z})$$

is asymptotically closed at $(\bar{x}, \bar{z}) = (\bar{y}, \bar{w}, \bar{z})$.

To deduce the statements above from the necessary optimality conditions of Theorem 10.9 applied to problem (10.9) with preference (10.10), it remains to check the validity of the corresponding PSNC properties from Exercise 10.29 imposed on the sets

$$\Omega_1 := \text{gph}\, F \quad \text{and} \quad \Omega_2 := \prod_{j=1}^{m} S_j \times W \times \prod_{i=1}^{n} \text{cl}\, P_i(\bar{z}) \tag{10.30}$$

under the SNC assumption on one of the sets in (10.28) made in the theorem. To proceed, we rename the sets and the reference points as follows:

$$\prod_{i=1}^{m+n+1} X_i := X \times Z = E^{m+n+1} \;\text{ with }\; X_i := E \;\text{ for }\; i = 1, \ldots, m+n+1,$$
$$\Theta_i := S_i \text{ for } i = 1, \ldots, m,\; \Theta_{m+1} := W,\; \Theta_{m+1+i} := \text{cl}\, P_i(\bar{z}) \text{ for } i = 1, \ldots, n,$$
$$\bar{x}_1 := -\bar{y}_1, \ldots, \bar{x}_m := -\bar{y}_m,\; \bar{x}_{m+1} := -\bar{w},\; \bar{x}_{m+2} := \bar{z}_1, \ldots,\; \bar{x}_{m+n+1} := \bar{z}_n.$$

Since one of the sets S_j, W, and $\text{cl}\, P_i(\bar{z})$ is SNC at $\bar{y}_j$, $\bar{w}$, and $\bar{z}_i$, respectively, there is $i_0 \in \{1, \ldots, m+n+1\}$ such that Ω_2 from (10.30) is represented as

$$\Omega_2 = \prod_{i=1}^{m+n+1} \Theta_i,$$

and it is *strongly PSNC* with respect to the index set $I := \{i_0\}$. Consider the mapping $f\colon \prod_{i=1}^{i_0-1} X_i \times \prod_{i_0+1}^{m+n+1} X_i \to X_{i_0}$ defined by

$$f(x_1, \ldots, x_{i_0-1}, x_{i_0+1}, \ldots, x_{m+n+1}) := -\sum_{i=1,\, i\neq i_0}^{m+n+1} x_i.$$

It follows from the structures of f and F in (10.9) that the set Ω_1 in (10.30) is represented as the collections of $(-x_1, \ldots, -x_{m+1}, x_{m+2}, \ldots, x_{m+n+1})$ with

$$(x_1, \ldots, x_{i_0-1}, x_{i_0+1}, \ldots, x_{m+n+1}, x_{i_0}) \in \operatorname{gph} f.$$

Since f is Lipschitzian, it is PSNC at $(\bar{x}_1, \ldots, \bar{x}_{i_0-1}, \bar{x}_{i_0+1}, \ldots, \bar{x}_{m+n+1}, \bar{x}_{i_0})$, and hence the set Ω_1 is PSNC at $(\bar{y}, \bar{w}, \bar{z})$ with respect to the index set $J := \{1, \ldots, m+n+1\}\backslash\{i_0\}$. Observing that $I \cup J = \{1, \ldots, m+n+1\}$ ensures that the PSNC assumptions of Exercise 10.29 are satisfied for the sets Ω_1 and Ω_2 in (10.30), and thus we can apply the necessary optimality conditions of Theorem 10.9 to the set-valued optimization problem (10.9) with respect to preference (10.10). It follows in this way by taking into account the structures of (10.9), (10.10), and (10.30) that there are $(0, 0) \neq (x^*, z^*) \in X^* \times Z^*$ and $p^* \in E^*$ satisfying the relationships

$$(-x^*, z^*) \in N\big((\bar{x}, \bar{z}); \Omega_2\big) = \prod_{j=1}^{m} N(\bar{y}_j; S_j) \times N(\bar{w}; W) \times \prod_{i=1}^{n} N\big(\bar{z}_i; \operatorname{cl} P_i(\bar{z})\big),$$
$$(x^*, -z^*) \in N\big((\bar{x}, \bar{z}); \Omega_1\big) = \prod_{j=1}^{m}\{-p^*\} \times \{-p^*\} \times \prod_{i=1}^{n}\{p^*\},$$

where the latter obviously yields $p^* \neq 0$. Thus we arrive at all the price conditions (10.29) of the theorem and so complete its proof. $\triangle$

10.3.2 Ordered Commodity Spaces

Now we consider economies with commodity spaces E that are partially *ordered* by their closed positive cones

$$E_+ := \big\{e \in E \big|\; 0 \preceq e\big\} \tag{10.31}$$

via some partial ordering relation $\preceq$. The *dual positive cone* is defined by

$$E_+^* := \big\{e^* \in E^* \big|\; \langle e, e^*\rangle \geq 0 \;\text{ for all }\; e \in E_+\big\}.$$

Recall that the ordering cone E_+ is *generating* in E if $E_+ - E_+ = E$.

The next useful result reveals a rather general setting in partially ordered spaces when the asymptotic closedness property holds.

Proposition 10.11 (Asymptotic Closedness Property in Ordered Banach Spaces). *Let E be an ordered Banach space with the generating closed positive cone E_+, let Ξ be a closed subset of E satisfying the condition*

$$\Xi - E_+ \subset \Xi, \tag{10.32}$$

and let $\bar{z} \in \operatorname{bd} \Xi$. Then the set Ξ is asymptotically closed at $\bar{z}$.

Proof. Since $\bar{z}$ is a boundary point of Ξ, we find a sequence $\{z_k\} \subset E$ with $z_k \to 0$ as $k \to \infty$ and $\bar{z} + z_k \notin \Xi$ for all $k \in \mathbb{N}$. The classical Krein-Šmulian theorem in ordered spaces with generating positive cones ensures the existence of a constant $M > 0$ such that for each $e \in E$ there are

$$u, v \in E_+ \;\text{ with }\; e = u - v \;\text{ and }\; \max\big\{\|u\|, \|v\|\big\} \le M\,\|e\|.$$

Hence, we get two sequences $\{u_k\} \subset E$ and $\{v_k\} \subset E$ satisfying

$$z_k = u_k - v_k,\; u_k \to 0, \;\text{ and }\; v_k \to 0 \;\text{ as }\; k \to \infty.$$

To justify the asymptotic closedness property of Ξ at $\bar{z}$, it suffices to show that condition (10.32) implies that $\bar{z} \notin \Xi - u_k$ for all large $k \in \mathbb{N}$. Arguing by contradiction, fix $k \in \mathbb{N}$ and suppose that there is some $z \in \Xi$ such that $\bar{z} = z - u_k$. This yields the relationships

$$z = \bar{z} + u_k = \bar{z} + z_k + v_k \in \Xi, \;\text{ i.e., }\; \bar{z} + z_k = z - v_k \in \Xi - E_+ \subset \Xi,$$

which contradict the choice of $\{z_k\}$ and thus completes the proof. △

According to the conventional terminology in microeconomics, we say that the economy $\mathcal{E}$ with the preference sets $P_i(z)$ exhibits:

- the *implicit free disposal of commodities* if

$$\operatorname{cl} W - E_+ \subset \operatorname{cl} W; \tag{10.33}$$

- the *free disposal of production* if

$$\operatorname{cl} S_j - E_+ \subset \operatorname{cl} S_j \;\text{ for some }\; j \in \{1, \ldots, m\}; \tag{10.34}$$

- the *desirability condition* if

$$\operatorname{cl} P_i(\bar{z}) + E_+ \subset \operatorname{cl} P_i(\bar{z}) \;\text{ for some }\; i \in \{1, \ldots, n\}. \tag{10.35}$$

The following consequence of Theorem 10.10 and Proposition 10.11 gives us effective implementations of marginal price equilibrium conditions for local *strict Pareto* and *strong Pareto* optimal allocations of economies with ordered commodity spaces while ensuring in addition the *price positivity*.

Corollary 10.12 (Extended Second Welfare Theorem for Local Strict Pareto and Strong Pareto Optimal Allocations with Ordered Commodities). *In addition to the general assumptions of Theorem 10.10, suppose that the commodity space E is partially ordered by the generating closed positive cone E_+. Then there is a positive price $p^* \in E^*_+ \setminus \{0\}$ satisfying relationships (10.29) in each of the following cases:*

• $(\bar{y}, \bar{z})$ *is a* LOCAL STRICT PARETO OPTIMAL ALLOCATION *of the economy $\mathcal{E}$ exhibiting the desirability condition (10.35) with respect to its preferences.*

• $(\bar{y}, \bar{z})$ *is a* LOCAL STRONG PARETO OPTIMAL ALLOCATION *of $\mathcal{E}$ exhibiting either the implicit free disposal of commodities (10.33), or the free disposal of production (10.34), or the desirability condition (10.35).*

Proof. Observe first that the price *positivity* $p^* \in E^*_+$ in ordered commodity spaces follows directly from assertions (10.29) of the extended second welfare theorem under the fulfillment of either one of the underlying conditions (10.33)–(10.35); see Exercise 10.36. Employing further Proposition 10.11 and the corresponding statements of Theorem 10.10, we arrive at both conclusions claimed in the corollary by showing that:

(i) There is a consumer index $i \in \{1, \ldots, n\}$ such that the associated component $\bar{z}_i$ of $(\bar{y}, \bar{z})$ is a *boundary point* of the set $\mathrm{cl}\, P_i(\bar{z})$ provided that $(\bar{y}, \bar{z})$ is a local *strict* Pareto optimal allocation of the economy $\mathcal{E}$ exhibiting the *desirability condition* (10.35).

(ii) Each component $\bar{z}_i$, $\bar{y}_j$ of $(\bar{y}, \bar{z})$ and their combination $\bar{w} = \sum_{i=1}^n \bar{z}_i - \sum_{j=1}^m \bar{y}_j$ is a *boundary point* of the corresponding set $\mathrm{cl}\, P_i(\bar{z})$, S_j, W provided that $(\bar{y}, \bar{z})$ is a local *strong* Pareto optimal allocation of $\mathcal{E}$ exhibiting *at least one* of the free disposal/desirability properties (10.33)–(10.35).

Both assertions (i) and (ii) above can be derived, arguing by contradiction, from the definitions of local strict Pareto and strong Pareto optimal allocations; see Exercise 10.37 to complete the proof of the corollary. △

10.3.3 Properness for Weak Pareto Optimal Allocations

Let us further establish relationships between the asymptotic closedness introduced in Definition 10.7 and some *properness* properties in economic modeling originated by Mas-Colell and then well recognized and developed in welfare economics. In this subsection, we consider the economy $\mathcal{E}$ as in (10.1) whose commodity space is a *Banach lattice* and suppose for simplicity that $C_1 = \ldots = C_n = E_+$, that $m = 1$ with S standing for the *total production set*, and that $W = \{\bar{w}\}$ (markets clear). Recall first the most advanced (to the best of our knowledge) properness properties developed in [270].

Definition 10.13 (Properness Properties). *Given the economy $\mathcal{E}$ as described above whose commodity space is a Banach lattice, we say that:*

(A1) *$\mathcal{E}$ satisfies the* PROPERNESS PROPERTY FOR PREFERENCES *if there are positive numbers δ, λ, and θ such that*

$$\Big(\big(P_i(\bar{z}) \cap (\bar{z}_i + \theta\mathbb{B})\big) + \Gamma\Big) \cap E_+ \subset P_i(\bar{z})$$
$$\textit{with } \Gamma := \bigcup_{t \in (0,\lambda]} t\Big(\frac{1}{(n+1)}\bar{w} + \delta\mathbb{B}\Big), \quad i = 1,\ldots,n. \tag{10.36}$$

(A2) *$\mathcal{E}$ satisfies the* PROPERNESS PROPERTY FOR PRODUCTION *if there are positive numbers δ, λ, and θ such that*

$$(y - \Gamma) \cap \big\{z \in E \big|\ z^+ \le y^+\big\} \subset Y \textit{ for all } y \in S \cap (\bar{y} + \theta\mathbb{B}),$$

where the cone Γ is defined as in (A1).

Note that the cone Γ in (10.36) depends on δ, λ, and θ while we do not indicate this dependence in notation for simplicity. When the preferred sets are derived from a transitive and complete preference on consumption sets, the properness condition for preferences $(A1)$ is implied by Mas-Collel's *uniform properness property*. Following this idea, we say that the preference relation $\prec$ is *proper* at $z \in E_+$ if there are positive numbers α and ε, a positive vector $\bar{w} \in E_+$, and a neighborhood of the origin $\mathcal{O}$ such that

$$u \in L \;\text{ and }\; \big[z - \alpha\bar{w} + u \prec x \Longrightarrow u \in \varepsilon\mathcal{O}\big].$$

The preference $\prec$ is *uniformly proper* if it is proper at every $z \in E_+$, while $\bar{w}$ and V can be chosen independent of z. Geometrically, the properness at z means that there is an open cone $\Gamma \subset E$ containing positive vectors such that

$$(-\Gamma) \cap \big\{u - z \in E_+ \big|\ u \prec z\big\} = \emptyset.$$

The next proposition establishes relationships between the properness for preference property (A1) and the asymptotic closedness of preference sets in the economic model under consideration.

Proposition 10.14 (Properness Implies Asymptotic Closedness). *The following assertions hold for the economy $\mathcal{E}$:*

(i) *The properness property* (A1) *of $P_i(\bar{z})$ at $\bar{z}_i$ implies the asymptotic closedness property of $P_i(\bar{z})$ at this point.*
(ii) *If property* (A1) *is satisfied, then the set*

$$\widetilde{P}_i(\bar{z}) := P_i(\bar{z}) \cap (\bar{z}_i + \theta\mathbb{B}) + \Gamma$$

is asymptotically closed at $\bar{z}_i$.

Proof. Let us first verify (i). Assume that $P_i(\bar{z})$ is proper at $\bar{z}_i$. To get the asymptotic closedness of $P_i(\bar{z})$ at $\bar{z}_i$ from (A1), we intend to show that

$$\big(\text{cl}\, P_i(\bar{z}) + c_k\big) \cap V \subset P_i(\bar{z}) \;\text{ for all }\; k \;\text{ sufficiently large,} \tag{10.37}$$

where the set $V \subset Z$ and the sequence $\{c_k\} \subset Z$ are defined by

$$V := \Big(\bar{z}_i + \frac{\theta}{2}\mathbb{B}\Big) \text{ and } c_k := \frac{\lambda}{k}\Big(\frac{1}{n+1}\bar{w} + \delta c\Big) \text{ with } c \in E_+ \cap \mathbb{B}.$$

To furnish this, fix $k \in \mathbb{N}$ so large that $c_k \in (\theta/2)\mathbb{B}$. Take further any $z \in (\text{cl}\, P_i(\bar{z}) + c_k) \cap V$ and find a sequence $\{z_m\} \subset P_i(\bar{z}) \subset E_+$ satisfying $z_m \to \widetilde{z} \in \text{cl}\, P_i(\bar{z})$, $z_m + c_k \in V$, and $z = \widetilde{z} + c_k$. Then $z_m \in -c_k + V \subset (\bar{z}_i + \theta\mathbb{B})$. Due to $c_k \in \imath\Gamma$, there is $\gamma > 0$ such that

$$z_m + c_k \in z_m + c_k + \gamma\mathbb{B} \subset \big(P_i(\bar{z}) \cap (\bar{z}_i + \theta\mathbb{B}) + \Gamma\big) \text{ for all } m \in \mathbb{N}.$$

Passing now to the limit as $m \to \infty$, we get from (10.36) that

$$z = \widetilde{z} + c_k \in \big(P_i(\bar{z}) \cap (\bar{z}_i + \theta\mathbb{B}) + \Gamma\big) \cap E_+ \subset P_i(\bar{z}).$$

Since $z \in (\text{cl}\, P_i(\bar{z}) + c_k) \cap V$ was chosen arbitrarily, the latter yields (10.37), which justifies the asymptotic closedness of $P_i(\bar{z})$ in (i) and ensures furthermore the asymptotic closedness property of $\widetilde{P}_i(\bar{z})$ asserted in (ii). $\triangle$

The following example shows that the converse in Proposition 10.14(i) fails even for economies with finite-dimensional commodity spaces.

Example 10.15 (Asymptotic Closedness Is Strictly Better Than Properness). Take $E = \mathbb{R}^3$ and define the preference set $P(0)$ at $0 \in \mathbb{R}^3$ by

$$P(0) := \big\{(a, b, 0) \in \mathbb{R}^3 \big|\, a, b \geq 0\big\} \setminus \{0\}.$$

Since $P(0) \cup \{0\}$ is a closed and convex cone, the set $P(0)$ clearly has the asymptotic closedness property at $\bar{z} = 0$. It turns out however that

$$\mathbb{R}^3_+ \cap \big(P(0) \cap \gamma\mathbb{B} + \Gamma\big) \not\subset P(0)$$

for every $\gamma > 0$ with Γ given in (10.36). To see this, take any triple $v = (v_1, v_2, v_3) \in \Gamma \cap \mathbb{R}^3_+$ with $v_3 > 0$; the existence of such v is ensured by the nonempty interior of Γ. Then for every $u = (u_1, u_2, 0) \in P(0) \cap \gamma\mathbb{B}$ we have $u_3 + v_3 = v_3 > 0$ and hence $u + v \notin P(0)$. This shows that $P(0)$ doesn't have the properness property at the origin. Observe in more generality that it happens when $\text{span}\, P_i(\bar{z}) \subset L$, where $L \neq E$ is a subspace of E.

The following proposition shows that the properness assumptions imposed above allow us to reduce *local weak Pareto* allocations of the economy $\mathcal{E}$ under consideration to those of a modified economy with *nonempty interior* properties of preference and production sets.

Proposition 10.16 (Weak Pareto Optimal Allocations Under Properness Assumptions). *Let* $(\bar{y}, \bar{z})$ *be a local weak Pareto optimal allocation of the economy*

$\mathcal{E} = (P_1, \ldots, P_n, S, \bar{w})$ *under the properness assumptions* (A1) *and* (A2) *from Definition* 10.13. *Then* $(\bar{y}, \bar{z})$ *is a local weak Pareto optimal allocation of the modified economy* $\widetilde{\mathcal{E}} = (\widetilde{P}_1, \ldots, \widetilde{P}_n, \widetilde{S}, \bar{w})$ *with*

$$\begin{aligned} &\widetilde{P}_i(\bar{z}) := P_i(\bar{z}) \cap (\bar{z}_i + \theta\mathbb{B}) + \Gamma,\ i = 1, \ldots, n, \\ &\widetilde{S} := S \cap (\bar{y} + \theta\mathbb{B}) - \Gamma. \end{aligned} \tag{10.38}$$

Proof. Arguing by contradiction, suppose that $(\bar{y}, \bar{z})$ is *not* a local weak Pareto optimal allocation of $\widetilde{\mathcal{E}}$ and thus find (y, z) satisfying

$$y \in \widetilde{S},\ \ \bar{w} = \sum_{i=1}^{n} z_i - y, \ \ \text{and}\ \ z_i \in \widetilde{P}_i(\bar{z})\ \ \text{for all}\ \ i = 1, \ldots, n.$$

The latter relationships clearly ensure the validity of the inclusion

$$\bar{w} \in \sum_{i=1}^{n} \Big(\big(P_i(\bar{z}) \cap (\bar{z}_i + \theta\mathbb{B})\big) + \Gamma\Big) - \Big(\big(S \cap (\bar{y} + \theta\mathbb{B})\big) - \Gamma\Big),$$

which contradicts the imposed properness assumptions (see Exercise 10.39) and thus verifies the claimed local weak Pareto optimality of $(\bar{y}, \bar{z})$. △

Based on the preceding results of this subsection, we deduce from Theorem 10.29 the following two versions of the extended second welfare theorem for the economy $\mathcal{E}$ for *local weak Pareto* optimal allocations. The first version, involving the properness property (A1) and the SNC assumptions as above, is a direct consequence of Theorem 10.29 and Proposition 10.14(i) in terms of the initial data (P_i, S) of the economy under consideration. On the other hand, the second version, which employs the nonempty interior of the convex cone Γ in (10.36) and the fact that $P_i(\bar{z}) \subset \widetilde{P}_i(\bar{z})$, allows us to derive from Theorem 10.29 and Propositions 10.14(ii), 10.16 a new version of the second welfare theorem for local weak Pareto optimal allocations *without SNC* requirements while for the modified economy $\widetilde{\mathcal{E}}$ with data (10.38).

Theorem 10.17 (Extended Second Welfare Theorem for Local Weak Pareto Optimal Allocations Under Properness Assumptions). *Let* $(\bar{y}, \bar{z})$ *be a local weak Pareto optimal allocation of the economy* $\mathcal{E} = (P_i, S, \bar{w})$ *with an ordered Asplund commodity space* E. *Then the following assertions hold:*

(i) *Assume that* $\mathcal{E}$ *satisfies the properness for preferences assumption* (A1) *and that one of the sets* $\text{cl}\, P_i(\bar{z})$*,* $i = 1, \ldots, n$*, and* S *is SNC at* $\bar{z}_i$ *and* $\bar{y}$*, respectively. Then there is a marginal price* $p^* \in E^* \setminus \{0\}$ *such that*

$$-\,p^* \in N\big(\bar{z}_i; \text{cl}\, P_i(\bar{z})\big),\ \ i = 1, \ldots, n, \ \ \textit{and}\ \ p^* \in N(\bar{y}; S). \tag{10.39}$$

(ii) *Assume that both properness properties* (A1) *and* (A2) *are satisfied. Then there is a marginal price* $p^* \in E^* \setminus \{0\}$ *such that*

$$-p^* \in N\big(\bar{z}_i; \operatorname{cl} \widetilde{P}_i(\bar{z})\big),\ i = 1, \ldots, n,\ \ \textit{and}\ \ p^* \in N(\bar{y}; \widetilde{S}) \tag{10.40}$$

via the modified preference and total production sets taken from (10.38).

Proof. To justify (i), employ Proposition 10.14(i) ensuring under (A1) the asymptotic closedness of each preference set $P_i(\bar{z})$ at $\bar{z}_i$ as $i = 1, \ldots, n$ imposed in Theorem 10.29 for weak Pareto optimal allocations. This allows us to get (10.39) with $p^* \neq 0$ under the SNC assumptions made.

To verify (ii), we use Proposition 10.16 telling us that $(\bar{y}, \bar{z})$ is a local weak Pareto optimal allocation for the modified economy $\widetilde{\mathcal{E}}$ with data (10.38). It follows from Proposition 10.14(ii) that each modified preference sets $\widetilde{P}_i(\bar{z})$ is asymptotically closed at $\bar{z}_i$, $i = 1, \ldots, n$. For applying now Theorem 10.29 to the weak Pareto optimal allocation $(\bar{y}, \bar{z})$ of the economy $\widetilde{\mathcal{E}}$, it is sufficient to demonstrate that at least *one* of the sets $\operatorname{cl} \widetilde{P}_i(\bar{z}_i)$, $i = 1, \ldots, n$, and $\widetilde{S}$ is SNC at the corresponding point. Let us show that in fact *all* of them are SNC at the references points. Taking into account the structures of the sets (10.38) and the construction of the regular normal cone in the definition of the SNC property, we consider without loss of generality the set

$$\widetilde{P}(\bar{z}) := P(\bar{z}) \cap (\bar{z} + \theta \mathbb{B}) + \Gamma \tag{10.41}$$

with the cone Γ defined in (10.36), and show that this set is SNC at $\bar{z} \in \operatorname{cl} \widetilde{P}(\bar{z})$. Pick a sequence $\{z_k, z_k^*\} \subset Z \times Z^*$ satisfying

$$z_k \to \bar{z},\ z_k^* \overset{w^*}{\to} 0\ \text{ as }\ k \to \infty\ \text{ with }\ z_k^* \in \widehat{N}\big(z_k; \widetilde{P}(\bar{z})\big),\ k \in \mathbb{N}.$$

Employing now the decreasing property of the regular normal cone with respect to set inclusions from Exercise 1.39(iii) and the convexity of the cone Γ, deduce from the above that

$$z_k^* \in N(\widetilde{z}_k; \Gamma) \subset N(0; \Gamma), \quad k \in \mathbb{N},$$

with some sequence $\{\widetilde{z}_k\} \subset Z$, which implies by the nonempty interior of Γ that $\|z_k^*\| \to 0$ as $k \to \infty$ due to Exercise 2.29. This justifies the SNC property of set $\widetilde{P}(\bar{z})$ in (10.41) at $\bar{z}$ (and of all the sets defined in (10.38) at the corresponding points) and thus completes the proof of the theorem. △

We refer the reader to Exercise 10.41 for more discussions on relationships between the two versions of the extended second welfare theorem under the properness assumptions given in Theorem 10.17.

10.4 Global Extended Second Welfare Theorems

In this section, we focus on deriving extended versions of the second welfare theorem for *global* Pareto-type optimal allocations of all the *four kinds* described in Definition 10.2. Since any global optimal allocation is a local one of the same type, the results of Section 10.3 obtained for local weak Pareto, strict Pareto, and strong Pareto optimal allocations surely hold for their global counterparts. Now we show, by using somewhat different arguments, that the global nature of all the optimal allocations under consideration in Definition 10.2 (including Pareto ones defined in (ii,v) therein, which are not studied in the local framework of Section 10.3), allows us to derive versions of the second welfare theorem held under *less restrictive qualification* conditions in comparison with those employed above.

10.4.1 Net Demand Qualification Conditions

Here are these *net demand qualification conditions* in the general welfare economic setting $\mathcal{E}$ of (10.1) defined in a similar while somewhat different way to serve each type of global optimal allocations from Definition 10.2.

Definition 10.18 (Net Demand Qualification Conditions). *Let $(\bar{y}, \bar{z})$ be a feasible allocation of the economy $\mathcal{E}$ in (10.1) with the preference sets $P_i(z)$, and let $\bar{w} := \sum_{i=1}^{n} \bar{z}_i - \sum_{j=1}^{m} \bar{y}_j \in W$. Rename the sets $P_i(\bar{z})$, S_j, and W as:*

$$\Xi_i := P_i(\bar{z}), i = 1, \ldots, n, \;\; \Xi_{n+j} := -S_j, j = 1, \ldots, m, \;\; \Xi_{m+n+1} := -W$$

with $\bar{x}_i := \bar{z}_i$, $\bar{x}_{n+j} := -\bar{y}_{n+j}$, $\bar{x}_{m+n+1} := -\bar{w}$. Given $\varepsilon > 0$, consider the set

$$\Delta_\varepsilon := \sum_{i=1}^{n+m+1} \mathrm{cl}\, \Xi_i \cap (\bar{z}_i + \varepsilon \mathbb{B}) \tag{10.42}$$

and define the following qualification conditions at the allocation $(\bar{y}, \bar{z})$, where the closure operation is redundant for all the sets Θ_i with $i > n$ due to the standing assumptions made for the economy $\mathcal{E}$:

(i) *The* NET DEMAND WEAK QUALIFICATION (NDWQ) CONDITION *holds at $(\bar{y}, \bar{z})$ if there are $\varepsilon > 0$ and $\{e_k\} \subset E$ with $e_k \to 0$ such that*

$$\Delta_\varepsilon + e_k \subset \sum_{i=1}^{n} \Xi_i + \sum_{i=n+1}^{n+m+1} \mathrm{cl}\, \Xi_i \;\; \textit{for large} \;\; k \in \mathbb{N}. \tag{10.43}$$

(ii) *The* NET DEMAND QUALIFICATION (NDQ) CONDITION *holds at $(\bar{y}, \bar{z})$ if there are $\varepsilon > 0$, $\{e_k\} \subset E$ with $e_k \to 0$, and $i_0 \in \{1, \ldots, n\}$ such that*

$$\Delta_\varepsilon + e_k \subset \Xi_{i_0} + \sum_{i=1,\, i\neq i_0}^{n+m+1} \mathrm{cl}\, \Xi_i \;\; \textit{for large } k \in \mathbb{N}. \tag{10.44}$$

(iii) *The* NET DEMAND STRICT QUALIFICATION (NDSQ) CONDITION *holds at* $(\bar{y}, \bar{z})$ *if there are* $\varepsilon > 0$, $e_k \to 0$, *and* $i_0 \in \{1, \ldots, n\}$ *such that*

$$\Delta_\varepsilon + e_k \subset \mathrm{cl}\, \Xi_{i_0} \setminus \{\bar{x}_{i_0}\} + \sum_{i=1,\, i\neq i_0}^{n+m+1} \mathrm{cl}\, \Xi_i \;\; \textit{for large } \; k \in \mathbb{N}. \tag{10.45}$$

(iv) *The* NET DEMAND STRONG QUALIFICATION (NDSNQ) CONDITION *holds at* $(\bar{y}, \bar{z})$ *if there are* $\varepsilon > 0$, $e_k \to 0$, *and* $i_0 \in \{1, \ldots, m+n+1\}$ *with*

$$\Delta_\varepsilon + e_k \subset \mathrm{cl}\, \Xi_{i_0} \setminus \{\bar{x}_{i_0}\} + \sum_{i=1,\, i\neq i_0}^{m+n+1} \mathrm{cl}\, \Xi_i \;\; \textit{for large } \; k. \tag{10.46}$$

We have the following relationships between net demand qualification conditions and the asymptotic closedness requirements on the corresponding sets; *three* of these requirements were employed in Section 10.3 to derive local versions of the second welfare theorem for Pareto-type optimal allocations.

Proposition 10.19 (Relationships Between Net Demand Qualification and Asymptotic Closedness Conditions). *Consider the economy* $\mathcal{E}$ *from* (10.1) *with the preference sets* $P_i(z)$. *Given a feasible allocation* $(\bar{y}, \bar{z})$ *of* $\mathcal{E}$, *assume that the production sets* S_j, $j = 1, \ldots, m$, *and the net demand set* W *are locally closed around the points in question. Then:*

(i) *The NDWQ condition* (10.43) *holds at* $(\bar{y}, \bar{z})$ *if all the preference sets* $P_i(\bar{z})$ *as* $i = 1, \ldots, n$ *are asymptotically closed at* $\bar{z}_i$, *respectively.*

(ii) *The NDQ condition* (10.44) *holds at* $(\bar{y}, \bar{z})$ *if there is* $i_0 \in \{1, \ldots, n\}$ *so that the preference set* $P_{i_0}(\bar{z})$ *is asymptotically closed at* $\bar{z}_{i_0}$.

(iii) *The NDSQ condition* (10.45) *holds at* $(\bar{y}, \bar{z})$ *if there is* $i_0 \in \{1, \ldots, n\}$ *so that the set* cl $P_{i_0}(\bar{z})$ *is asymptotically closed at* $\bar{z}_{i_0}$.

(iv) *The NDSNQ condition* (10.46) *holds at* $(\bar{y}, \bar{z})$ *if one of the sets*

$$\mathrm{cl}\, P_i(\bar{z}), \ldots, \mathrm{cl}\, P_n(\bar{z}),\;\; S_1, \ldots, S_j,\;\; W$$

is asymptotically closed at the corresponding point.

Proof. Assertions (i) and (ii) follow from the proof of [523, Proposition 8.4], while the other two assertions (iii) and (iv) can be verified similarly. △

It occurs that the converse implications in Proposition 10.19 fail in quite common finite-dimensional situations.

Example 10.20 (Net Demand Qualification Is Strictly Weaker Than Asymptotic Closedness). Consider the markets clear economy $\mathcal{E}$ with $E = \mathbb{R}^2$, $n = m = 1$, $W = \{0\}$, and

$$C = \mathbb{R}^2_+,\ S = \Xi := \left\{(a,b) \in \mathbb{R}^2_+ \,\middle|\, ab = 0\right\},\ P(z) \equiv \Xi \setminus \{(0,0)\},\ \bar{z} = (0,0).$$

It is obvious that the set $P(0)$ is not asymptotically closed at $0 \in \mathbb{R}^2$, while the NDQ condition (10.44) is satisfied at $(\bar{y}, \bar{z}) = (0,0) \in \mathbb{R}^4$ by

$$\left(k^{-1}, k^{-1}\right) + \Xi - \Xi \subset \Xi \setminus \{(0,0)\} - \Xi.$$

10.4.2 Global Optimality in Welfare Economics

The next major theorem establishes extended versions of the second welfare theorem for all the four types of *global* Pareto optimal allocations from Definition 10.2. These extended versions are valid under the corresponding net demand qualification conditions from Definition 10.18.

Theorem 10.21 (Extended Second Welfare Theorems for Global Pareto-Type Optimal Allocations). *Let $(\bar{y}, \bar{z})$ be a global optimal allocation of economy (10.1) in the senses listed below with respect to the preference sets $P_i(z)$ under the local satiation requirement (10.27). Assume that the commodity space E is Asplund and that one of the sets (10.28) is SNC at $\bar{z}_i$, $\bar{y}_j$, and $\bar{w} = \sum_{i=1}^n \bar{z}_i - \sum_{j=1}^m \bar{y}_j$, respectively. Then there exists a nonzero price $p^* \in E^*$ satisfying all the relationships (10.29) of the extended second welfare theorem in each of the following cases of global optimal allocations of $\mathcal{E}$:*

- *$(\bar{y}, \bar{z})$ is a* GLOBAL WEAK PARETO OPTIMAL ALLOCATION *provided that the net demand weak qualification condition (10.43) is satisfied.*
- *$(\bar{y}, \bar{z})$ is a* GLOBAL PARETO OPTIMAL/EFFICIENT ALLOCATION *provided that the net demand qualification condition (10.44) is satisfied.*
- *$(\bar{y}, \bar{z})$ is a* GLOBAL STRICT PARETO OPTIMAL ALLOCATION *provided that the net demand strict qualification condition (10.45) is satisfied.*
- *$(\bar{y}, \bar{z})$ is a* GLOBAL STRONG PARETO OPTIMAL ALLOCATION *provided that the net demand strong qualification condition (10.46) is satisfied.*

Proof. Consider two subsets of the Asplund spaces E^{m+n+1} defined by

$$\begin{cases} \Omega_1 := \prod_{j=1}^m S_j \times \prod_{i=1}^n \mathrm{cl}\, P_i(\bar{z}) \times W, \\ \Omega_2 := \left\{(y,z,w) \in E^{m+n+1} \,\middle|\, \sum_{i=1}^n z_i - \sum_{j=1}^m y_j - w = 0\right\}, \end{cases} \tag{10.47}$$

which are locally closed around the point $(\bar{y}, \bar{z}, \bar{w})$, and show that this point is *locally extremal* for the system $\{\Omega_1, \Omega_2\}$ under the fulfillment of the NDWQ/NDQ/NDSQ/NDSGQ conditions held for the corresponding weak Pareto/Pareto/strict Pareto/strong Pareto optimal allocation of the economy $\mathcal{E}$. Indeed, we always have $(\bar{y}, \bar{z}, \bar{w}) \in \Omega_1 \cap \Omega_2$, and it remains to check that there exist a sequence $\{a_k\} \subset E^{m+n+1}$ with $a_k \to 0$ as $k \to \infty$ and a neighborhood U of $(\bar{y}, \bar{z}, \bar{w})$ such that

$$\Omega_1 \cap (\Omega_2 - a_k) \cap U = \emptyset \;\text{ for all large }\; k \in \mathbb{N} \tag{10.48}$$

when the corresponding qualification condition (10.43)–(10.46) is satisfied. Arguing in a unified way, take $\{e_k\} \subset E$ from the selected qualification condition and form the sequence $a_k := (0, \ldots, 0, e_k) \in E^{m+n+1}$. Denote further

$$U := \prod_{j=1}^{m} (\bar{y}_j + \varepsilon \mathbb{B}) \times \prod_{i=1}^{n} (\bar{z}_i + \varepsilon \mathbb{B}) \times (\bar{w} + \varepsilon \mathbb{B})$$

and show that the extremality relationship (10.48) holds under this choice. Supposing the contrary, find a sequence of triples $(y_k, z_k, w_k) \in \Omega_1$ with $(y_k, z_k, w_k) + a_k \in \Omega_2$ and get by the structure of $\{\Omega_1, \Omega_2\}$ in (10.47) and the above choice of $\{a_k\}$ and U that

$$\begin{cases} y_{jk} \in S_j \cap (\bar{y}_j + \varepsilon \mathbb{B}), \;\; j = 1, \ldots, m, \\ z_{ik} \in \text{cl}\, P_i(\bar{z}) \cap (\bar{z}_i + \varepsilon \mathbb{B}), \;\; i = 1, \ldots, n, \\ w_k \in W \cap (\bar{w} + \varepsilon \mathbb{B}), \quad \text{and} \\ 0 = \displaystyle\sum_{i=1}^{n} z_i - \sum_{j=1}^{m} y_j - w_k + e_k \in \Delta_\varepsilon + e_k \end{cases} \tag{10.49}$$

for the set Δ_ε defined in (10.42). We check now that the relationships in (10.49) lead to a contradiction with the global Pareto-type optimality provided the fulfillment of the corresponding net demand qualification conditions.

• Assuming that $(\bar{y}, \bar{z})$ is a *global weak Pareto* optimal allocation and that the net demand weak qualification condition (10.43) is satisfied, we get

$$0 \in \sum_{i=1}^{n} P_i(\bar{z}) - \sum_{j=1}^{m} S_j - W,$$

i.e., there are $z_i \in P_i(\bar{z})$, $y_j \in S_j$, and $w \in W$ such that

$$w = \sum_{i=1}^{n} z_i - \sum_{j=1}^{m} y_j. \tag{10.50}$$

This tells us that the allocation (y, z) is feasible for $\mathcal{E}$, while the inclusions $z_i \in P_i(\bar{z})$ as $i = 1, \ldots, n$ imply that $(\bar{y}, \bar{z})$ is not a global weak Pareto optimal allocation of $\mathcal{E}$ by Definition 10.2, a contradiction.

• Assuming that $(\bar{y}, \bar{z})$ is a *global Pareto* optimal allocation and that the net demand qualification condition (10.44) holds, we get

$$0 \in P_{i_0}(\bar{z}) + \sum_{i=1\, i\neq i_0}^{n} \operatorname{cl} P_i(\bar{z}) - \sum_{j=1}^{m} S_j - W,$$

i.e., there are $z_{i_0} \in P_{i_0}(\bar{z})$, $z_i \in \operatorname{cl} P_i(\bar{z})$ for $i = 1, \ldots, n$, $y_j \in S_j$ for $j = 1, \ldots, m$, and $w \in W$ satisfying (10.50). These relationships clearly contradict the definition of the global Pareto optimality of the allocation $(\bar{y}, \bar{z})$.

• Assuming that $(\bar{y}, \bar{z})$ is a *global strict Pareto* optimal allocation and that the net demand strict qualification condition (10.45) holds, we get

$$0 \in \operatorname{cl} P_{i_0}(\bar{z}) \setminus \{\bar{z}_{i_0}\} + \sum_{i=1,\, \neq i_0}^{n} \operatorname{cl} P_i(\bar{z}) - \sum_{j=1}^{m} S_j - W,$$

i.e., there are $z_i \in \operatorname{cl} P_i(\bar{z})$ as $i = 1, \ldots, n$ with $z \neq \bar{z}$, $y_j \in S_j$ as $j = 1, \ldots, m$, and $w \in W$ satisfying (10.50). Thus the allocation (y, z) is feasible for the economy $\mathcal{E}$ with $z_i \in \operatorname{cl} P_i(\bar{z})$ for all $i = 1, \ldots, n$. This implies that $(\bar{y}, \bar{z})$ is not a global strict Pareto optimal allocation of $\mathcal{E}$ by its definition.

• Assuming finally that $(\bar{y}, \bar{z})$ is a *global strong Pareto* optimal allocation and that the net demand strong qualification condition (10.46) holds, we get:

$$\begin{cases} \text{either } \ 0 \in \operatorname{cl} P_{i_0}(\bar{z}) \setminus \{\bar{z}_{i_0}\} + \displaystyle\sum_{\substack{i=1 \\ i\neq i_0}}^{n} \operatorname{cl} P_i(\bar{z}) - \sum_{j=1}^{m} S_j - W, \\ \text{or } \ 0 \in \displaystyle\sum_{i=0}^{n} \operatorname{cl} P_i(\bar{z}) - S_{j_0} \setminus \{\bar{y}_{j_0}\} - \sum_{\substack{j=1 \\ j\neq j_0}}^{m} S_j - W, \\ \text{or } \ 0 \in \displaystyle\sum_{i=0}^{n} \operatorname{cl} P_i(\bar{z}) - \sum_{j=1}^{m} S_j - W \setminus \{\bar{w}\}. \end{cases}$$

Each of the latter conditions allows us to find $z_i \in \operatorname{cl} P_i(\bar{z})$ as $i = 1, \ldots, n$, $y_j \in S_j$ as $j = 1, \ldots, m$, and $w \in W$ satisfying (10.50) such that $(y, z) \neq (\bar{y}, \bar{z})$. Thus we clearly arrive at a contradiction with the global strong Pareto optimality of the feasible allocation $(\bar{y}, \bar{z})$ of the economy $\mathcal{E}$.

The rest of the proof for all the four kinds of global Pareto-type optimal allocations are similar to the proofs of Theorem 10.10 with the usage of Theorem 10.9 based on the application of the product extremal principle and SNC calculus to the set system (10.47). △

Remark 10.22 (Global vs. Local Pareto-Type Optimal Allocations). Let us demonstrate that the net demand qualification conditions from Definition 10.18 are appropriate to deal with *global*, not *local* Pareto-type optimal allocations. Consider the economy $\mathcal{E}$ with the initial data given by $E = \mathbb{R}^2$, $n = m = 1$, $C = \mathbb{R}^2_+$, $W = \{0\}$, and

$$S := \{(a, -a+2) \in \mathbb{R}^2 \mid a \le 0\}$$
$$\cup\{(a,b) \in \mathbb{R}^2_+ \mid a^2 + (b-1)^2 = 1\} \cup \{(a,-a) \mid a \ge 0\}.$$

The customer uses the preference generated by a nonconvex cone Θ as follows:

$$P(z) := z + \Theta \setminus \{0\}, \quad \text{where} \quad \Theta := \{z = (a,b) \in \mathbb{R}^2 \mid ab = 0\}.$$

Since $P(z) \cup \{z\}$ is a closed set, there is no difference between weak Pareto, Pareto, and strict Pareto optimal allocations as well as between NDWQ, NDQ, and NDSQ conditions from Definition 10.18. It is easy to check that $(\bar{y}, \bar{z}) = (0,0) \in \mathbb{R}^2 \times \mathbb{R}^2$ is a *local* (not global) Pareto optimal allocation of the economy $\mathcal{E}$ with the ball neighborhood $\mathcal{O} := \operatorname{int}\mathbb{B} \times \operatorname{int}\mathbb{B}$. Furthermore, the net demand qualification condition is satisfied since the set $P(0) - S$ contains the unit ball of $\mathbb{R}^2$, and thus the underlying inclusion

$$\Delta_1 + e_k \in P(0) - S$$

holds with $\varepsilon > 0$ sufficiently small (in fact $\varepsilon \le 1$) and $k \in \mathbb{N}$ sufficiently large. For this example, the inclusion $0 \in \Delta_1 + e_k$ implies that $0 \in P(0) - S$. The latter gives us $z = (0,2) \in P(0)$ and $y = (0,2) \in S$. Observe that there is *no contradiction* with *local optimality* of the reference *Pareto* optimal allocation $(\bar{y}, \bar{z}) = (0,0)$ since the feasible allocation (y,z) found above doesn't belong to the aforementioned neighborhood $\mathcal{O}$ of $(\bar{y}, \bar{z})$.

10.5 Exercises for Chapter 10

Exercise 10.23 (Relationships Between Various Pareto-Type Optimal Allocations). Consider the four types of local Pareto-type optimal allocations and their global versions listed in Definition 10.2.

(i) Show that all the implications (iv)⇒(iii)⇒(ii)⇒(i) are strict.

(ii) Assuming that the extended preference sets $P_i(\bar{z}) \cup \{\bar{z}\}$ are locally closed around $\bar{z}$ for all $i = 1, \ldots, n$, show that the notions of Pareto, weak Pareto, and strict Pareto optimal allocations are identical.

(iii) Clarify relationships between the aforementioned Pareto optimal allocations and the strong Pareto one in the setting of (ii).

Exercise 10.24 (Relationships Between Fully Localized Solutions in Set-Valued Optimization). Show that both implications (iii)⇒(ii)⇒(i) in Definition 10.3 are strict even for unconstrained problems in finite-dimensional spaces.

Exercise 10.25 (Fully Localized Multiobjective Optimization Description of Local Pareto Optimal Allocations in Welfare Economics). Find an appropriate notion of fully localized optimal solutions to the set-valued optimization problem (10.9) that is equivalent to local Pareto optimal allocations of the economy $\mathcal{E}$. Compare it with a fully localized counterpart of super minimizers from Definition 9.46.

Exercise 10.26 (Existence of Pareto-Type Optimal Solutions in Set-Valued Optimization and Welfare Economics).

(i) Investigate the possibility of establishing the existence of (global) optimal solutions of the set-valued optimization problem (10.3) with respect to the preference relation (10.4) by modifying the constructions of Section 9.3.

(ii) Considering the set-valued optimization framework of the welfare economics in (10.9) and (10.10), find in this way efficient conditions ensuring the existence of global Pareto-type optimal allocations of the economy $\mathcal{E}$ given in Definition 10.2.

Exercise 10.27 (Exact Extremal Principle in Product Spaces). Give a detailed proof of Lemma 10.6. *Hint:* Follow the lines in the proof of [523, Lemma 5.58] by passing to the limit in the relationships of the approximate extremal principle with the usage of the PSNC and strong PSNC conditions from Definition 10.5.

Exercise 10.28 (Sufficient Conditions for Asymptotic Closedness of Sets). Justify the sufficiency for the asymptotic closedness property of each item (i)–(iii) listed right after Definition 10.7; compare this with [59, Section 13.3].

Exercise 10.29 (Necessary Conditions for Fully Localized Minimizers in Product Spaces). Establish a version of Theorem 10.9 in the case where both spaces X and Z are represented in the product forms: $X = \prod_{i=1}^{n} X_i$ and $Z = \prod_{j=1}^{m} Z_j$. *Hint*: Proceed as in the proof of Theorem 10.9 with the usage of Proposition 10.8 and the observation that the PSNC assumptions on the sets Ω_1 and Ω_2 can be replaced by the following: Ω_1 is PSNC at $(\bar{x}, \bar{z})$ with respect to some $I \subset \{1, \ldots, n; 1, \ldots, m\}$ and Ω_2 is strongly PSNC at $(\bar{x}, \bar{z})$ with respect to some $J \subset \{1, \ldots, n; 1, \ldots, m\}$, where $I \cup J = \{1, \ldots, n; 1 \ldots, m\}$.

Exercise 10.30 (Comparison Between Necessary Conditions for Pareto-Type Local Minimizers in Multiobjective Optimization).

(i) Check that the necessary optimality conditions obtained in Theorem 10.9 for fully localized strong minimizers imply those for fully localized minimizers, while the latter conditions yield those for fully localized weak minimizers.

(ii) Give examples showing that the converse implications in (i) fail in both finite and infinite dimensions.

(iii) In the case of local Pareto minimizers for problem with single-valued objectives, compare the optimality conditions of Theorem 10.9 with the corresponding results from [523, Theorem 5.73].

Exercise 10.31 (Relationships Between Local Versions of the Second Welfare Theorem for Pareto-Type Optimal Allocations in General Commodity Spaces). Consider the setting of Theorem 10.10.

(i) Show that the imposed asymptotic closedness properties are essential for the validity of the obtained versions of the second welfare theorem.

(ii) Check that the marginal price conditions established for local strong Pareto optimal allocations imply those for local strict ones, which yield in turn the results for local weak Pareto optimal allocations of $\mathcal{E}$ while the converse implications fail.

(iii) Give an example showing that the obtained conditions for local strict Pareto optimal allocations are not necessary for local Pareto ones.

Exercise 10.32 (Excess Demand Condition). Consider the economy $\mathcal{E}$ with the net demand constraint set W given by

$$W = \omega + \Gamma \quad \text{for some} \quad \omega \in W. \tag{10.51}$$

In particular, for the case of commodity spaces partially ordered by the closed positive cone E_+ representation (10.51) with $\Gamma := -E_+$ corresponds to the *implicit free disposal of commodities*.

Show that the last condition on the marginal price p^* in (10.29) yields in case (10.51) the *zero value of excess demand* condition

$$\Big\langle p^*, \sum_{i=1}^{n} \bar{x}_i - \sum_{j=1}^{m} \bar{y}_j - \omega \Big\rangle = 0.$$

Exercise 10.33 (Second Welfare Theorem for Economies with Convex Preference and Production Sets). In the setting of Theorem 10.10 consider the economy $\mathcal{E}$, where all the preference $P_i(\bar{x})$ and production S_j sets are convex.

(i) Show that, under the validity of the corresponding assumptions of the theorem concerning local weak, strict, and strong Pareto optimal allocations, there exists a nonzero price $p^* \in N(\bar{w}; W)$ satisfying the conditions:

$$\begin{aligned} &\bar{x}_i \ \text{ minimize } \ \langle p^*, x_i \rangle \ \text{ over } \ x_i \in \operatorname{cl} P_i(\bar{x}), \quad i = 1, \ldots, n, \\ &\bar{y}_j \ \text{ maximize } \ \langle p^*, y_j \rangle \ \text{ over } \ y_j \in S_j, \quad j = 1, \ldots, m. \end{aligned}$$

(ii) Compare the result in (i) for weak Pareto optimal allocations with the classical version of the second welfare theorem in convex economies under the *nonempty interiority* assumptions; see, e.g., [490] and the references therein.

Exercise 10.34 (Approximate Versions of the Second Welfare Theorem). Consider the welfare economy (10.1) in the Asplund space setting.

(i) Derive approximate versions of assertions (i)–(iii) of Theorem 10.29 via the regular normal cone (1.5) without any SNC/PCNC assumptions. *Hint:* Use the approximate extremal principle similarly to the proof of [523, Theorem 8.5].

(ii) Are the results from (i) equivalent to those in Theorem 10.29 in finite-dimensional commodity spaces?

Exercise 10.35 (Decentralized Equilibria in Nonconvex Economies via Nonlinear Prices). Using the smooth variational descriptions of regular normals to nonconvex sets from Theorem 1.10(ii) and its infinite-dimensional extensions from Exercise 1.51, give interpretations of the appropriate second welfare theorems from Exercise 10.34 via approximate decentralized (convex-type) equilibria by using nonlinear (convex-concave) marginal prices. *Hint:* Compare with [523, Theorem 8.7] obtained for the cases of global Pareto and weak Pareto optimal allocations.

Exercise 10.36 (Price Positivity). Prove that condition (10.32) in a partially ordered Banach space E implies that $N(\bar{z}; \Xi) \subset E^*_+$ for any $\bar{z} \in \Xi$. *Hint:* Employ the Banach space definition of the basic normal cone (1.58) and the decreasing property of regular normals from Exercise 1.39(iii).

Exercise 10.37 (Boundary Point Descriptions of Local Pareto-Type Optimal Allocations). Consider the setting of Corollary 10.12.

(i) Give detailed verifications of both items (i) and (ii) in the proof of Corollary 10.12 concerning local strict and strong Pareto optimal allocations.

(ii) Does such a boundary point description (and hence the second welfare result of Corollary 10.12) hold in the local Pareto optimal allocations?

Exercise 10.38 (Asymptotic Closedness of Modified Preferences Sets). Give a detailed proof of assertion (ii) in Proposition 10.14.

Exercise 10.39 (Admissible Commodities Under Properness). Show that under assumptions (A1) and (A2) in the setting of Proposition 10.16, we have

$$\bar{w} \notin \sum_{i=1}^{n} \Big(\big(P_i(\bar{z}) \cap (\bar{z}_i + \theta \mathbb{B})\big) + \Gamma \Big) - \Big(\big(S \cap (\bar{y} + \theta \mathbb{B})\big) - \Gamma \Big),$$

Hint: Follow the proof of [270, Claim 4.1] based on the decomposition property of vector lattices that is standard in the literature on properness properties.

Exercise 10.40 (Order Stability of Sets with Respect to Cones). We say that a subset $\Xi \subset Z$ of a Banach space Z partially ordered by a convex cone $\Theta \subset Z$ is *order stable* at $\bar{z} \in \Xi$ if the following holds:

$$\text{for all } \{z_k\} \subset \Xi + \Theta \text{ with } z_k \to \bar{z} \text{ there is } \{\widetilde{z}_k\} \subset \Xi \text{ such that } \widetilde{z}_k \to \bar{z}.$$

Assuming that Ξ is order stable at $\bar{z} \in \Xi$, show that

$$N(\bar{z}; \Xi + \Theta) \subset N(\bar{z}; \Omega). \tag{10.52}$$

Exercise 10.41 (Relationships Between Versions of the Second Welfare Theorem Under Properness Properties).

(i) Construct examples showing that assertions (i) and (ii) of Theorem 10.17 are generally independent in both finite and infinite dimensions.

(ii) Using (10.52), show that (ii)$\Rightarrow$(i) in Theorem 10.17 provided that all the sets $\mathrm{cl}\,P_i(\bar{z})$, $i = 1, \ldots, n$, and S are order stable at $\bar{z}_i$ and $\bar{y}$, respectively.

(iii) Show that the implication in (ii) fails in the commodity space $E = \mathbb{R}^2$ with the ordering cone $\Theta = \mathbb{R}^2_+$ without the imposed order stability assumption.

Exercise 10.42 (Net Demand and Asymptotical Closedness). Consider nonconvex economies in the setting of Proposition 10.19.

(i) Verify assertions (iii) and (iv) of this proposition.

(ii) Show that the converse implications fail for economies with $E = \mathbb{R}^2$.

Exercise 10.43 (Global Versions of the Second Welfare Theorem from the Extremal Principle). Give a detailed proof of Theorem 10.21 for all the four kinds of Pareto-type global optimal allocations by applying the product extremal principle from Lemma 10.6 to the set system (10.47). *Hint:* Proceed similarly to the proof of Theorem 10.10 by taking into account the established extremality of the triple $(\bar{y}, \bar{z}, \bar{w})$ for (10.47) in all the cases under consideration.

Exercise 10.44 (Extended Second Welfare Theorem for Local Pareto Optimal Allocations). Let $(\bar{y}, \bar{z})$ be a local Pareto optimal allocation of the economy (10.1) in either finite-dimensional or Asplund space setting.

(i) Clarify whether the NDQ condition from Definition 10.18(ii) ensures the validity of the second welfare theorem (10.29) for $(\bar{y}, \bar{z})$.

(ii) Clarify whether the following condition ensures the validity of the second welfare theorem for local Pareto optimal allocation of $\mathcal{E}$: There is a consumer index $i_0 \in \{1, \ldots, n\}$ such that the preference set $P_{i_0}(\bar{z})$ is asymptotically closed at $\bar{z}_{i_0}$.

(iii) Find sufficient conditions under which the relationships in (10.29) hold for the local Pareto optimal allocation $(\bar{y}, \bar{z})$ of $\mathcal{E}$, which is not a global one.

10.6 Commentaries to Chapter 10

Competitive models of *welfare economics*, starting with the classical Walrasian equilibrium model and the subsequent fundamental developments by Pareto, Lange, Hicks, Samuelson, Arrow and Debreu, have also been among the strongest motivations for developments of new mathematical techniques and forms of analysis. We refer the reader to the surveys in Khan [412] and in Chapter 8 of the author's book [523] with the extended bibliographies therein for detailed discussions of different approaches, genesis of ideas, results, and applications known at that time.

It has been well realized that methods of modern variational analysis and generalized differentiation provide useful tools for better understanding of such and related microeconomic models with discovering new mechanics of *decentralized price equilibria* in the *absence of convexity*. In particular, the usage of our basic/limiting normal cone, initiated by Cornet [180] and Khan [412] (see also Khan's preprint of 1987 referred to in [412]) produced the most adequate version of the *second welfare theorem* for the *marginal price* equilibrium in nonconvex finite-dimensional models. The application of the *extremal principle* to these issues was suggested in Mordukhovich [515] providing new versions of such results in both finite and infinite dimensions. This approach led the author [521] to developing a *nonlinear price* mechanism supporting a *limiting decentralized* (maxmin type) equilibrium in fully nonconvex competitive models; see [523] for further details. More recent results in these and related directions of microeconomics can be found, e.g., in [56, 57, 76, 98, 99, 265, 266, 270, 322, 392, 393, 409, 722].

Our approach in Chapter 10 employs the aforementioned variational techniques while being significantly different from all the previous developments. It is based on the paper by Bao and Mordukhovich [56], which establishes two-sided relationships between models of welfare economics and *set-valued optimization*. This approach generates new notions and results in both of these areas that are strongly interrelated and in fact are motivated by each other as shown above.

Section 10.1. The basic model of welfare economics formulated in Subsection 10.1.1 is rather conventional now in microeconomic; see, e.g., the books [11, 490]. Note that introducing here the "net demand constraint set" W as in [484, 515, 523] allows us to unify various market requirements (markets clear, implicit free disposal of commodities, etc.) and to take into account a possible incomplete/uncertain information on the initial aggregate endowment of scarce resources. On other hand, it doesn't create additional mathematical difficulties in comparison with the treatment of nonconvex production sets in the model under consideration.

The notions of (local and global) Pareto and weak Pareto optimal allocations in Definition 10.1 are pretty standard in welfare economics , while it is not the case of the strict and strong Pareto ones therein. To the best of our knowledge, the *strong Pareto* notion first appeared in Khan [411], and *strict Pareto* optimal allocations were defined but not investigated in [523, Remark 8.15]. Both of these notions are economically meaningful and play a significant role in our study of welfare economies from the viewpoint of constrained set-valued optimization.

Observe to this end that over the years, Pareto-type optimal allocations in models of welfare economics and Pareto-type optimal solutions in multiobjective optimization have been studied *separately*, without establishing any connections between them. Most probably it is due to the fact that multiobjective objective optimization dealt mainly with (single-objective) *vector* problems. As shown above in Theorem 10.4, there exists an *equivalence* between the local optimal allocations under consideration in the welfare economic model and the corresponding *set-valued* optimization problem with the explicit *geometric constraint*. As one can see, there is *no way* of making the equivalent optimization problem (10.9) to be either vector-objective or unconstrained. Furthermore, the "minimization" in (10.3) and thus in (10.9) is defined not in a conventional route via some ordering cone as in Chapter 9 but via the *level-set* preference relation (10.4) induced by the preference mappings (10.10) of the welfare economic model $\mathcal{E}$.

To proceed with the study of economically meaningful notions of local Pareto optimal allocations in the welfare economy $\mathcal{E}$ via multiobjective optimization, we need to introduce *new concepts* of local optimal solutions to constrained set-valued optimization problems of type (10.3). This is done in Definition 10.3 via the notions of *fully localized minimizers* that take into account the set-valued nature of the cost mapping in (10.3). The equivalence Theorem 10.4 tells us that local *weak Pareto* and *strong Pareto* optimal allocations of $\mathcal{E}$ are in agreement with the corresponding fully localized notions for set-valued optimization problem (10.9), while fully localized *Pareto* solutions to (10.9) reduce to local *strict Pareto* optimal allocations of $\mathcal{E}$. It is a challenging *open question* to find a set-valued optimization counterpart of local Pareto optimal allocations in models of welfare economics; see Exercise 10.25.

Section 10.2. This section presents *necessary optimality conditions* for *fully localized* minimizers of the constrained *set-valued optimization* problem (10.3) with the nonstandard preference relation (10.4) therein. Such conditions are certainly of their own interest for multiobjective optimization, independently of applications to welfare economics while being strongly motivated by these applications. The reader would not be surprised that we derive such conditions by using an appropriate version of the *extremal principle* for set systems, which is the underlying tool of our dual-space variational analysis. Due to the structures in (10.3) and (10.4), the most appropriate information in this setting is provided by the *product* extremal principle from Lemma 10.6, which is a small extension of [523, Lemma 5.58].

To apply the product extremal principle to the set system that naturally arises in the proof of Theorem 10.9, we need to check the extremality of this system for each kind of fully localized minimizers described in Definition 10.3. To proceed in this direction, a new (local) *asymptotic closedness* property of sets was introduced in [56] being motivated by applications to welfare economics and then was studied in [59, 64, 409] and other publications. This property has nothing to do with the standard local closedness of sets while extending and unifying previously known "asymptotic" properties of this type; see, e.g., [56, 523] and the references therein. Theorem 10.9 taken from [56] tells us the asymptotic closedness property imposed on different sets, which are associated with the multiobjective optimization problem in (10.3) and (10.4), distinguishes the three kinds of fully localized minimizers from Definition 10.3. Apart from that, the necessary optimality conditions of Theorem 10.9 are identical for all the types of fully localized minimizers under consideration.

Section 10.3. In this section, we present several extended versions of the *second welfare theorem* for nonconvex economies in Banach spaces. So much has been done and written on this topic that there is no need to reproduce it in this book. We refer the reader to [412, 490, 523] and the large bibliographies therein for detailed discussions and historical remarks. Acknowledging it, the material of this section based on [56] has something new to add to the previous developments.

To the best of our knowledge, Theorem 10.10 taken from [56] provides the most advanced versions of the second welfare theorem for *local* weak Pareto optimal allocations of nonconvex economies in Asplund spaces together with the new results for local strict and local strong Pareto optimal allocations in this settings. These results are derived in our approach as direct consequences of the established necessary optimality conditions for constrained multiobjective problems with level-set preferences and the equivalence between the aforementioned Pareto-type optimal allocations in welfare economics and the fully localized solutions in set-valued optimization.

The obtained results easily yield the marginal price *positivity* $p^* \in E^*_+ \setminus \{0\}$ when the commodity space E is *ordered* by its positive cone E_+ with respect to the given preference (10.31). If furthermore the positive cone E_+ is *generating*, i.e., $E_+ = E_+ = E$, then the underlying asymptotic closedness assumption at the optimal allocation in question holds *automatically* under the validity of either the *free disposal of production*, or the *implicit free disposal of commodities*, or the *desirability condition*. All these conditions have been well recognized in microeconomics; see, e.g., [180, 412, 490]. In this way, we arrive at the enhanced version of the second welfare theorem for local *strong* Pareto optimal allocations, which was first established in [521, 523] by a direct proof, and at the new result for local *strict* Pareto optimal allocations of the welfare economy $\mathcal{E}$; see Corollary 10.12.

The last subsection of this section concerns local *weak* Pareto optimal allocations in connection with the *uniform properness* conditions introduced by Mas-Collel [489, 490] and then largely developed by Florenzano, Gourdel, and Jofré [270] among others. Following [56], we show that the such properness properties (viz., their most advanced forms taken from [270]) yield the asymptotic closedness property of the corresponding sets, while not vice versa. This allows us to derive improved versions of the second welfare theorem for local weak Pareto optimal allocation of nonconvex economies with ordered Asplund commodity spaces; see Theorem 10.17.

Section 10.4. The final section of this chapter (and of the whole book) deals with refined versions nonconvex models of second welfare theorem for *global* Pareto optimal allocations of *all the four* types introduced in Definition 10.2. The results obtained are similar to those for local Pareto optimal allocations in Theorem 10.10 but with replacing the asymptotic closedness properties by the corresponding *net demand qualification conditions*. The major advantage of the "global" Theorem 10.21 in comparison with its "local" counterpart in Theorem 10.10 is that now we cover (global) *Pareto optimal allocations* from Definition 10.2(ii,v) contrary to the case of Theorem 10.10; see more discussions in Remark 10.22.

The proof of Theorem 10.21 doesn't involve any reduction to problems of set-valued optimization while applying directly the product extremal principle and SNC calculus. Both NDWQ and NDQ from Definition 10.18 appeared in Mordukhovich [515], while the equivalent version of NDQ was formulated and used by Jofré under the name of "asymptotically included condition"; see [180, 391, 411, 412, 515, 523] and the references therein for previously known qualification conditions of this type. The other two conditions, NDSQ and NDSNQ, were defined in [56] for global strict Pareto and strong Pareto optimal allocations. Refined versions of these conditions were used in [57] to establish enhanced versions of the second welfare theorem for nonconvex economies with Asplund commodity spaces.

Note in conclusion that the difference between necessary optimality conditions for *local* and *global* minimizers has been realized in optimization theory but mainly for infinite-dimensional problems such as the classical calculus of variations and optimal control. Quite recently [194], Dempe and Dutta revealed a striking dissimilarity between local and global solutions for problems of *bilevel programming* finite dimensions. To the best of our knowledge, a clear discrepancy between local and global solutions *microeconomic modeling* in the setting of the second welfare theorem was first illuminated by Bao and Mordukhovich [56].

Section 10.5. The material presented in this section is complementary to the main results of the chapter. Besides rather simple exercises, there are major *unsolved* problems formulated in Exercises 10.24, 10.26, and 10.44. A number of exercises (with the hints therein) relate to the results given in [523, Chapter 8], where the reader can find more details and discussions.

References

1. W. van Ackooij and R. Henrion (2017), (Sub-)gradient formulae for probability functions of random inequality systems under Gaussian distribution, *SIAM/ASA J. Uncert. Qualif.* **5**, 63–87.
2. L. Adam, M. Červinka and M. Pištěk (2016), Normally admissible stratifications and calculation of normal cones to a finite union of polyhedral sets, *Set-Valued Var. Anal.* **24**, 207–229.
3. L. Adam, R. Henrion and J. V. Outrara (2017), On M-stationarity conditions in MPECs and the associated qualification conditions, *Math. Program.*, DOI 10.1007/s10107-017-1146-3.
4. L. Adam and T. Kroupa (2017), The intermediate set and limiting superdifferential for coalitional games: between the core and the Weber set, *Int. J. Game Theory* **46**, 891–918.
5. L. Adam and J. V. Outrata (2014), On optimal control of a sweeping process coupled with an ordinary differential equation, *Disc. Cont. Dyn. Syst. Ser. B*, **19**, 2709–2738.
6. S. Adly, R. Cibulka and H. Massias (2013), Variational analysis and generalized equations in electronic: stability and simulation issues, *Set-Valued Anal.* **21**, 333–358.
7. S. Adly, F. Nacry and L. Thibault (2016), Preservation of prox-regularity of sets with applications to constrained optimization, *SIAM J. Optim.* **26**, 448–473.
8. S. Adly and J. V. Outrata (2013), Qualitative stability of a class of non-monotone variational inclusions: application in electronics, *J. Convex Anal.* **20**, 43–66.
9. S. Adly and A. Seeger (2011), A nonsmooth algorithm for cone-constrained eigenvalue problems, *Comput. Optim. Appl.* **49**, 299–318.
10. S. Al-Homidan, Q. H. Ansari and J.-C. Yao (2008), Some generalizations of Ekeland-type variational principle with applications to equilibrium problems and fixed point theory, *Nonlinear Anal.* **69**, 126–139.

B. S. Mordukhovich, *Variational Analysis and Applications*, Springer Monographs in Mathematics, https://doi.org/10.1007/978-3-319-92775-6

11. C. D. Aliprantis and K. C. Border (2006), *Infinite Dimensional Analysis: A Hitchhiker's Guide*, Springer, Berlin.
12. F. Alizadeh and D. Goldfarb (2003), Second-order cone programming, *Math. Program.* **95**, 3–51.
13. J. Alonso H. Martini and M. Spirova (2012), Minimal enclosing discs, circumcircles, and circumcenters in normed planes, part I, *Comput. Geom.* **45**, 258–274.
14. D. T. V. An and N. D. Yen (2015), Differential stability of convex optimization problems under inclusion constraints, *Applic. Anal.* **94**, 108–128.
15. E. J. Anderson and P. Nash (1987), *Linear Programming in Infinite-Dimensional Spaces*, Wiley, Chichester, United Kingdom.
16. R. Andreani, G. Haeser, M. L. Schuverdt and P. J. S. Silva (2012), Two new weak constraint qualifications and applications, *SIAM J. Optim.* **22**, 1109–1135.
17. Q. H. Ansari, C. S. Lalitha and M. Mehta (2014), *Generalized Convexity, Nonsmooth variational inequalities, and Nonsmooth Optimization*, CRC Press, Boca Raton, Florida.
18. M. Apetrii, M. Durea and R. Strugariu (2013), On subregularity properties of set-valued mappings, *Set-Valued Var. Anal.* **21**, 93–126.
19. F. J. Aragón Artacho, A. L. Dontchev, M. Gaydu, M. H. Geoffroy and V. M. Veliov (2011), Metric regularity of Newton's iteration, *SIAM J. Control Optim.* **49** (2011), 339–362.
20. F. J. Aragón Artacho and M. H. Geoffroy (2008), Characterizations of metric regularity of subdifferentials, *J. Convex Anal.* **15**, 365–380.
21. F. J. Aragón Artacho and M. H. Geoffroy (2014), Metric subregularity of the convex subdifferential in Banach spaces, *J. Nonlinear Convex Anal.* **15**, 35–47.
22. F. J. Aragón Artacho and B. S. Mordukhovich (2010), Metric regularity and Lipschitzian stability of parametric variational systems, *Nonlinear Anal.* **72**, 1149–1170.
23. F. J. Aragón Artacho and B. S. Mordukhovich (2011), Enhanced metric regularity and Lipschitzian properties of variational systems, *J. Global Optim.* **50**, 145–167.
24. A. V. Arutyunov (2005), Covering of nonlinear maps on cone in neighborhood of abnormal point, *Math. Notes* **77**, 447–460.
25. A. V. Arutyunov (2007), Covering mappings in metric spaces and fixed points, *Dokl. Math.* **6**, 665–668.
26. A. V. Arutyunov, E. R. Avakov, B. D. Gelman, A. V. Dmitruk and V. V. Obukhovskii (2009), Locally covering maps in metric spaces and coincidence points, *J. Fixed Point Theory Appl.* **5**, 106–127.
27. A. V. Arutyunov, E. R. Avakov and A. F. Izmailov (2007), Directional regularity and metric regularity, *SIAM J. Optim.* **18**, 810–833.
28. A. V. Arutyunov, E. R. Avakov and S. E. Zhukovskiy (2015), Stability theorem for estimating the distance to a set of coincidence points, *SIAM J. Optim.* **25**, 807–828.

29. A. V. ARUTYUNOV AND A. F. IZMAILOV (2006), Directional stability theorem and directional metric regularity, *Math. Oper. Res.* **31**, 526–543.
30. A. V. ARUTYUNOV AND S. E. ZHUKOVSKIY (2010), The existence of inverse mappings and their properties, *Proc. Steklov Inst. Math.* **271**, 12–22.
31. H. ATTOUCH, J. BOLTE, P. REDONT AND A. SOUBEYRAN (2010), Proximal alternating minimization and projection methods for nonconvex problems: an approach based on the Kurdyka-Łojasiewicz inequality, *Math. Oper. Res.* **35** (2010), 438–457.
32. H. ATTOUCH, J. BOLTE AND B. F. SVAITER (2013), Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward-backward splitting, and regularized Gauss-Seidel methods, *Math. Program.* **137** (2013), 91–129.
33. H. ATTOUCH, G. BUTTAZZO AND G. MICHAILLE (2014), *Variational Analysis in Sobolev and BV Spaces*, 2nd edition, SIAM, Philadelphia, Pennsylvania.
34. J.-P. AUBIN (1981), Contingent derivatives of set-valued maps and existence of solutions to nonlinear inclusions and differential inclusions, in *Mathematical Analysis and Applications*, edited by L. Nachbin, pp. 159–229, Academic Press, New York.
35. J.-P. AUBIN (1984), Lipschitz behavior of solutions to convex minimization problems, *Math. Oper. Res.* **9**, 87–111.
36. J.-P. AUBIN AND H. FRANKOWSKA (1990), *Set-Valued Analysis*, Birkhäuser, Boston, Massachusetts.
37. A. AUSLENDER AND M. TEBOULLE (2003), *Asymptotic Cones and Functions in Optimization and Variational Inequalities*, Springer, New York.
38. D. AUSELL, M. ČERVINKA AND M. MARÉCHAL (2016), Deregulated electricity markets with thermal losses and production bounds: models and optimality conditions, *RAIRO-Oper. Res.* **50**, 19–38.
39. D. AUSSEL, J.-N. CORVELLEC AND M. LASSONDE (1995), Mean value property and subdifferential criteria for lower semicontinuous functions, *Trans. Amer. Math. Soc.* **347**, 4147–4161.
40. D. AUSSEL, A. DANIILIDIS AND L. THIBAULT (2005), Subsmooth sets: functional characterizations and related concepts, *Trans. Amer. Math. Soc.* **357**, 1275–1301.
41. D. AUSSEL, Y. GARCIA AND N. HADJISAVVAS (2009), Single-directional property of multivalued maps and variational systems, *SIAM J. Optim.* **20**, 1274–1285.
42. D. AZÉ AND J.-N. CORVELLEC (2004), Characterization of error bounds for lower semicontinuous functions on metric spaces, *ESAIM Control Optim. Calc. Var.* **10**, 409–425.
43. D. AZÉ AND J.-N. CORVELLEC (2014), Nonlinear local error bounds via a change of metric, *J. Fixed Point Theory Appl.* **16**, 351–372.
44. D. AZÉ, J.-N. CORVELLEC AND R. LUCCHETTI (2002), Variational pairs and applications to stability in nonsmooth analysis, *Nonlinear Anal.* **49**, 643–670.

45. M. BACÁK, J. M. BORWEIN, A. EBERHARD AND B. S. MORDUKHOVICH (2010), Infimal convolutions and Lipschitzian properties of subdifferentials for prox-regular functions in Hilbert spaces, *J. Convex Anal.* **17** (2010), 737–763.
46. R. BAIER, E. FARKHI AND V. ROSHCHINA (2010), On computing the Mordukhovich subdifferential using directed sets in two dimensions, in *Variational Analysis and Generalized Differentiation in Optimization and Control*, edited by R. S. Burachik and J.-C. Yao, pp. 59–93, Springer, New York.
47. R. BAIER, E. FARKHI AND V. ROSHCHINA (2012), The directed and Rubinov subdifferentials of quasidifferentiable functions, I: definition and examples, *Nonlinear Anal.* **75**, 1074–1088.
48. R. BAIER, E. FARKHI AND V. ROSHCHINA (2012), The directed and Rubinov subdifferentials of quasidifferentiable functions, II: calculus, *Nonlinear Anal.* **75**, 1058–1073.
49. L. BAN, B. S. MORDUKHOVICH AND W. SONG (2011), Lipschitzian stability of parametric variational inequalities over generalized polyhedra in Banach spaces, *Nonlinear Anal.* **74**, 441–461.
50. T. Q. BAO (2013), On a nonconvex separation theorem and the approximate extremal principle in Asplund spaces, *Acta Math. Vietnam.* **38**, 279–291.
51. T. Q. BAO, P. GUPTA AND B. S. MORDUKHOVICH (2007), Necessary conditions for multiobjective optimization with equilibrium constraints, *J. Optim. Theory Appl.* **135**, 179–203.
52. T. Q. BAO AND B. S. MORDUKHOVICH (2007), Variational principles for set-valued mappings with applications to multiobjective optimization, *Control Cybern.* **36**, 531–562.
53. T. Q. BAO AND B. S. MORDUKHOVICH (2007), Existence of minimizers and necessary conditions in set-valued optimization with equilibrium constraints, *Appl. Math.* **52**, 453–472.
54. T. Q. BAO AND B. S. MORDUKHOVICH (2009), Necessary conditions for super minimizers in constrained multiobjective optimization, *J. Global Optim.* **43**, 533–552.
55. T. Q. BAO AND B. S. MORDUKHOVICH (2010), Relative Pareto minimizers for multiobjective problems: existence and optimality conditions, *Math. Program* **122**, 301–347.
56. T. Q. BAO AND B. S. MORDUKHOVICH (2010), Set-valued optimization in welfare economics, *Adv. Math. Econ.* **13**, 114–153.
57. T. Q. BAO AND B. S. MORDUKHOVICH (2011), Refined necessary conditions in multiobjective optimization with applications to microeconomic modeling, *Discrete Contin. Dyn. Syst.* **31**, 1069–1096.
58. T. Q. BAO AND B. S. MORDUKHOVICH (2012), Sufficient conditions for global weak Pareto solutions in multiobjective optimization, *Positivity* **16**, 579–602.
59. T. Q. BAO AND B. S. MORDUKHOVICH (2012), Extended Pareto optimality in multiobjective problems, in *Recent Developments in Vector Optimization*, edited by Q. H. Ansari and J.-C. Yao, pp. 467–515, Springer, Berlin.

60. T. Q. Bao and B. S. Mordukhovich (2014), Necessary nondomination conditions in set and vector optimization with variable ordering structures, *J. Optim. Theory Appl.* **162**, 350–370.
61. T. Q. Bao, B. S. Mordukhovich and A. Soubeyran (2015), Variational analysis in psychological modeling, *J. Optim. Theory Appl.* **164**, 290–315.
62. T. Q. Bao, B. S. Mordukhovich and A. Soubeyran (2015), Fixed points and variational principles with applications to capability theory of wellbeing via variational rationality, *Set-Valued Var. Anal.* **23**, 375–398.
63. T. Q. Bao, B. S. Mordukhovich and A. Soubeyran (2015), Minimal points, variational principles, and variable preferences in set optimization, *J. Nonlinear Convex Anal.* **16**, 1511–1537.
64. T. Q. Bao and C. Tammer (2012), Lagrange necessary conditions for Pareto minimizers in Asplund spaces and applications, *Nonlinear Anal.* **75**, 1089–1103.
65. A. Barbagallo and P. Mauro (2016), A general quasi-variational problem of Cournot-Nash type and its inverse formulation, *J. Optim. Theory Appl.* **170** (2016), 476–492.
66. M. Bardi (1989), A boundary value problem for the minimal time function, *SIAM J. Control Optim.* **27**, 776–785.
67. M. Bardi and I. Capuzzo Dolcetta (1997), *Optimal Control and Viscosity Solutions of Hamilton-Jacobi Equations*, Birkhäuser, Boston, Massachusetts.
68. P. I. Barton, K. A. Khan, P. G. Stechinski and H. A. J. Watson (2017), Computationally relevant generalized derivatives: theory, evaluation and applications, *Optim. Methods Softw.*, DOI 10.1080/10556788.2017.1374385
69. H. H. Bauschke, J. M. Borwein and W. Li (1999), Strong conical hull intersection property, bounded linear regularity, Jameson's property (G), and error bounds in convex optimization, *Math. Program.* **86**, 135–160.
70. H. H. Bauschke and P. L. Combettes (2017), *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, 2nd edition, Springer, New York.
71. H. H. Bauschke, D. R. Luke, H. M. Phan and X. Wang (2013), Restricted normal cones and the method of alternating projections: theory, *Set-Valued Var. Anal.* **21**, 431–473.
72. H. H. Bauschke, D. R. Luke, H. M. Phan and X. Wang (2013), Restricted normal cones and the method of alternating projections: applications, *Set-Valued Var. Anal.* **21**, 475–501.
73. H. H. Bauschke, D. R. Luke, H. M. Phan and X. Wang (2014), Restricted normal cones and sparsity optimization with affine constraints, *Found. Comput. Math.* **14**, 63–83.
74. M. S. Bazaraa, J. J. Goode and M. Z. Nashed (1974), On the cone of tangents with applications to mathematical programming, *J. Optim. Theory Appl.***13**, 389–426.
75. E. M. Bednarczuk and D. Zagrodny (2009), Vector variational principle, *Arch. Math.* **93**, 577–586.

76. S. BELLAASSALI AND A. JOURANI (2008), Lagrange multipliers for multiobjective programs with a general preference, *Set-Valued Anal.* **16**, 229–243.
77. R. E. BELLMAN (1957), *Dynamic Programming*, Princeton University Press, Princeton, New Jersey.
78. M. BENKO AND H. GFRERER (2017), New verifiable stationarity concepts for a class of mathematical programs with disjunctive constraints, to appear in *Optimization*, arXiv https://arxiv.org/pdf/1611.08206.pdf.
79. F. BENITA, S. DEMPE AND P. MEHLITZ (2016), Bilevel optimal control problems with pure state constraints and finite-dimensional lower level, *SIAM J. Optim.* **26**, 564–588.
80. F. BENITA AND P. MEHLITZ (2016), Bilevel optimal control with final-state-dependent finite-dimensional lower level, *SIAM J. Optim.* **26**, 718–752.
81. G. C. BENTO AND A. SOUBEYRAN (2015), A generalized inexact proximal point method for nonsmooth functions that satisfies Kurdyka-Łojasiewicz inequality, *Set-Valued Var. Anal.* **23** (2015), 501–517.
82. P. BEREMLIJSKI, J. HASLINGER, J. V. OUTRATA AND R. PATHÓ (2014), Shape optimization in contact problems with Coulomb friction and a solution-dependent friction coefficient, *SIAM J. Control Optim.* **52**, 3371–3400.
83. F. BERNARD AND L. THIBAULT (2004), Prox-regularity of functions and sets in Banach spaces, to appear in *Set-Valued Anal.* **12**, 25–47.
84. D. P. BERTSEKAS AND A. E. OZDAGLAR (2002), Pseudonormality and a Lagrange multiplier theory for constrained optimization, *J. Optim. Theory Appl.* **114**, 287–343.
85. M. BIANCHI, N. HADJISAVVAS AND S. SCHAIBLE (1997), Vector equilibrium problems with generalized monotone bifunctions, *J. Optim. Theory Appl.* **92**, 527–542.
86. M. BIANCHI, G. KASSAY AND R. PINI (2013), An inverse map result and some applications to sensitivity of generalized equations, *J. Math. Anal. Appl.* **399**, 279–290.
87. M. BIANCHI, G. KASSAY AND R. PINI (2016), Linear openness of the composition of set-valued maps and applications to variational systems, *Set-Valued Var. Anal.* **24**, 581–595.
88. G. BIGI, M. CASTELLANI, M. PAPPALARDO AND M. PASSACANTANDO (2013), Existence and solution methods for equilibria, *European J. Oper. Res.* **227**, 1–11.
89. E. BLUM AND W. OETTLI (1994), From optimization and variational inequalities to equilibrium problems, *Math. Student* **63**, 123–145.
90. J. BOLTE, A. DANIILIDIS AND A. S. LEWIS (2006), The Morse-Sard theorem for nondifferentiable subanalytic functions, *J. Math. Anal. Appl.* **321**, 729–740.
91. J. BOLTE, T. P. NGUYEN, J. PEYPOUQUET AND B. W. SUTER (2017), From error bounds to the complexity of first-order descent methods for convex functions, *Math. Program.* **165**, 471–507.

92. J. BOLTE, S. SABACH AND M. TEBOULLE (2014), Proximal alternating linearized minimization for nonconvex and nonsmooth problems, *Math. Program* **146**, 459–494.
93. V. BONDAREVSKY, A. LESCHOV AND L. MINCHENKO (2016), Value functions and their directional derivatives in parametric nonlinear programming, *J. Optim. Theory Appl.* **171**, 440–464.
94. J. F. BONNANS (1994), Local analysis of Newton-type methods for variational inequalities and nonlinear programming, *Appl. Math. Optim.* **29**, 161–186.
95. J. F. BONNANS AND H. RAMÍREZ C. (2005), Perturbation analysis of second-order cone programming problems, *Math. Program.* **104**, 205–227.
96. J. F. BONNANS AND A. SHAPIRO (2000), *Perturbation Analysis of Optimization Problems*, Springer, New York.
97. H. BONNEL, A. N. IUSEM AND B. F. SVAITER (2005), Proximal methods in vector optimization, *SIAM J. Optim.* **15**, 953–970.
98. J.-M. BONNISSEAU, B. CORNET AND M.-O. CZARNECKI (2007), The marginal pricing rule revisited, *Econom. Theory* **33**, 579–589.
99. J.-M. BONNISSEAU AND O. LACHIRI (2006), About the second theorem of welfare economics with stock markets, *Pac. J. Optim.* **2**, 469–485.
100. J.-M. BONNISSEAU AND C. LE VAN (1996), On the subdifferential of the value function in economic optimization problems, *J. Math. Econom.* **25**, 55–73.
101. J. M. BORWEIN AND S. P. FITZPATRICK (1995), Weak* sequential compactness and bornological limit derivatives, *J. Convex Anal.* **2**, 59–68.
102. J. M. BORWEIN AND A. JOFRÉ (1988), Nonconvex separation property in Banach spaces, *Math. Methods Oper. Res.* **48**, 169–179.
103. J. M. BORWEIN AND R. GOEBEL (2003), Notions of relative interior in Banach spaces, *J. Math. Sci.* **115**, 2542–2553.
104. J. M. BORWEIN AND A. S. LEWIS (1992), Partially finite convex programming, I: quasi relative interiors and duality theory, *Math. Prog.* **57**, 15–48.
105. J. M. BORWEIN AND A. S. LEWIS (2000), *Convex Analysis and Nonlinear Optimization: Theory and Examples*, 2nd edition, Springer, New York.
106. J. M. BORWEIN, Y. LUCET AND B. S. MORDUKHOVICH (2000), Compactly epi-Lipschitzian convex sets and functions in normed spaces, *J. Convex Anal.* **7**, 375–393.
107. J. M. BORWEIN, B. S. MORDUKHOVICH AND Y. SHAO (1999), On the equivalence of some basic principles of variational analysis, *J. Math. Anal. Appl.* **229**, 228–257.
108. J. M. BORWEIN AND D. PREISS (1987), A smooth variational principle with applications to subdifferentiability and differentiability of convex functions, *Trans. Amer. Math. Soc.* **303**, 517–527.
109. J. M. BORWEIN AND H. M. STRÓJWAS (1985), Tangential approximations, *Nonlinear Anal.* **9**, 1347–1366.

110. J. M. BORWEIN AND H. M. STRÓJWAS (1987), Proximal analysis and boundaries of closed sets in Banach spaces, II: applications, *Canad. J. Math.* **39**, 428–472.
111. J. M. BORWEIN, J. S. TREIMAN AND Q. J. ZHU (1998), Necessary conditions for constrained optimization problems with semicontinuous and continuous data, *Trans. Amer. Math. Soc.* **350**, 2409–2429.
112. J. M. BORWEIN AND J. D. VANDERWERFF (2010), *Convex Functions*, Cambridge University Press, Cambridge, United Kingdom.
113. J. M. BORWEIN AND Q. J. ZHU (1999), A survey of subdifferential calculus with applications, *Nonlinear Anal.* **38**, 687–773.
114. J. M. BORWEIN AND Q. J. ZHU (2005), *Techniques of Variational Analysis*, Springer, New York.
115. J. M. BORWEIN AND D. ZHUANG (1993), Super efficiency in vector optimization, *Trans. Amer. Math. Soc.* **338**, 105–122.
116. R. I. BOŢ (2010), *Conjugate Duality in Convex Optimization*, Springer, Berlin.
117. R. I. BOŢ (2012), An upper estimate for the Clarke subdifferential of an infimal value function proved via the Mordukhovich subdifferential, *Nonlinear Anal.* **75**, 1141–1146.
118. R. I. BOŢ AND E. R. CSETNEK (2016), An inertial Tseng's type proximal algorithm for nonsmooth and nonconvex optimization problems, *J. Optim. Theory Appl.* **171**, 600–616.
119. R. I. BOŢ, E. R. CSETNEK AND G. WANKA (2008), Regularity conditions via quasi-relative interior in convex programming, *SIAM J. Optim.* **19**, 217–233.
120. R. I. BOŢ, S. M. GRAD AND G. WANKA (2008), New regularity conditions for strong and total Fenchel-Lagrange duality in infinite-dimensional spaces, *Nonliner Anal.* **69**, 323–336.
121. R. I. BOŢ, S. M. GRAD AND G. WANKA (2009), Generalized Moreau-Rockafellar results for composed convex functions, *Optimization* **58**, 917–938.
122. R. I. BOŢ AND C. HENDRICH (2013), A Douglas-Rachford type primal-dual method for solving inclusions with mixtures of composite and parallel-sum type monotone operators, *SIAM J. Optim.* **23**, 2541–2565.
123. G. BOULIGAND (1930), Sur les surfaces dépourvues de points hyperlimits, *Ann. Soc. Polon. Math.* **9**, 32–41 (in French).
124. M. BOUNKHEL (2012), *Regularity Concepts in Nonsmooth Analysis: Theory and Applications*, Springer, New York.
125. M. BOUNKHEL AND L. THIBAULT (2002), On various notions of regularity of sets in nonsmooth analysis, *Nonlinear Anal.* **48**, 223–246.
126. H. BRÉZIS (1973), *Opérateurs Maximaux Monotones et Semigroupes de Contractions dans les Espaces de Hilbert*, North-Holland, Amsterdam.
127. L. M. BRICEÑO-ARIAS, N. D. HOANG AND J. PEYPOUQUET (2016), Existence, stability and optimality for optimal control problems governed by maximal monotone operators, *J. Diff. Eqs.* **260**, 733–757.

128. S. BUNDFUSS AND M. DÜR (2009), An adaptive linear approximation algorithm for copositive programs, *SIAM J. Optim.* **20**, 30–53.
129. R. W. BROCKETT (1983), Asymptotic stability and feedback stabilization, in *Differential Geometric Control Theory* edited by R. W. Brockett et al., pp. 181–191, Birkhäuser, Boston, Massachusetts.
130. R. S. BURACHIK AND A. N. IUSEM (2008), *Set-Valued Mappings and Enlargements of Monotone Operators*, Springer, New York.
131. R. S. BURACHIK AND V. JEYAKUMAR (2005), A dual condition for the convex subdifferential sum formula with applications, J. Convex Anal. **12**, 279–290.
132. J. V. BURKE AND S. DENG (2005), Weak sharp minima revisited, II: applications to linear regularity and error bounds, *Math. Program.* **104**, 235–261.
133. J. V. BURKE AND M. C. FERRIS (1993), Weak sharp minima in mathematical programming, *SIAM J. Control Optim.* **31**, 1340–1359.
134. J. V. BURKE AND M. L. OVERTON (2001), Variational analysis of non-Lipschitz spectral functions, *Math. Program.* **90**, 317–352.
135. A. CABOT AND L. THIBAULT (2014), Sequential formulae for the normal cone to sublevel sets, *Trans. Amer. Math. Soc.* **366**, 6591–6628.
136. P. CANNARSA AND C. SINESTRARI (2004), *Semiconvex Functions, Hamilton-Jacobi Equations, and Optimal Control*, Birkhäuser, Boston, Massachusetts.
137. M. J. CÁNOVAS, A. L. DONTCHEV, M. A. LÓPEZ AND J. PARRA (2005), Metric regularity of semi-infinite constraint systems, *Math. Program.* **104**, 329–346.
138. M. J. CÁNOVAS, R. HENRION, M. A. LÓPEZ AND J. PARRA (2016), Outer limit of subdifferentials and calmness moduli in linear and nonlinear programming, *J. Optim. Theory Appl.* **169**, 925–952.
139. M. J. CÁNOVAS, A. Y. KRUGER, M. A. LÓPEZ, J. PARRA AND M. THÉRA (2014), Calmness modulus of linear semi-infinite programs, *SIAM J. Optim.* **24**, 29–48.
140. M. J. CÁNOVAS, M. A. LÓPEZ, B. S. MORDUKHOVICH AND J. PARRA (2009), Variational analysis in semi-infinite and infinite programming, I: stability of linear inequality systems of feasible solutions, *SIAM J. Optim.* **20**, 1504–1526.
141. M. J. CÁNOVAS, M. A. LÓPEZ, B. S. MORDUKHOVICH AND J. PARRA (2010), Variational analysis in semi-infinite and infinite programming, II: necessary optimality conditions, *SIAM J. Optim.* **20**, 2788–2806.
142. M. J. CÁNOVAS, M. A. LÓPEZ, B. S. MORDUKHOVICH AND J. PARRA (2012), Quantitative stability of linear infinite inequality systems under block perturbations with applications to convex systems, *TOP* **20**, 310–327.
143. T. H. CAO AND B. S. MORDUKHOVICH (2016), Optimal control of a perturbed sweeping process via discrete approximations, *Disc. Cont. Dyn. Syst. Ser. B*, **21**, 3331–3358.

144. T. H. Cao and B. S. Mordukhovich (2017), Optimality conditions for a controlled sweeping process with applications to the crowd motion model, *Disc. Cont. Dyn. Syst. Ser. B*, **22**, 267–306.
145. L. C. Ceng, B. S. Mordukhovich and J.-C. Yao (2010), Hybrid approximate proximal method with auxiliary variational inequality for vector optimization, *J. Optim. Theory Appl.* **146** (2010), 267–303.
146. C. R. Chen, S. J. Li and K. L. Teo (2009), Solution semicontinuity of parametric generalized vector equilibrium problems, *J. Global Optim.* **45**, 309–318.
147. G. Y. Chen, X. X. Huang and X. Q. Yang (2005), *Vector Optimization*, Springer, Berlin.
148. X. Chen (2012), Smoothing methods for nonsmooth, nonconvex minimization, *Math. Program.* **134**, 71–99.
149. N. H. Chieu (2009), The Fréchet and limiting subdifferentials of integral functionals on the spaces $L_1(\Omega, E)$, *J. Math. Anal. Appl.* **360**, 704–710.
150. N. H. Chieu, T. D. Chuong, J.-C. Yao and N. D. Yen (2011), Characterizing convexity of a function by its Fréchet and limiting second-order subdifferentials, *Set-Valued Var. Anal.* **19**, 75–96.
151. N. H. Chieu and L. V. Hien (2017), Computation of graphical derivative for a class of normal cone mappings under a very weak condition, *SIAM J. Optim.* **27**, 190–204.
152. N. H. Chieu and N. Q. Huy (2011), Second-order subdifferentials and convexity of real-valued functions, *Nonlinear Anal.* **74**, 154–160.
153. N. H. Chieu, G. M. Lee, B. S. Mordukhovich and T. T. A. Nghia (2016), Coderivative characterizations of maximal monotonicity for set-valued mappings, *J. Convex Anal.* **23**, 461–480.
154. N. H. Chieu and N. T. Q. Trang (2012), Coderivative and monotonicity of continuous mappings, *Taiwanese J. Math.* **16**, 353–365.
155. N. M. Chieu, J.-C. Yao and N. D. Yen (2010), Relationships between Robinson metric regularity and Lipschitz-like behavior of implicit multifunctions, *Nonlinear Anal.* **72**, 3594–3601.
156. C. K. Chui, F. Deutsch and J. D. Ward (1990), Constrained best approximation in Hilbert space, *Constr. Approx.* **6**, 35–64.
157. T. D. Chuong, N. Q. Huy and J.-C. Yao (2009), Subdifferentials of marginal functions in semi-infinite programming, *SIAM J. Optim.* **20**, 1462–1477.
158. T. D. Chuong and D. S. Kim (2015), Hölder-like property and metric regularity of a positive-order for implicit multifunctions, *Math. Oper. Res.* **41**, 596–611.
159. T. D. Chuong, B. S. Mordukhovich and J.-C. Yao (2011), Hybrid approximate proximal algorithms for efficient solutions in vector optimization, *J. Nonlinear Convex Anal.* **12**, 257–286.
160. R. Cibulka, A. D. Dontchev and A. Y. Kruger (2018), Strong metric subregularity of mappings in variational analysis and optimization, *J. Math. Anal. Appl.* **457**, 1247–1282.

161. R. Cibulka and M. Fabian (2013), A note on Robinson-Ursescu and Lyusternik–Graves theorems, *Math. Program.* **139**, 89–101.
162. F. H. Clarke (1973), *Necessary Conditions for Nonsmooth Problems in Optimal Control and the Calculus of Variations*, Ph.D. dissertation, Department of Mathematics, University of Washington, Seattle, Washington.
163. F. H. Clarke (1975), Generalized gradients and applications, *Trans. Amer. Math. Soc.* **205**, 247–262.
164. F. H. Clarke (1976), A new approach to Lagrange multipliers, *Math. Oper. Res.* **2**, 165–174.
165. F. H. Clarke (1983), *Optimization and Nonsmooth Analysis*, Wiley-Interscience, New York.
166. F. H. Clarke and Y. S. Ledyaev (1994), Mean value inequality in Hilbert spaces, *Trans. Amer. Math. Soc.* **344**, 307–324.
167. F. H. Clarke, Y. S. Ledyaev, R. J. Stern and P. R. Wolenski (1998), *Nonsmooth Analysis and Control Theory*, Springer, New York.
168. F. H. Clarke, R. J. Stern and P. R. Wolenski (1993), Subgradient criteria for monotonicity and the Lipschitz condition, *Canad. J. Math.* **45**, 1167–1183.
169. C. Clason and T. Valkonen (2017), Stability of saddle points via explicit coderivatives of pointwise subdifferentials, *Set-Valued Var. Anal.* **25**, 69–112.
170. S. Cobzaş (2013), *Functional Analysis in Asymmetric Normed Spaces*, Birkhüser, Basel, Switzerland.
171. G. Colombo, V. V. Goncharov and B. S. Mordukhovich (2010), Well-posedness of minimal time problems with constant dynamics in Banach spaces, *Set-Valued Var. Anal.* **18**, 349–972.
172. G. Colombo, R. Henrion, N. D. Hoang, and B. S. Mordukhovich (2012), Discrete approximations and optimality conditions for optimal control of the sweeping process, *Dyn. Contin. Discrete Impuls. Syst. Ser. B* **19**, 117–159.
173. G. Colombo, R. Henrion, N. D. Hoang, and B. S. Mordukhovich (2015), Discrete approximations of a controlled sweeping process, *Set-Valued Var. Anal.* **23**, 69–86.
174. G. Colombo, R. Henrion, N. D. Hoang, and B. S. Mordukhovich (2016), Optimal control of the sweeping process over polyhedral controlled sets, *J. Diff. Eqs.* **260**, 3397–3447.
175. G. Colombo and L. Thibault (2010), Prox-regular sets and applications, in *Handbook of Nonconvex Analysis*, edited by D. Y. Gao and D. Motreanu, pp. 99-182, International Press, Boston, Massachusetts.
176. G. Colombo and P. R. Wolenski (2004), The subgradient formula for the minimal time functions in the case of constant dynamics in Hilbert spaces, *J. Global Optim.* **28**, 269–282.
177. B. Colson, P. Marcotte and G. Savard (2007), An overview of bilevel optimization, *Ann. Oper. Res.* **153**, 235–256.
178. R. Cominetti (1990), Metric regularity, tangent cones, and second-order optimality conditions, *Appl. Math. Optim.* **21**, 265–287.

179. H. W. CORLEY (1988), Optimality conditions for maximization of set-valued functions, *J. Optim. Theory Appl.* **58**, 1–10.
180. B. CORNET (1990), Marginal cost pricing and Pareto optimality, in *Essays in Honor of Edmond Malinvaud*, edited by P. Champsaur, Vol. 1, pp. 14–53, MIT Press, Cambridge, Massachusetts.
181. R. CORREA, A. JOFRÉ AND L. THIBAULT (1994), Subdifferential monotonicity as characterization of convex functions, *Numer. Funct. Anal. Optim.* **15**, 1167–1183.
182. M. G. CRANDALL, H. ISHII AND P.-L. LIONS (1992), User's guide to viscosity solutions of second-order partial differential equations, *Bull. Amer. Math. Soc.* **27**, 1–67.
183. M. G. CRANDALL AND P.-L. LIONS (1983), Viscosity solutions of Hamilton-Jacobi equations, *Trans. Amer. Math. Soc.* **277**, 1–42.
184. G. P. CRESPI, I. GINCHEV AND M. ROCCA (2006), First-order optimality conditions in set-valued optimization, *Math. Methods Oper. Res.* **63**, 87–106.
185. J.-P. CROUZEIX AND E. OCAÑ A ANAYA (2010), Maximality is nothing but continuity, *J. Convex Anal.* **17** (2010), 521–534.
186. Y. CUI, D. SUN AND K.-C. TOH (2016), On the asymptotic superlinear convergence of the augmented Lagrangian method for semidefinite programming with multiple solutions, arXiv:1610.00875.
187. M. CÚTH AND M. FABIAN (2016), Asplund spaces characterized by rich families and separable reduction of Fréchet subdifferentiability, *J. Funct. Anal.* **270**, 1361–1378.
188. P. DANIELE, S. GIUFFRÉ, G. IDONE AND A. MAUGERI (2007), Infinite dimensional duality and applications, *Math. Ann.* **339**, 221–239.
189. A. DANIILIDIS AND P. GEORGIEV (2004), Approximate convexity and submonotonicity, *J. Math. Anal. Appl.* **291**, 292–301.
190. C. DAVIS AND W. HARE (2013), Exploiting known structures to approximate normal cones, *Math. Oper. Res.* **38**, 665–681.
191. E. DE GIORGI, A. MARINO AND M. TOSQUES (1980), Problemi di evolusione in spazi metrici e curve di massima pendenza, *Atti. Accad. Lincei Rend. Cl. Sci. Fiz. Mat. Natur.* **68**, 180–187 (in Italian).
192. J. E. DENNIS, JR. AND J. J. MORÉ (1974), A characterization of superlinear convergence and its application to quasi-Newton methods, *Math. Comp.* **28**, 549–560.
193. S. DEMPE (2003), *Foundations of Bilevel Programming*, Kluwer, Dordrecht, The Netherlands.
194. S. DEMPE AND J. DUTTA (2012), Is bilevel programming a special case of mathematical programming with complementarity constraints?, *Math. Program.* **131**, 37–48.
195. S. DEMPE, J. DUTTA AND B. S. MORDUKHOVICH (2007), New necessary optimality conditions in optimistic bilevel programming, *Optimization* **56**, 577–604.

196. S. Dempe, N. A. Gadhi and L. Lafhim (2010), Fuzzy and exact optimality conditions for a bilevel set-valued problem via extremal principles, *Numer. Funct. Anal. Optim.* **31**, 907–920.
197. S. Dempe, V. Kalashnikov, G. A. Pérez-Valdés and N. Kalashnykova (2015), *Bilevel Programming Problems*, Springer, New York.
198. S. Dempe, B. S. Mordukhovich and A. B. Zemkoho (2012), Sensitivity analysis for two-level value functions with applications to bilevel programming, *SIAM J. Optimization* **22**, 1309–1343.
199. S. Dempe, B. S. Mordukhovich and A. B. Zemkoho (2014), Necessary optimality conditions in pessimistic bilevel programming, *Optimization* **63**, 505–533.
200. S. Dempe and A. B. Zemkoho (2012), On the Karush-Kuhn-Tucker reformulation of the bilevel optimization problem, *Nonlinear Anal.* **75**, 1202–1218.
201. S. Dempe and A. B. Zemkoho (2013), The bilevel programming problem: reformulations, constraint qualifications and optimality conditions, *Math. Program.* **138**, 447–473.
202. S. Dempe and A. B. Zemkoho (2014), KKT reformulation and necessary conditions for optimality in nonsmooth bilevel optimization, *SIAM J. Optim.* **24**, 1639–1669.
203. V. F. Demyanov and A. M. Rubinov (2000), *Quasidifferentiable and Related Topics*, Kluwer, Dordrecht, The Netherlands.
204. F. Deutsch, W. Li and J. D. Ward (1997), A dual approach to constrained interpolation from a convex subset of Hilbert space, *J. Approx. Theory* **90**, 385–414.
205. R. Deville, G. Godefroy and V. Zizler (1993), *Smoothness and Renorming in Banach Spaces*, Wiley, New York.
206. S. Dias and G. Smirnov (2012), On the Newton method for set-valued maps, *Nonlinear Anal.* **75**, 1219–1230.
207. J. Diestel (1984), *Sequences and Series in Banach Spaces*, Springer, New York.
208. C. Ding, D. Sun and J. J. Ye (2014), First order optimality conditions for mathematical programs with semidefinite cone complementarity constraints, *Math. Program.* **147**, 539–579.
209. C. Ding, D. Sun and L. Zhang (2017), Characterization of the robust isolated calmness for a class of conic programming problems, *SIAM J. Optim.* **27**, 67–90.
210. N. Dinh, M. A. Goberna and M. A. López (2006), From linear to convex systems: consistency, Farkas' lemma and applications, *J. Convex Anal.* **13**, 279–290.
211. N. Dinh, M. A. Goberna and M. A. López (2010), On the stability of the feasible set in optimization problems, *SIAM J. Optim.* **20**, 2254–2280.
212. N. Dinh, M. A. Goberna, M. A. López and T. Q. Son (2007), New Farkas-type results with applications to convex infinite programming, *ESAIM: Control Optim. Cal. Var.* **13**, 580–597.

213. N. Dinh and V. Jeaykumar (2014), Farkas' lemma: three decades of generalizations for mathematical optimization, *TOP* **22**, 1–22.
214. N. Dinh, B. S. Mordukhovich and T. T. A. Nghia (2009), Qualification and optimality conditions for convex and DC programs with infinite constraints, *Acta Math. Vietnamica* **34**, 123–153.
215. N. Dinh, B. S. Mordukhovich and T. T. A. Nghia (2010), Subdifferentials of value functions and optimality conditions for some classes of DC and bilevel infinite and semi-infinite programs, *Math. Program.* **123**, 101–138.
216. A. B. Dmitruk and A. Y. Kruger (2008), Metric regularity and systems of generalized equations, *J. Math. Anal. Appl.* **342**, 864–873.
217. A. V. Dmitruk, A. A. Milyutin and N. P. Osmolovskii (1980), Lyusternik's theorem and the theory of extrema, *Russian Math. Surveys* **35**, 11–51.
218. S. Dolecki and G. H. Greco (2007), Towards historical roots of necessary conditions of optimality: Regula of Peano, *Control Cybern.* **36**, 491–518.
219. M. B. Donato, M. Milasi and C. Vitanza (2008), An existence result of a quasi-variational inequality associated to an equilibrium problem, *J. Global Optim.* **40**, 87–97.
220. A. L. Dontchev (1995), Characterizations of Lipschitz stability in optimization, in *Recent Developments in Well-Posed Variational Problems*, edited by R. Lucchetti and J. Revalski, pp. 95–115, Kluwer, Dordrecht, The Netherlands.
221. A. L. Dontchev (2012), Generalizations of the Dennis-More theorem, *SIAM J. Optim.* **22** (2012), 821–830.
222. A. L. Dontchev and H. Frankowska (2012), Lyusternik-Graves theorem and fixed points II, *J. Convex Anal.* **19**, 975–997.
223. A. L. Dontchev and W. W. Hager (1994), Implicit functions, Lipschitz maps and stability in optimization, *Math. Oper. Res.* **19**, 753–768.
224. A. L. Dontchev, A. S. Lewis and R. T. Rockafellar (2003), The radius of metric regularity, *Trans. Amer. Math. Soc.* **355**, 493–517.
225. A. L. Dontchev and R. T. Rockafellar (1996), Characterizations of strong regularity for variational inequalities over polyhedral convex sets, *SIAM J. Optim.* **6**, 1087–1105.
226. A. L. Dontchev and R. T. Rockafellar (1997), Characterizations of Lipschitzian stability in nonlinear programming, in *Mathematical Programming with Data Perturbations*, edited by A. V. Fiacco, pp. 65–82, Marcel Dekker, New York.
227. A. L. Dontchev and R. T. Rockafellar (2014), *Implicit Functions and Solution Mappings: A View from Variational Analysis*, 2nd edition, Springer, New York.
228. D. Drusvyatskiy and A. D. Ioffe (2015), Quadratic growth and critical point stability of semi-algebraic functions, *Math. Program.* **153**, 635–653.
229. D. Drusvyatskiy, A. D. Ioffe and A. S. Lewis (2015), Transversality and alternating projections for nonconvex sets, *Found. Comput. Math.* **15**, 1637–1651.

230. D. Drusvyatskiy and A. S. Lewis (2013), Tilt stability, uniform quadratic growth, and strong metric regularity of the subdifferential, *SIAM J. Optim.* **23**, 256–267.
231. D. Drusvyatskiy and A. S. Lewis (2016), Error bounds, quadratic growth, and linear convergence of proximal methods, http: arXiv:1602.06661.
232. D. Drusvyatskiy, B. S. Mordukhovich and T. T. A. Nghia (2014), Second-order growth, tilt stability, and metric regularity of the subdifferential, *J. Convex Anal.* **21**, 1165–1192.
233. D. Drusvyatskiy and C. Paquette (2018), Variational analysis of spectral functions simplified, to appear in *J. Convex Anal.* **25**, No. 1.
234. A. Y. Dubovitskii and A. A. Milyutin (1965), Extremum problems in the presence of restrictions, *USSR Comput. Maths. Math. Phys.* **5**, 1–80.
235. M. Dür (2003), A parametric characterization of local optimality, *Math. Methods Oper. Res.* **57**, 101–109.
236. M. Durea, J. Dutta and C. Tammer (2010), Lagrange multipliers and ε-Pareto solutions in vector optimization with nonsolid cones in Banach spaces, *J. Optim. Theory Appl.* **145**, 196–211.
237. M. Durea and R. Strugariu (2010), Necessary optimality conditions for weak sharp minima in set-valued optimization, *Nonlinear Anal.* **73**, 2148–2157.
238. M. Durea and R. Strugariu (2012), Openness stability and implicit multifunction theorems: applications to variational systems, *Nonlinear Anal.* **75**, 1246–1259.
239. M. Durea and R. Strugariu (2012), Chain rules for linear openness in general Banach spaces, *SIAM J. Optim.* **22**, 899–913.
240. M. Durea and R. Strugariu (2016), Metric subregularity of composition set-valued mappings with applications to fixed point theory, *Set-Valued Anal.* **24**, 231–251.
241. J. Dutta (2005), Optimality conditions for maximizing a locally Lipschitz function, *Optimization* **54**, 377–389.
242. J. Dutta (2012), Strong KKT, second order conditions and non-solid cones in vector optimization, in *Recent Developments in Vector Optimization*, edited by Q. H. Ansari and J.-C. Yao, pp. 127–167, Springer, Berlin.
243. A. C. Eberhard and R. Wenczel (2012), A study of tilt-stable optimality and sufficient conditions, *Nonlinear Anal.* **75**, 1240–1281.
244. M. Ehrgott (2005), *Multicriteria Optimization*, 2nd edition, Springer, Berlin.
245. G. Eichfelder (2014), *Variable Ordering Structures in Vector Optimization*, Springer, Berlin.
246. G. Eichfelder and T. X. D. Ha (2013), Optimality conditions for vector optimization problems with variable ordering structures, *Optimization* **62**, 1468–1476.

247. G. EICHFELDER AND R. KASIMBEYLI (2013), Properly optimal elements in vector optimization with variable ordering structures, *J. Global Optim.* **60**, 689–712.
248. I. EKELAND (1972), Sur les problémes variationnels, *C. R. Acad. Sci. Paris* **275**, 1057–1059.
249. I. EKELAND (1974), On the variational principle, *J. Math. Anal. Appl.* **47**, 324–353.
250. I. EKELAND (1979), Nonconvex minimization problems, *Bull. Amer. Math. Soc.* **1**, 432–467.
251. B. EL ABDOINI AND L. THIBAULT (1996), Optimality conditions for problems with set-valued objectives, *J. Appl. Anal.* **2**, 183–201.
252. K. EMICH AND R. HENRION (2014), A simple formula for the second-order subdifferential of maximum functions, *Vietnam J. Math.* **42**, 467–478.
253. E. ERNST AND M. THÉRA (2007), Boundary half-strips and the strong CHIP, *SIAM. J. Optim.* **18**, 834–852.
254. M. FABIAN (1989), Subdifferentiability and trustworthiness in the light of a new variational principle of Borwein and Preiss, *Acta Univ. Carolina, Ser. Math. Phys.* **30**, 51–56.
255. M. FABIAN ET AL. (2011), *Functional Analysis and Infinite-Dimensional Geometry*, 2nd edition, Springer, New York.
256. M. FABIAN, R. HENRION, A. Y. KRUGER AND J. V. OUTRATA (2010), Error bounds: necessary and sufficient conditions, *Set-Valued Var. Anal.* **18**, 121–149.
257. M. FABIAN AND B. S. MORDUKHOVICH (1998), Smooth variational principles and characterizations of Asplund spaces, *Set-Valued Anal.* **6**, 381–406.
258. M. FABIAN AND B. S. MORDUKHOVICH (2002), Separable reduction and extremal principles in variational analysis, *Nonlinear Anal.* **49**, 265–292.
259. M. FABIAN AND B. S. MORDUKHOVICH (2003), Sequential normal compactness versus topological normal compactness in variational analysis, *Nonlinear Anal.* **54**, 1057–1067.
260. F. FACCHINEI AND J.-S. PANG (2003), *Finite-Dimensional Variational Inequalities and Complementary Problems*, published in two volumes, Springer, New York.
261. D. H. FANG, C. LI AND K. F. NG (2009), Constraint qualifications for extended Farkas's lemmas and Lagrangian dualities in convex infinite programming, *SIAM J. Optim.* **20**, 1311–1332.
262. D. H. FANG, C. LI AND K. F. NG (2010), Constraint qualifications for optimality conditions and total Lagrange dualities in convex programming, *Nonlinear Analysis* **73**, 1143–1159.
263. H. FEDERER (1959), Curvature measures, *Trans. Amer. Math. Soc.* **93**, 418–491.
264. M. C. FERRIS (1988), *Weak Sharp Minima and Penalty Functions in Mathematical Programming*, Ph.D. dissertation, University of Cambridge, Cambridge, United Kingdom.

265. S. D. Flåm (2006), Upward slopes and inf-convolutions, *Math. Oper. Res.* **31**, 188–198.
266. S. D. Flåm, J.-B. Hiriart-Urruty and A. Jourani (2009), Feasibility in finite time, *J. Dyn. Contr. Syst.* **15**, 537–555.
267. M. L. Flegel and C. Kanzow (2005), On M-stationarity for mathematical programs with equilibrium constraints, *J. Math. Anal. Appl.* **310**, 286–302.
268. W. H. Fleming and H. M. Soner (1993), *Controlled Markov Processes and Viscosity Solutions*, Springer, New York.
269. E.-A. Florea (2016), Coderivative necessary optimality conditions for sharp and robust efficiencies in vector optimization with variable ordering structure, *Optimization* **65** (2016), 1417–1435.
270. M. Florenzano, P. Gourdel and A. Jofré (2006), Supporting weakly Pareto optimal allocations in infinite-dimensional nonconvex economies, *J. Economic Theory* **29**, 549–564.
271. F. Flores-Bazán, A. Jourani and G. Mastroeni (2014), On the convexity of the value function for a class of nonconvex variational problems: existence and optimality conditions, *SIAM J. Control Optim.* **52**, 3673–3693.
272. F. Flores-Bazán and E. Hernández (2013), Optimality conditions for a unified vector optimization problem with not necessarily preordering relations, *J. Global Optim.* **56**, 299–315.
273. F. Flores-Bazán and G. Mastroeni (2015), Characterizing FJ and KKT conditions in nonconvex mathematical programming with applications, *SIAM J. Optim.* **25** (2015), 647–676.
274. H. Frankowska and M. Quincampoix (2012), Hölder metric regularity of set-valued maps, *Math. Program.* **132**, 333–354.
275. M. Fukushima and J.-S. Pang (2005), Quasi-variational inequalities, generalized Nash equilibria, and multi-leader-follower games, *Comput. Management Sci.* **1**, 21–56.
276. M. Gaydu, M. H. Geoffroy and C. Jean-Alexis (2011), Metric subregularity of order q and the solving of inclusions, *Cent. European J. Math.* **9**, 147–161.
277. W. Geremew, B. S. Mordukhovich and N. M. Nam (2009), Coderivative calculus and metric regularity for constraint and variational systems, *Nonlinear Anal.* **70**, 529–552.
278. C. Gerstewitz (Tammer) (1983), Ninchtkonvexe dualität in der vectoroptimierung, *Wissenschaftliche Zeitschrift der TH Leuna-Merseburg* **25**, 357–364 (in German).
279. C. Gerth (Tammer) and P. Weidner (1990), Nonconvex separation theorems and some applications in vector optimization, *J. Optim. Theory Appl.* **67**, 297–320.
280. H. Gfrerer (2011), First order and second order characterizations of metric subregularity and calmness of constraint set mappings, *SIAM J. Optim.* **21**, 1439–1474.

281. H. GFRERER (2013), On directional metric subregularity and second-order optimality conditions for a class of nonsmooth mathematical programs, *SIAM J. Optim.* **23**, 632–665.
282. H. GFRERER (2013), On directional metric regularity, subregularity and optimality conditions for nonsmooth mathematical programs, *Set-Valued Var. Anal.* **21**, 151–176.
283. H. GFRERER AND D. KLATTE (2016), Lipschitz and Hölder stability of optimization problems and generalized equations Lipschitz and Hölder stability of optimization problems and generalized equations, *Math. Program.* **158**, 35–75.
284. H. GFRERER AND B. S. MORDUKHOVICH (2015), Complete characterizations of tilt stability in nonlinear programming under weakest qualification conditions, *SIAM J. Optim.* **25**. 2081–2119.
285. H. GFRERER AND B. S. MORDUKHOVICH (2017), Robinson regularity of parametric constraint systems via variational analysis, *SIAM J. Optim.* **27**, 438–465.
286. H. GFRERER AND J. V. OUTRATA (2016), On computation of generalized derivatives of the normal-cone mapping and their applications, *Math. Oper. Res.* **41**, 1535–1556.
287. H. GFRERER AND J. V. OUTRATA (2016), On Lipschitzian properties of implicit multifunctions, *SIAM J. Optim.* **26**, 2160–2189.
288. H. GFRERER AND J. V. OUTRATA (2016), On computation of limiting coderivatives of the normal-cone mapping to inequality systems and their applications, *Optimization* **65**, 671–700.
289. H. GFRERER AND J. V. OUTRATA (2017), On the Aubin property of a class of parameterized variational systems, *Math. Meth. Oper. Res.*, DOI 10.1007/s00186-017-0596-y.
290. H. GFRERER AND J. J. YE (2017), New constraint qualifications for mathematical programs with equilibrium constraints via variational analysis, *SIAM J. Optim.* **27**, 842–865.
291. F. GIANNESSI (1980), Theorems of alternative, quadratic programs and complementarity problems, in *Variational Inequalities and Complementarity Problems*, edited by R. Cottle et al., pp. 151–186, Wiley, New York.
292. F. GIANNESSI (2005), *Constrained Optimization and Image Space Analysis*, Springer, London, United Kingdom.
293. I. GINCHEV AND B. S. MORDUKHOVICH (2011), On directionally dependent subdifferentials, *C. R. Acad. Bulg. Sci.* **64**, 497–508.
294. I. GINCHEV AND B. S. MORDUKHOVICH (2012), Directional subdifferentials and optimality conditions, *Positivity* **16**, 707–737.
295. E. GINER (2017), Clarke and limiting subdifferentials of integral functionals, *J. Convex Anal.* **24**, No. 3.
296. I. V. GIRSANOV (1972), *Lectures on Mathematical Theory of Extremum Problems*, Springer, Berlin.

297. M. A. GOBERNA, V. JEYAKUMAR AND M. A. LÓPEZ (2008), Necessary and sufficient constraint qualifications for solvability of systems of infinite convex inequalities, *Nonlinear Anal.* **68**, 1184–1194.
298. M. A. GOBERNA AND M. A. LÓPEZ (1998), *Linear Semi-Infinite Optimization*, Wiley, Chichester, United Kingdom.
299. M. A. GOBERNA AND M. A. LÓPEZ (2014), *Post-Optimal Analysis in Linear Semi-Infinite Optimization*, Springer, New York.
300. A. GÖPFERT, H. RIAHI, C. TAMMER AND C. ZĂLINESCU (2003), *Variational Methods in Partially Ordered Spaces*, Springer, New York.
301. V. V. GOROKHOVIK (1990), *Convex and Nonsmooth Problems of Vector Optimization*, Nauka i Tekhnika, Minsk, Belarus.
302. V. V. GOROKHOVIK AND M. TRAFIMOVICH (2016), Positively homogeneous functions revisited, *J. Optim. Theory Appl.* **171**, 481–503.
303. S.-M. GRAD (2015), *Vector Optimization and Monotone Operators via Convex Duality*, Springer, Berlin.
304. L. M. GRAÑA DRUMMOND AND B. F. SVAITER (2005), A steepest descent method for vector optimization, *J. Comput. Appl. Math.* **175** (2005), 395–414.
305. L. M. GRAVES (1950), Some mapping theorems, *Duke Math. J.* **17**, 111–114.
306. A. GREENBAUM, A. S. LEWIS AND M. L. OVERTON (2017), Variational analysis of the Crouzeix ratio, *Math. Program.* **164**, 229–243.
307. A. GRIEWANK (2013), On stable piecewise linearization and generalized algorithmic differentiation, *Optim. Methods Softw.* **28** (2013), 1139–1178.
308. A. GRIEWANK AND A. WALTHER (2008), *Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation*, 2nd edition, SIAM, Philadelphia, Pennsylvania.
309. A. GRIEWANK AND A. WALTHER (2016), First- and second-order optimality conditions for piecewise smooth objective functions, *Optim. Methods Softw.* **31**, 904–930.
310. S. GRUNDEL AND M. L. OVERTON (2014), Variational analysis of the spectral abscissa at a matrix with a nongeneric multiple eigenvalue, *Set-Valued Var. Anal.* **22**, 19–43.
311. J. GRZYZBOWSKI, D. PALLASCHKE AND R. URBANSKI (2017), Characterization of differences of sublinear functions, to appear in *Pure Appl. Funct. Anal.*
312. A. GUERRAGGIO AND D. T. LUC (2006), Properly maximal points in product spaces, *Math. Oper. Res.* **31**, 305–315.
313. R. GUESNERIE (1975), Pareto optimality in non-convex economies, *Econometrica* **43**, 1–29.
314. L. GUO AND J. J. YE (2016), Necessary optimality conditions for optimal control problems with equilibrium constraints, *SIAM J. Control Optim.* **54**, 2710–2733.
315. M. GÜRBÜZBALABAN AND M. L. OVERTON (2012), On Nesterov's nonsmooth Chebyshev-Rosenbrock functions, *Nonlinear Anal.* **75**, 1282–1289.

316. R. GUPTA, F. JAFARI, R. J. KIPKA AND B. S. MORDUKHOVICH (2017), Linear openness and feedback stabilization of nonlinear control systems, to appear in *Dyn. Cont. Control. Syst. Ser. S*, arXiv:1704.00867.
317. C. GUTIÉRREZ, L. HUERGA, B. JIMÉNEZ AND V. NOVO (2013), Proper approximate solutions and ε-subdifferentials in vector optimization: basic properties and limit behaviour, *Nonlinear Anal.* **79**, 52–67.
318. T. X. D. HA (2005), Lagrange multipliers for set-valued problems associated with coderivatives, *J. Math. Anal. Appl.* **311**, 647–663.
319. T. X. D. HA (2005), Some variants of the Ekeland variational principle for a set-valued map, *J. Optim. Theory Appl.* **124**, 187–206.
320. T. X. D. HA (2012), Optimality conditions for various efficient solutions involving coderivatives: from set-valued optimization problems to set-valued equilibrium problems, *Nonlinear Anal.* **75**, 1305–1323.
321. T. X. D. HA (2012), The Fermat rule and Lagrange multiplier rule for various effective solutions to set-valued optimization problems expressed in terms of coderivatives, in *Recent Developments in Vector Optimization*, edited by Q. H. Ansari and J.-C. Yao, pp. 417–466, Springer, Berlin.
322. A. HABTE AND B. S. MORDUKHOVICH (2011), Extended second welfare theorem for nonconvex economies with infinite commodities and public goods, *Adv. Math. Econ.* **14**, 93–126.
323. N. HADJISAVVAS (2005), Generalized convexity, generalized monotonicity and nonsmooth analysis, *Handbook of Generalized Convexity and Generalized Monotonicity*, edited by N. Hadjisavvas et al., pp. 465–499, Springer, New York.
324. A. H. HAMEL, F. HEYDE, A. LÖHNE, B. RUDLOFF AND C. SCHRAGE (2015), Set optimization–a rather short introduction, in *Set Optimization and Applications–the State of the Art*, edited by A. H. Hamel et al., pp. 65–141, Springer, Berlin.
325. A. H. HAMEL AND C. SCHRAGE (2014), Directional derivatives, subdifferentials and optimality conditions for set-valued convex functions, *Pac. J. Optim.* **10**, 667–689.
326. N. T. V. HANG (2014), The penalty functions method and multiplier rules based on the Mordukhovich subdifferential, *Set-Valued Var. Anal.* **22**, 299–312.
327. N. T. V. HANG AND J.-C. YAO (2016), Sufficient conditions for error bounds of difference functions and applications, *J. Global Optim.* **66**, 439–456.
328. N. T. V. HANG AND N. D. YEN (2015), Optimality conditions and stability analysis via the Mordukhovich subdifferential, *Numer. Funct. Anal. Optim.* **36**, 364–386.
329. N. T. V. HANG AND N. D. YEN (2016), On the problem of minimizing a difference of polyhedral convex functions under linear constraints, *J. Optim. Theory Appl.* **171**, 617–642.
330. A. HANTOUTE, R. HENRION AND P. PÉREZ-AROS (2017), Subdifferential characterization of continuous probability functions under Gaussian distribution, to appear in *Math. Program.*, arXiv:1705.10160.

331. A. Hantoute, M. A. López and C. Zălinescu (2008), Subdifferential calculus rules in convex analysis: a unifying approach via pointwise supremum functions, *SIAM J. Optim.* **19**, 863–882.
332. W. L. Hare and C. Planiden (2014), Parametrically prox-regular functions, *J. Convex Anal.* **21**, 901–923.
333. W. Hare and C. Sagastizábal (2005), Computing proximal points of nonconvex functions, *Math. Program.* **116** (2009), 221–258.
334. Y. He and K. F. Ng (2006), Subdifferentials of a minimal time function in Banach spaces, *J. Math. Anal. Appl.* **321**, 896–910.
335. R. Henrion, A. Jourani and J. V. Outrata (2002), On the calmness of a class of multifunctions, *SIAM J. Optim.* **13**, 603–618.
336. R. Henrion, B. S. Mordukhovich and N. M. Nam (2010), Second-order analysis of polyhedron systems in finite and infinite dimensions with applications to robust stability of variational inequalities, *SIAM J. Optim.* **20**, 2199–2227.
337. R. Henrion and J. V. Outrata (2001), A subdifferential condition for calmness of multifunctions, *J. Math. Anal. Appl.* **258**, 110–130.
338. R. Henrion and J. V. Outrata (2005), Calmness of constraint systems with applications, *Math. Program.* **104**, 437–464.
339. R. Henrion, J. V. Outrata and T. Surowiec (2009), On the coderivative of normal cone mappings to inequality systems, *Nonlinear Anal.* **71**, 1213–1226.
340. R. Henrion, J. V. Outrata and T. Surowiec (2012), Analysis of M-stationary points to an EPEC modeling oligopolistic competition in an electricity spot market, *ESAIM Control Optim. Calc. Var.* **18**, 295–317.
341. R. Henrion and T. Surowiec (2011), On calmness conditions in convex bilevel programming, *Applic. Anal.* **90**, 951–970.
342. R. Henrion and W. Römisch (2007), On M-stationary points for a stochastic equilibrium problem under equilibrium constraints in electricity spot market modeling, *Appl. Math.* **52** (2007), 473–494.
343. E. Hernández, L. Rodríguez-Marín and M. Sama (2010), On solutions of set-valued optimization problems, *Comput. Math. Appl.* **65**, 1401–1408.
344. R. Hesse and D. R. Luke (2013), Nonconvex notions of regularity and convergence of fundamental algorithms for feasibility problems, *SIAM J. Optim.* **23**, 2397–2419.
345. R. Hettich and K. O. Kortanek (1993), Semi-infinite programming: theory, methods, and applications, *SIAM Rev.* **35**, 380–429.
346. M. Hintermüller, B. S. Mordukhovich and T. Surowiec (2014), Several approaches for the derivation of stationarity conditions for elliptic MPECs with upper-level control constraints, *Math. Program.* **146**, 555–582.
347. M. Hintermüller, and T. Surowiec (2011), First-order optimality conditions for elliptic mathematical programs with equilibrium constraints via variational analysis, *SIAM J. Optim.* **21**, 1561–1593.

348. J.-B. HIRIART-URRUTY (1982), ε-Subdifferential, in *Convex Analysis and Optimization*, edited by J.-P. Aubin and R. Vinter, pp. 43–92, Pitman, London, United Kingdom.
349. J.-B. HIRIART-URRUTY (1983), A short proof of the variational principle for approximate solutions of a minimization problem, *Amer. Math. Monthly* **90**, 206–207.
350. J.-B. HIRIART-URRUTY (1989), From convex optimization to nonconvex optimization: necessary and sufficient conditions for global optimality, in *Nonconvex Optimization and Related Topics*, pp. 219–239, Plenum Press, New York.
351. J.-B. HIRIART-URRUTY AND Y. S. LEDYAEV (1996), A note on the characterization of the global maxima of a (tangentially) convex function over a convex set, *J. Convex Anal.* **3**, 55–61.
352. J.-B. HIRIART-URRUTY AND C. LEMARÉCHAL (1993), *Convex Analysis and Minimization Algorithms*, published in two volumes, Springer, Berlin.
353. A. J. HOFFMAN (1952), On approximate solutions of systems of linear inequalities, *J. Res. Nat. Bureau Stand.* **e49**, 263–265.
354. T. HOHEISEL, C. KANZOW, B. S. MORDUKHOVICH AND H. M. PHAN (2012), Generalized Newton's methods for nonsmooth equations based on graphical derivatives, **75**, 1324–1340; Erratum in *Nonlinear Anal.* **86** (2013), 157–158.
355. R. HORST, P. D. PARDALOS AND N. V. THOAI (2000), *Introduction to Global Optimization*, 2nd edition, Springer, New York.
356. X. HU AND D. RALPH (2007), Using EPECs to model bilevel games in restructured electricity markets with locational prices, *Oper. Res.* **55**, 809–827.
357. H. HUANG (2008), The Lagrange multiplier rule for super efficiency in vector optimization, *J. Math. Anal. Appl.* **342**, 503–513.
358. N. Q. HUY AND N. V. TUYEN (2016), New second-order optimality conditions for a class of differentiable optimization problems, *J. Optim. Theory Appl.* **171**, 27–44.
359. N. Q. HUY, D. S. KIM AND N. V. TUYEN (2017), New second-order Karush-Kuhn-Tucker optimality conditions for vector optimization, *Appl. Math. Optim.*, DOI 10.1007/s00245-017-9432-2.
360. N. Q. HUY, B. S. MORDUKHOVICH AND J.-C. YAO (2008), Coderivatives of frontier and solution maps in parametric multiobjective optimization, *Taiwanese J. Math.* **12**, 2083–2111.
361. N. Q. HUY AND J.-C. YAO (2013), Exact formulae for coderivatives of normal cone mappings to perturbed polyhedral convex sets, *J. Optim. Theory Appl.* **157**, 25–43.
362. J. IDE, E. KÖBIS, D. KUROIWA, A. SCHÖBEL AND C. TAMMER (2014), The relationships between multicriteria robustness concepts and set-valued optimization, *Fixed Point Theory Appl.* **83**.
363. A. D. IOFFE (1979), Regular points of Lipschitz functions, *Trans. Amer. Math. Soc.* **251**, 61–69.

364. A. D. IOFFE (1981), Calculus of Dini subdifferentials, CEREMADE Publication 8110, Universiteé de Paris IX “Dauphine”.
365. A. D. IOFFE (1981), Approximate subdifferentials of nonconvex functions, CEREMADE Publication 8120, Universiteé de Paris IX “Dauphine.”
366. A. D. IOFFE (1981), Nonsmooth analysis: differential calculus of nondifferentiable mappings, *Trans. Amer. Math. Soc.* **266** (1981), 1–56.
367. A. D. IOFFE (1983), On subdifferentiability spaces, *Ann. New York Acad. Sci.* **410**, 107–119.
368. A. D. IOFFE (1984), Approximate subdifferentials and applications, I: the finite dimensional theory, *Trans. Amer. Math. Soc.* **281**, 389–415.
369. A. D. IOFFE (1989), Approximate subdifferentials and applications, III: the metric theory, *Mathematika* **36**, 1–38.
370. A. D. IOFFE (1990), Proximal analysis and approximate subdifferentials, *J. London Math. Soc.* **41**, 175–192.
371. A. D. IOFFE (2000), Codirectional compactness, metric regularity and subdifferential calculus, in *Constructive, Experimental and Nonlinear Analysis*, edited by M. Théra, pp. 123–164, American Mathematical Society, Providence, Rhode Island.
372. A. D. IOFFE (2010), On regularity concepts in variational analysis, *J. Fixed Point Theory Appl.* **8**, 339–363.
373. A. D. IOFFE (2012), On stability of solutions to convex inequalities, *J. Convex Anal.* **19**, 1017–1032.
374. A. D. IOFFE (2013), Convexity and variational analysis, in *Computational and Analytical Mathematics*, edited by D. Bailey et al., pp. 397–428, Springer, New York.
375. A. D. IOFFE (2017), *Variational Analysis of Regular Mappings: Theory and Applications* (2017), Springer, Cham, Switzerland.
376. A. D. IOFFE AND J. V. OUTRATA (2008), On metric and calmness qualification conditions in subdifferential calculus, *Set-Valued Anal.* **16**, 199–228.
377. A. D. IOFFE AND Y. SEKIGUCHI (2009), Regularity estimates for convex multifunctions, *Math. Program.* **117**, 255–270.
378. A. D. IOFFE AND V. M. TIKHOMIROV (1973), *Theory of Extremal Problems*, Nauka, Moscow (in Russian); English translation in North-Holland, Amsterdam, 1979.
379. A. N. IUSEM AND V. SOSA (2010), On the proximal point method for equilibrium problems in Hilbert spaces, *Optimization* **59**, 1259–1274.
380. G. E. IVANOV AND L. THIBAULT (2017), Infimal convolution and optimal time control problem: Fréchet and proximal subdifferentials, *Set-Valued Var. Anal.*, DOI 10.1007/s11228-016-0398-z.
381. G. E. IVANOV AND L. THIBAULT (2017), Infimal convolution and optimal time control problem: limiting subdifferential, *Set-Valued Var. Anal.* **25**, 517–542.
382. A. F. IZMAILOV (2014), Strongly regular nonsmooth generalized equations, *Math. Program.* **147**, 581–590.

383. A. F. IZMAILOV AND M. V. SOLODOV (2012), Stabilized SQP revisited, *Math. Program.* **133**, 93–120.
384. A. F. IZMAILOV AND M. V. SOLODOV (2014), *Newton-Type Methods for Optimization and Variational Problems*, Springer, New York.
385. J. JAHN (2004), *Vector Optimization: Theory, Applications and Extensions*, Springer, Berlin.
386. J. JAHN AND A. A. KHAN (2003), Some calculus rules for contingent epiderivatives, *Optimization* **52**, 113–125.
387. J. JEYAKUMAR (1997), Asymptotic dual conditions characterizing optimality for convex programs, *J. Optim. Theory Appl.* **93**, 153–165.
388. V. JEYAKUMAR, G. M. LEE AND N. DINH (2003), New sequential Lagrange multiplier conditions characterizing optimality without constraint qualification for convex programs, *SIAM J. Optim.* **14**, 534–547.
389. V. JEYAKUMAR AND D. T. LUC (2008), *Nonsmooth Vector Functions and Continuous Optimization*, Springer, New York.
390. Y. JIANG, Y.-J. LIU AND L. ZHANG (2015), Variational geometry of the complementarity set for second-order cone, *Set-Valued Var. Anal.* **23**, 399–414.
391. A. JOFRÉ (2000), A second welfare theorem in nonconvex economies, in *Constructive, Experimental, and Nonlinear Analysis*, edited by M. Théra, pp. 175–184, American Mathematical Society, Providence, Rhode Island.
392. A. JOFRÉ AND A. JOURANI (2015), Characterizations of the free disposal condition for nonconvex economies on infinite dimensional commodity spaces, *SIAM J. Optim.* **25** (2015), 699–712.
393. A. JOFRÉ AND J. RIVERA CAYUPI (2006), A nonconvex separation property and some applications, *Math. Program.* **108** (2006), 37–51.
394. H. TH. JONGEN, F. TWILT AND G. M. WEBER (1992), Semi-infinite optimization: structure and stability of the feasible set, *J. Optim. Theory Appl.* **72**, 529–552.
395. H. TH. JONGEN AND J.-J. RÜCKMANN (1998), On stability and deformation in semi-infinite optimization, in *Semi-Infinite Programming*, edited by R. Reemtsen and J.-J. Rückmann, pp. 29–67, Kluwer, Boston, Massachusetts.
396. H. TH. JONGEN, J.-J. RÜCKMANN AND O. STEIN (1998), Generalized semi-infinite optimization: a fist order optimality condition and examples, *Math. Program.*, **83**, 145–158.
397. A. JOURANI AND L. THIBAULT (1995), Verifiable conditions for openness, metric regularity of multivalued mappings, *Trans. Amer. Math. Soc.* **347**, 1225–1268.
398. A. JOURANI AND L. THIBAULT (1996), Metric regularity and subdifferential calculus in Banach spaces, *Set-Valued Anal.* **3**, 87–100.
399. A. JOURANI AND L. THIBAULT (1996), Extensions of subdifferential calculus rules in Banach spaces, *Canad. J. Math.* **48**, 834–848.
400. A. JOURANI AND L. THIBAULT (1998), Chain rules for coderivatives of multivalued mappings in Banach spaces, *Proc. Amer. Math. Soc.* **126**, 1479–1485.

401. A. JOURANI AND L. THIBAULT (1998), Qualification conditions for calculus rules of coderivatives of multivalued mappings, *J. Math. Anal. Appl.* **218**, 66–81.
402. A. JOURANI AND L. THIBAULT (1999), Coderivatives of multivalued mappings, locally compact cones and metric regularity, *Nonlinear Anal.* **35**, 925–945.
403. A. JOURANI AND L. THIBAULT (2011), Noncoincidence of approximate and limiting subdifferentials of integral functionals, *SIAM J. Control Optim.* **49**, 1435–1453.
404. A. JOURANI, L. THIBAULT AND D. ZAGRODNY (2012), $C^{1,\omega(\cdot)}$-regularity and Lipschitz-like properties of subdifferential, *Proc. London Math. Soc.* **105**, 189–223.
405. N. KANZI AND S. NOBAKHTIAN (2010), Optimality conditions for nonsmooth semi-infinite programming, *Optimization* **59**, 717–727.
406. N. KANZI AND S. NOBAKHTIAN (2010), Necessary optimality conditions for nonsmooth generalized semi-infinite programming problems, *European J. Oper. Res.* **205**, 253–261.
407. R. KASIMBEYLI (2010), A nonlinear cone separation theorem and scalarization in nonconvex vector optimization, *SIAM J. Optim.* **20**, 1591–1619.
408. P. KENDEROV (1975), Semi-continuity of set-valued monotone mappings, *Fundamenta Mathematicae* **88**, 61–69.
409. A. A. KHAN, C. TAMMER AND C. ZĂLINESCU (2015), *Set-Valued Optimization. An Introduction with Applications*, Springer, Berlin.
410. K. A. KHAN AND P. I. BARTON (2015), A vector forward mode of automatic differentiation for generalized derivative evaluation, *Optim. Methods Softw.* **30**, 1185–1212.
411. M. A. KHAN (1991), Ioffe's normal cone and the foundation of welfare economics: the infinite-dimensional theory, *J. Math. Anal. Appl.* **161**, 284–298.
412. M. A. KHAN (1999), The Mordukhovich normal cone and the foundations of welfare economics, *J. Public Econ. Theory* **1**, 309–338.
413. P. D. KHANH, J.-C. YAO AND N. D. YEN (2017), The Mordukhovich subdifferentials and directions of descent, *J. Optim. Theory Appl.* **172**, 518–534.
414. P. Q. KHANH, A. KRUGER AND N. H. THAO (2015), An induction theorem and nonlinear regularity models, *SIAM J. Optim.* **25**, 2561–2588.
415. P. Q. KHANH AND D. N. QUY (2011), On generalized Ekeland's variational principle and equivalent formulations for set-valued mappings, *J. Global Optim.* **49** (2011), 381–396.
416. B. T. KIEN, Y. C. LIOU, N.-C. WONG AND J.-C. YAO (2009), Subgradients of value functions in parametric dynamic programming, *European J. Oper. Res.* **193**, 12–22.
417. A. J. KING AND R. T. ROCKAFELLAR (2002), Sensitivity analysis for nonsmooth generalized equations, *Math. Program.* **55**, 193–212.
418. D. KLATTE AND R. HENRION (1998), Regularity and stability in nonlinear semi-infinite optimization, in *Semi-Infinite Programming*, edited by R. Reemtsen and J.-J. Rückmann, pp. 69–102, Kluwer, Boston, Massachusetts.

419. D. KLATTE, A. Y. KRUGER AND B. KUMMER (2012), From convergence principles to stability and optimality conditions, *J. Convex Anal.* **19**, 1043–1072.
420. D. KLATTE AND B. KUMMER (2002), *Nonsmooth Equations in Optimization: Regularity, Calculus, Methods, and Applications*, Kluwer, Boston, Massachusetts.
421. D. KLATTE AND B. KUMMER (2009), Optimization methods and stability of inclusions in Banach spaces, *Math. Program.* **117**, 305–330.
422. M. KNOSSALLA (2018), Minimization of marginal functions in mathematical programming based on continuous outer subdifferentials, *Optimization*, DOI 10.1080/02331934.2018.1426579.
423. M. KOČVARA, M. KRUŽIK AND J. V. OUTRATA (2005), On the control of an evolutionary equilibrium in micromagnetics, in *Optimization with Multivalued Mappings: Theory, Applications and Algorithms*, edited by S. Dempe and V. V. Kalashnikov, pp. 143–168, Springer, Berlin.
424. I. V. KONNOV (2001), *Combined Relaxation Methods for Variational Inequalities*, Springer, Berlin.
425. N. N. KRASOVSKII AND A. I. SUBBOTIN (1988), *Game-Theoretical Control Problems*, Springer, New York.
426. A. Y. KRUGER (1981), *Generalized Differentials of Nonsmooth Functions and Necessary Conditions for an Extremum*, Ph.D. dissertation, Department of Applied Mathematics, Belarus State University, Minsk, Belarus (in Russian).
427. A. Y. KRUGER (1981), Epsilon-semidifferentials and epsilon-normal elements, Depon. VINITI #1331-81, Moscow (in Russian).
428. A. Y. KRUGER (1981), Generalized differentials of nonsmooth functions, Depon. VINITI #1332-81, Moscow (in Russian).
429. A. Y. KRUGER (1982), On characterizing the covering property for nonsmooth operators, *Abstracts of the School on the Operator Theory in Functional Spaces*, pp. 94–95, Minsk, Belarus (in Russian).
430. A. Y. KRUGER (1985), Generalized differentials of nonsmooth functions and necessary conditions for an extremum, *Siberian Math. J.* **26**, 370–379.
431. A. Y. KRUGER (1985), Properties of generalized differentials, *Siberian Math. J.* **26**, 822–832.
432. A. Y. KRUGER (1988), A covering theorem for set-valued mappings, *Optimization* **19**, 763–780.
433. A. Y. KRUGER (2009), About stationarity and regularity in variational analysis, *Taiwanese J. Math.* **13**, 1737–1785.
434. A. Y. KRUGER (2015), Error bounds and metric subregularity, *Optimization* **64**, 49–79.
435. A. Y. KRUGER (2015), Error bounds and Hölder metric subregularity, *Set-Valued Var. Anal.* **23**, 705–736.
436. A. Y. KRUGER AND M. A. LÓPEZ (2012), Stationarity and regularity of infinite collections of sets, *J. Optim. Theory Appl.* **154**, 339–369.

437. A. Y. KRUGER AND M. A. LÓPEZ (2012), Stationarity and regularity of infinite collections of sets. Applications to infinitely constrained optimization, *J. Optim. Theory Appl.* **155**, 390–416.
438. A. Y. KRUGER, D. R. LUKE AND N. M. THAO (2017), Set regularities and feasibility problems, *Math. Program.*, DOI 10.1007/s10107-016-1039-x.
439. A. Y. KRUGER AND B. S. MORDUKHOVICH (1978), Minimization of nonsmooth functionals in optimal control problems, *Eng. Cybernetics* **16**, 126–133.
440. A. Y. KRUGER AND B. S. MORDUKHOVICH (1980), Generalized normals and derivatives, and necessary optimality conditions in nondifferential programming, I&II, Depon. VINITI: I# 408-80, II# 494-80, Moscow (in Russian).
441. A. Y. KRUGER AND B. S. MORDUKHOVICH (1980), Extremal points and the Euler equation in nonsmooth optimization, *Dokl. Akad. Nauk BSSR* **24**, 684–687 (in Russian).
442. A. KRULIKOWSKI (1997), *Fundamentals of Geometric Dimensioning and Tolerancing*, Cengage Learning, Stamford, Connecticut.
443. B. KUMMER (1991), Lipschitzian inverse functions, directional derivatives and applications in $C^{1,1}$ optimization, *J. Optim. Theory Appl.* **70**, 561–582.
444. B. KUMMER (1999), Metric regularity: characterizations, nonsmooth variations and successive approximation, *Optimization* **46**, 247–281.
445. B. KUMMER (2000), Inverse functions of pseudo regular mappings and regularity conditions, *Math. Program.* **88**, 313–339.
446. D. KUROIWA (1998), The natural criteria in set-valued optimization, *RIMS Kokyuroku* **1031**, 85–90.
447. A. G. KUSRAEV AND S. S. KUTATELADZE (1995), *Subdifferentials: Theory and Applications*, Kluwer, Dordrecht, The Netherlands.
448. S. LAHRECH AND A. BENBRIK (2005), On the Mordukhovich subdifferential in binormal spaces and some applications, *Int. J. Pure Appl. Math* **20**, 31–39.
449. G. LEBOURG (1975), Valeur moyenne pour gradient généraliseé, *C. R. Acad. Sci. Paris* **281**, 795–798.
450. Y. S. LEDYAEV AND Q. J. ZHU (1999), Implicit multifunction theorem, *Set-Valued Anal.* **7**, 209–238.
451. G. M. LEE AND N. D. YEN (2012), Coderivatives of a Karush-Kuhn-Tucker point set map and applications, *Nonlinear Analysis* **95**, 191–201.
452. C. LEMARÉCHAL, F. OUSTRY AND C. SAGASTIZÁBAL (2000), The $\mathcal{U}$-Lagrangian of a convex function, *Trans. Amer. Math. Soc.* **352**, 711–729.
453. A. B. LEVY (1996), Implicit multifunction theorems for the sensitivity analysis of variational conditions, *Math. Program.* **74**, 333–350.
454. A. B. LEVY AND B. S. MORDUKHOVICH (2004), Coderivatives in parametric optimization, *Math. Program.* **99**, 311–327.
455. A. B. LEVY AND R. A. POLIQUIN (1997), Characterizing the single-valuedness of multifunctions, *Set-Valued Anal.* **5**, 351–364.
456. A. B. LEVY, R. A. POLIQUIN AND R. T. ROCKAFELLAR (2000), Stability of locally optimal solutions, *SIAM J. Optim.* **10**, 580–604.

457. A. S. LEWIS (2003), The mathematics of eigenvalue optimization, *Math. Program.* **97**, 155–176.
458. A. S. LEWIS, D. R. LUKE, AND J. MALICK (2009), Local linear convergence for alternating and averaged nonconvex projections, *Found. Comput. Math.* **9**, 485–513.
459. A. S. LEWIS AND J. MALICK (2008), Alternating projection on manifolds, *Math. Oper. Res.* **33**, 216–234.
460. A. S. LEWIS AND S. J. WRIGHT (2011), Identifying activity, *SIAM J. Optim.* **21**, 597–614.
461. A. S. LEWIS AND S. ZHANG (2013), Partial smoothness, tilt stability, and generalized Hessians, *SIAM J. Optim.* **23**, 74–94.
462. C. LI, B. S. MORDUKHOVICH, J. WANG AND J.-C. YAO (2011), Weak sharp minima on Riamennian manifolds, *SIAM J. Optim.* **21**, 1523–1560.
463. C. LI, K. F. NG AND T. K. PONG (2007), The SECQ, linear regularity, and the strong CHIP for an infinite system of closed convex sets in normed linear spaces, *SIAM. J. Optim.* **18**, 643–665.
464. C. LI, K. F. NG AND T. K. PONG (2008), Constraint qualifications for convex inequality systems with applications to constrained optimization, *SIAM J. Optim.* **19**, 163–187.
465. G. LI AND B. S. MORDUKHOVICH (2012), Hölder metric subregularity with applications to proximal point method, *SIAM J. Optim.* **22**, 1655–1684.
466. G. LI, B. S. MORDUKHOVICH, T. T. A. NGHIA AND T. S. PHAM (2017), Error bounds for parametric polynomial systems with applications to higher-order stability analysis and convergence rates, *Math. Program.*, DOI 10.1007/s10107-016-1014-6.
467. G. LI, B. S. MORDUKHOVICH AND T. S. PHAM (2015), New fractional error bounds for polynomial systems with applications to Höderian stability in optimization and spectral theory of tensors, *Math. Program.* **153**, 333–362.
468. G. LI, K. F. NG AND X. Y. ZHENG (2007), Unified approach to some geometric results in variational analysis, *J. Funct. Anal.* **248** (2007), 317–343.
469. M. B. LIGNOLA AND J. MORGAN (2017), Inner regularizations and viscosity solutions for pessimistic bilevel optimization problems, *J. Optim. Theory Appl.* **173**, 183–202.
470. P. D. LOEWEN (1988), The proximal subgradient formula in Banach spaces, *Canad. Math. Bull.* **31**, 353–361.
471. P. D. LOEWEN (1993), *Optimal Control via Nonsmooth Analysis*, American Mathematical Society, Providence, Rhode Island.
472. P. D. LOEWEN (1994), A mean value theorem for Fréchet subgradients, *Nonlinear Anal.* **23**, 1365–1381.
473. M. S. ŁOJASIEWICZ (1959), Sur la probléme de la division, *Studia Math.* **18**, 87–136.
474. A. LÖHNE (2011), *Vector Optimization with Infimum and Supremum*, Springer, Berlin.
475. M. A. LÓPEZ AND G. STILL (2007), Semi-infinite programming, *European J. Oper. Res.* **180**, 491–518.

476. M. A. LÓPEZ AND M. VOLLE (2010), A formula for the set of optimal solutions of a relaxed minimization problem. Applications to subdifferential calculus, *J. Convex Anal.* **17**, 1057–1075.
477. S. LU (2011), Implications of the constant rank constraint qualification, *Math. Program.* **126**, 365–392.
478. D. T. LUC (1989), *Theory of Vector Optimization*, Springer, Berlin.
479. R. LUCCHETTI (2006), *Convexity and Well-Posed Problems*, Springer, New York.
480. D. R. LUKE (2012), Local linear convergence of approximate projections onto regularized sets, *Nonlinear Anal.* **75**, 1531–1546.
481. D. R. LUKE, N. H. THAO AND M. K. TAM (2017), Quantitative convergence analysis of iterated expansive, set-valued mappings, to appear in *Math. Oper. Res.*, arXiv:1605.05725.
482. Z. Q. LUO, J.-S. PANG AND D. RALPH (1996), *Mathematical Programs with Equilibrium Constraints*, Cambridge University Press, Cambridge, United Kingdom.
483. L. A. LYUSTERNIK (1934), On conditional extrema of functionals, *Math. Sbornik* **41**, 390–401.
484. G. G. MALCOLM AND B. S. MORDUKHOVICH (2001), Pareto optimality in nonconvex economies with infinite-dimensional commodity spaces, *J. Global Optim.* **20**, 323–346.
485. M. MARÉCHAL AND R. CORREA (2016), Error bounds, metric subregularity and stability in generalized Nash equilibrium problems with nonsmooth payoff functions, *Optimization* **65**, 1829–1854.
486. J.-E. MARTÍNEZ-LEGAZ AND B. F. SVAITER (2005), Monotone operators representable by l.s.c. convex functions, *Set-Valued Anal.* **13** (2005), 21–46.
487. J.-E. MARTÍNEZ-LEGAZ AND M. VOLLE (1999). Duality in DC programming: the case of several DC constraints, *J. Math. Anal. Appl.* **237**, 657–671.
488. H. MARTINI, K. J. SWANEPOEL AND G. WEISS (2002), The Fermat-Torricelli problem in normed planes and spaces, *J. Optim. Theory Appl.* **115**, 283–314.
489. A. MAS-COLELL (1985), Pareto optima and equilibria: the finite-dimensional case, in *Advances in Equilibrium Theory*, edited by C. D. Aliprantis, O. Burkinshaw and N. J. Rothman, Lecture Notes in Econom. Math. Systems **244**, pp. 25–42, Springer, Berlin.
490. A. MAS-COLLEL, M. D. WHINSTON AND J. R. GREEN (1995), *Mircoeconomic Theory*, Oxford University Press, Oxford, United Kingdom.
491. G. MASTROENI (2003), Gap functions for equilibrium problems, *J. Global Optim.* **27**, 411–426.
492. A. MAUGERI (2001), Equilibrium problems and variational inequalities, in *Equilibrium Problems: Nonsmooth Optimization and Variational Inequality Methods*, edited by F. Giannessi et al., pp. 187–205, Kluwer, Dordrecht, The Netherlands.

493. A. MAUGERI AND D. PUGLISI (2014), A new necessary and sufficient condition for the strong duality and the infinite dimensional Lagrange multiplier rule, *J. Math. Anal. Appl.* **415**, 661–676.
494. P. MEHLITZ AND G. WACHSMUTH (2017), The limiting normal cone to pointwise defined sets in Lebesgue spaces, *Set-Valued Var. Anal.*, DOI 10.1007/s11228-016-0393-4.
495. K. MENG AND X. YANG (2016), Variational analysis of weak sharp minima via exact penalization, *Set-Valued Var. Anal.* **24**, 619–635.
496. P. MICHEL AND J.-P. PENOT (1992), A generalized derivative for calm and stable functions, *Differ. Integr. Equ.* **5**, 433–454.
497. A. MIELKE, R. ROSSI AND G. SAVARÉ (2013), Nonsmooth analysis of doubly nonlinear evolution equations, *Calc. Var. Partial Diff. Eq.* **46**, 253–310.
498. E. MIGLIERINA, E. MOLHO AND M. ROCCA (2005), Well-posedness and scalarization in vector optimization, *J. Optim. Theory Appl.* **126**, 391–409.
499. L. MINCHENKO AND S. STAKHOVSKI (2011), Parametric nonlinear programming problems under the relaxed constant rank condition, *SIAM J. Optim.* **21**, 314–332.
500. S. K. MISHRA AND V. LAHA (2016), On Minty variational principle for nonsmooth vector optimization problems with approximate convexity, *Optim. Lett.* **10**, 577–589.
501. O. MLEŞNIŢE AND A. PETRUŞEL (2015), Metric regularity and Ulam-Hyers stability results for coincidence problems with multivalued operators, *J. Nonlinear Convex Anal.* **16** (2015), 1397–1413.
502. B. S. MORDUKHOVICH (1976), Maximum principle in problems of time optimal control with nonsmooth constraints, *J. Appl. Math. Mech.* **40**, 960–969.
503. B. S. MORDUKHOVICH (1977), Approximation and maximum principle for nonsmooth problems of optimal control, *Russian Math. Surveys* **196**, 263–264.
504. B. S. MORDUKHOVICH (1980), Metric approximations and necessary optimality conditions for general classes of extremal problems, *Soviet Math. Dokl.* **22**, 526–530.
505. B. S. MORDUKHOVICH (1984), Nonsmooth analysis with nonconvex generalized differentials and adjoint mappings, *Dokl. Akad. Nauk BSSR* **28**, 976–979 (in Russian).
506. B. S. MORDUKHOVICH (1985), On necessary conditions for an extremum in nonsmooth optimization, *Soviet Math. Dokl.* **32**, 215–220.
507. B. S. MORDUKHOVICH (1988), *Approximation Methods in Problems of Optimization and Control*, Nauka, Moscow (in Russian).
508. B. S. MORDUKHOVICH (1992), Sensitivity analysis in nonsmooth optimization, in *Theoretical Aspects of Industrial Design*, edited by D. A. Field and V. Komkov, SIAM Proc. Appl. Math. **58**, pp. 32–46, Philadelphia, Pennsylvania.
509. B. S. MORDUKHOVICH (1992), On variational analysis of differential inclusions, in *Optimization and Nonlinear Analysis*, edited by A. D. Ioffe, L.

Marcus and S. Reich, Pitman Research Notes Math. Ser. **244**, pp. 199–213, Longman, Harlow, Essex, United Kingdom.
510. B. S. MORDUKHOVICH (1993), Complete characterization of openness, metric regularity, and Lipschitzian properties of multifunctions, *Trans. Amer. Math. Soc.* **340**, 1–35.
511. B. S. MORDUKHOVICH (1994), Generalized differential calculus for nonsmooth and set-valued mappings, *J. Math. Anal. Appl.* **183**, 250–288.
512. B. S. MORDUKHOVICH (1994), Lipschitzian stability of constraint systems and generalized equations, *Nonlinear Anal.* **22**, 173–206.
513. B. S. MORDUKHOVICH (1994), Stability theory for parametetic generalized equations and variational inequalities via nonsmooth analysis, *Trans. Amer. Math. Soc.* **343**, 609–658.
514. B. S. MORDUKHOVICH (1997), Coderivatives of set-valued mappings: calculus and applications, *Nonlinear Anal.* **30**, 3059–3070.
515. B. S. MORDUKHOVICH (2000), Abstract extremal principle with applications to welfare economics, *J. Math. Anal. Appl.* **251**, 187–216.
516. B. S. MORDUKHOVICH (2001), The extremal principle and its applications to optimization and economics, in *Optimization and Related Topics*, edited by A. Rubinov and B. Glover, pp. 343–369, Kluwer, Dordrecht, The Netherlands.
517. B. S. MORDUKHOVICH (2002), Calculus of second-order subdifferentials in infinite dimensions, *Control and Cybernetics* **31**, 558–573.
518. B. S. MORDUKHOVICH (2004), Coderivative analysis of variational systems, *J. Global Optim.* **28**, 347–362.
519. B. S. MORDUKHOVICH (2004), Necessary conditions in nonsmooth minimization via lower and upper subgradients, *Set-Valued Anal.* **12**, 163–193.
520. B. S. MORDUKHOVICH (2004), Equilibrium problems with equilibrium constraints via multiobjective optimization, *Optim. Meth. Soft.* **19**, 479–492.
521. B. S. MORDUKHOVICH (2005), Nonlinear prices in nonconvex economies with classical Pareto and strong Pareto optimal allocations, *Positivity* **9**, 541–568.
522. B. S. MORDUKHOVICH (2006), *Variational Analysis and Generalized Differentiation, I: Basic Theory*, Springer, Berlin.
523. B. S. MORDUKHOVICH (2006), *Variational Analysis and Generalized Differentiation, II: Applications*, Springer, Berlin.
524. B. S. MORDUKHOVICH (2008), Failure of metric regularity for major classes of variational systems, *Nonlinear Anal.* **69**, 918–924.
525. B. S. MORDUKHOVICH (2009), Methods of variational analysis in multiobjective optimization, *Optimization* **58**, 413–430.
526. B. S. MORDUKHOVICH (2009), Multiobjective optimization problems with equilibrium constraints, *Math. Program.* **117**, 331–354.
527. B. S. MORDUKHOVICH (2018), *Second-Order Variational Analysis and Applications*, book in progress.

528. B. S. MORDUKHOVICH AND A. Y. KRUGER (1976), Necessary optimality conditions for a terminal control problem with nonfunctional constraints, *Dokl. Akad. Nauk BSSR* **20**, 1064–1067 (in Russian).
529. B. S. MORDUKHOVICH AND L. MOU (2009), Necessary conditions for nonsmooth optimization problems with operator constraints in metric spaces, *J. Convex Anal.* **16**, 913–938.
530. B. S. MORDUKHOVICH AND N. M. NAM (2005), Subgradients of distance functions with some applications, *Math. Program.* **104** (2005), 635–668.
531. B. S. MORDUKHOVICH AND N. M. NAM (2005), Subgradients of distance functions at out-of-state points, *Taiwanese J. Math.* **10** (2006), 299–326.
532. B. S. MORDUKHOVICH AND N. M. NAM (2005), Variational stability and marginal functions via generalized differentiation, *Math. Oper. Res.* **30**, 1–18.
533. B. S. MORDUKHOVICH AND N. M. NAM (2009), Variational analysis of generalized equations via coderivative calculus in Asplund spaces, *J. Math. Anal. Appl.* **350**, 663–679.
534. B. S. MORDUKHOVICH AND N. M. NAM (2010), Limiting subgradients of minimal time functions in Banach spaces, *J. Global Optim.* **46**, 615–633.
535. B. S. MORDUKHOVICH AND N. M. NAM (2011), Subgradients of minimal time functions under minimal requirements, *J. Convex Anal.* **18**, 915–947.
536. B. S. MORDUKHOVICH AND N. M. NAM (2011), Applications of variational analysis to a generalized Fermat-Torricelli problem, *J. Optim. Theory Appl.* **148**, 431–454.
537. B. S. MORDUKHOVICH AND N. M. NAM (2014), *An Easy Path to Convex Analysis and Applications*, Morgan & Claypool Publishers, San Rafael, California.
538. B. S. MORDUKHOVICH AND N. M. NAM (2017), Extremality of convex sets with some applications, *Optim. Lett.* **11**, 1201–1215.
539. B. S. MORDUKHOVICH, N. M. NAM AND N. T. Y. NHI (2014), Partial second-order subdifferentials in variational analysis and optimization, *Numer. Funct. Anal. Appl.* **35**, 1113–1151.
540. B. S. MORDUKHOVICH, N. M. NAM AND H. M. PHAN (2011), Variational analysis of marginal functions with applications to bilevel programming, *J. Optim. Theory Appl.* **152**, 557–586.
541. B. S. MORDUKHOVICH, N. M. NAM, R. B. RECTOR AND T. TRAN (2017), Variational geometric approach to generalized differential and conjugate calculi in convex analysis, *Set-Valued Var. Anal.* **25**, 731–755.
542. B. M. MORDUKHOVICH, N. M. NAM AND J. SALINAS, JR. (2012), Solving a generalized Heron problem by means of convex analysis, *Amer. Math. Monthly* **119**, 87–99.
543. B. M. MORDUKHOVICH, N. M. NAM AND J. SALINAS, JR. (2012), Applications of variational analysis to a generalized Heron problem, *Applic. Anal.* **91**, 1915–1942

544. B. M. MORDUKHOVICH, N. M. NAM AND M. C. VILLALOBOS (2012), The smallest enclosing ball problem and the smallest intersecting ball problem: existence and uniqueness of solutions, *Optim. Lett.* **154**, 768–791.
545. B. M. MORDUKHOVICH, N. M. NAM AND B. WANG (2009), Metric regularity of mappings and generalized normals to set images, *Set-Valued Var. Anal.* **17**, 359–387.
546. B. S. MORDUKHOVICH, N. M. NAM AND N. D. YEN (2006), Fréchet subdifferential calculus and optimality conditions in nondifferentiable programming, *Optimization* **55**, 685–396.
547. B. S. MORDUKHOVICH, N. M. NAM AND N. D. YEN (2009), Subgradients of marginal functions in parametric mathematical programming, *Math. Program.* **116**, 369–396.
548. B. S. MORDUKHOVICH AND T. T. A. NGHIA (2012), DC optimization approach to metric regularity of convex multifunctions with applications to stability of infinite systems, *J. Optim. Theory Appl.* **155**, 762–784.
549. B. S. MORDUKHOVICH AND T. T. A. NGHIA (2013), Constraint qualifications and optimality conditions in nonlinear semi-infinite and infinite programming, *Math. Program.* **139**, 271–300.
550. B. S. MORDUKHOVICH AND T. T. A. NGHIA (2013), Subdifferentials of nonconvex supremum functions and their applications to semi-infinite and infinite programs with Lipschitzian data, *SIAM J. Optim.* **23**, 406–431.
551. B. S. MORDUKHOVICH AND T. T. A. NGHIA (2013), Second-order variational analysis and characterizations of tilt-stable optimal solutions in infinite-dimensional spaces, *Nonlinear Anal.* **86**, 159–180.
552. B. S. MORDUKHOVICH AND T. T. A. NGHIA (2014), Full Lipschitzian and Hölderian stability in optimization with applications to mathematical programming and optimal control, *SIAM J. Optim.* **24**, 1344-1381.
553. B. S. MORDUKHOVICH AND T. T. A. NGHIA (2014), Nonsmooth cone-constrained optimization with applications to semi-infinite programming, *Math. Oper. Res.* **39**, 301–337.
554. B. S. MORDUKHOVICH AND T. T. A. NGHIA (2015), Second-order characterizations of tilt stability with applications to nonlinear programming, *Math. Program.* **149**, 83-104.
555. B. S. MORDUKHOVICH AND T. T. A. NGHIA (2016), Local monotonicity and full stability of parametric variational systems, *SIAM J. Optim.* **26**, 1032–1059.
556. B. S. MORDUKHOVICH, T. T. A. NGHIA AND R. T. ROCKAFELLAR (2016), Full stability in finite-dimensional optimization, *Math. Oper. Res.* **40**, 226–252.
557. B. S. MORDUKHOVICH AND J. V. OUTRATA (2001), On second-order subdifferentials and their applications, *SIAM J. Optim.* **12**, 139–169.
558. B. S. MORDUKHOVICH AND J. V. OUTRATA (2007) Coderivative analysis of quasi-variational inclusions with applications to stability and optimization, *SIAM J. Optim.* **18**, 389–412.

559. B. S. MORDUKHOVICH AND J. V. OUTRATA (2013), Tilt stability in nonlinear programming under Mangasarian-Fromovitz constraint qualification, *Kybernetika* **49**, 446–464.
560. B. S. MORDUKHOVICH, J. V. OUTRATA AND M. ČERNINKA (2007), Equilibrium problems with complementarity constraints: case study with applications to oligopolistic markets, *Optimization* **56**, 479–494.
561. B. S. MORDUKHOVICH, J. V. OUTRATA AND H. RAMÍREZ C. (2015), Second-order variational analysis in conic programming with applications to optimality conditions and stability, *SIAM J. Optim.* **25**, 76–101.
562. B. S. MORDUKHOVICH, J. V. OUTRATA AND H. RAMÍREZ C. (2015), Graphical derivatives and stability analysis for parameterized equilibria with conic constraints, *Set-Valued Var. Anal.* **23**, 687–704.
563. B. S. MORDUKHOVICH, J. V. OUTRATA AND M. E. SARABI (2014), Full stability in second-order cone programming, *SIAM J. Optim.* **24**, 1581–1613.
564. B. S. MORDUKHOVICH AND W. OUYANG (2015), Higher-order metric subregularity and its applications, *J. Global Optim.* **63**, 777–795.
565. B. S. MORDUKHOVICH, B. PANICUCCI, M. PASSACANTANDO AND M. PAPPALARDO (2012), Hybrid proximal methods for equilibrium problems, *Optim. Lett.* **6**, 1535–1550.
566. B. S. MORDUKHOVICH, J. PĒNA AND V. ROCHSHINA (2010), Applying metric regularity to compute a condition measure of smooth algorithms for matrix games, *SIAM J. Optim.* **20**, 3490–3511.
567. B. S. MORDUKHOVICH AND H. M. PHAN (2011), Rated extremal principle for finite and infinite systems with applications to optimization, *Optimization* **60**, 893–924.
568. B. S. MORDUKHOVICH AND H. M. PHAN (2012), Tangential extremal principle for finite and infinite systems, I: basic theory, *Math. Program.* **136**, 31–63.
569. B. S. MORDUKHOVICH AND H. M. PHAN (2012), Tangential extremal principle for finite and infinite systems, II: applications to semi-infinite and multiobjective optimization, *Math. Program.* **136**, 31–63.
570. B. S. MORDUKHOVICH AND R. T. ROCKAFELLAR (2012), Second-order subdifferential calculus with applications to tilt stability in optimization, *SIAM J. Optim.* **22**, 953–986.
571. B. S. MORDUKHOVICH, R. T. ROCKAFELLAR AND M. E. SARABI (2013), Characterizations of full stability in constrained optimization, *SIAM J. Optim.* **23**, 1810–1849.
572. B. S. MORDUKHOVICH AND N. SAGARA (2018), Subdifferentiation of nonconvex integral functionals on Banach spaces with applications to stochastic dynamic programming, *J. Convex Anal.* **25**, No. 2.
573. B. S. MORDUKHOVICH AND M. E. SARABI (2015), Variational analysis and full stability of optimal solutions to constrained and minimax problems, *Nonlinear Anal.* **121**, 36–53.

574. B. S. MORDUKHOVICH AND M. E. SARABI (2016), Generalized differentiation of piecewise linear functions in second-order variational analysis, *Nonlinear Anal.* **132** (2016), 240–273.
575. B. S. MORDUKHOVICH AND M. E. SARABI (2016), Second-order analysis of piecewise linear functions with applications to optimization and stability, *J. Optim. Theory Appl.* **171**, 504–526.
576. B. S. MORDUKHOVICH AND M. E. SARABI (2017), Stability analysis for composite optimization problems and parametric variational systems, *J. Optim. Theory Appl.* **172**, 554–577.
577. B. S. MORDUKHOVICH AND M. E. SARABI (2017), Critical multipliers in variational systems via second-order generalized differentiation, *Math. Program.*, DOI 10.1007/s10107-017-1155-2.
578. B. S. MORDUKHOVICH AND Y. SHAO (1995), Differential characterizations of covering, metric regularity, and Lipschitzian properties of multifunctions between Banach spaces, *Nonlinear Anal.* **25**, 1401–1424.
579. B. S. MORDUKHOVICH AND Y. SHAO (1996), Extremal characterizations of Asplund spaces, *Proc. Amer. Math. Soc.* **124**, 197–205.
580. B. S. MORDUKHOVICH AND Y. SHAO (1996), Nonsmooth sequential analysis in Asplund spaces, *Trans. Amer. Math. Soc.* **348**, 1235–1280.
581. B. S. MORDUKHOVICH AND Y. SHAO (1996), Nonconvex coderivative calculus for infinite-dimensional multifunctions, *Set-Valued Anal.* **4**, 205–136.
582. B. S. MORDUKHOVICH AND Y. SHAO (1997), Stability of multifunctions in infinite dimensions: point criteria and applications, *SIAM J. Control Optim.* **35**, 285–314.
583. B. S. MORDUKHOVICH AND Y. SHAO (1997), Fuzzy calculus for coderivatives of multifunctions, *Nonlinear Anal.* **29**, 605–626.
584. B. S. MORDUKHOVICH AND Y. SHAO (1998), Mixed coderivatives of set-valued mappings in variational analysis, *J. Appl. Anal.* **4**, 269–294.
585. B. S. MORDUKHOVICH, Y. SHAO AND Q. J. ZHU (2000), Viscosity coderivatives and their limiting behavior in smooth Banach spaces, *Positivity* **4**, 1–39.
586. B. S. MORDUKHOVICH, J. S. TREIMAN AND Q. J. ZHU (2003), An extended extremal principle with applications to multiobjective optimization, *SIAM J. Optim.* **14**, 359–379.
587. B. S. MORDUKHOVICH AND B. WANG (2002), Necessary optimality and suboptimality conditions in nondifferentiable programming via variational principles, *SIAM J. Control Optim.* **41**, 623–640.
588. B. S. MORDUKHOVICH AND B. WANG (2002), Extensions of generalized differential calculus in Asplund spaces, *J. Math. Anal. Appl.* **272**, 164–186.
589. B. S. MORDUKHOVICH AND B. WANG (2003), Calculus of sequential normal compactness in variational analysis, *J. Math. Anal. Appl.* **282**, 63–84.
590. B. S. MORDUKHOVICH AND B. WANG (2003), Differentiability and regularity of Lipschitzian mappings, *Proc. Amer. Math. Soc.* **131**, 389–399.
591. B. S. MORDUKHOVICH AND B. WANG (2004), Restrictive metric regularity and generalized differential calculus in Banach spaces, *Int. J. Maths. Math. Sci.* **50**, 2650–2683.

592. B. S. MORDUKHOVICH AND B. WANG (2008) Generalized differentiation of parameter-dependent sets and mappings, *Optimization* **57**, 17–40.
593. J.-J. MOREAU (1963), Fonctionelles sous-différentiables, *C. R. Acad. Sci. Paris* **257**, 4117–4119 (in French).
594. N. MOVAHEDIAN (2012), Calmness of set-valued mappings between Asplund spaces and application to equilibrium problems, *Set-Valued Var. Anal.* **20**, 499–518.
595. N. MOVAHEDIAN (2014), Nonsmooth calculus of semismooth functions and maps, *J. Optim. Theory Appl.* **160**, 415–438.
596. N. MOVAHEDIAN (2016), Necessary optimality conditions for countably infinite Lipschitz problems with equality constraint mappings, *Optim. Lett.* **10**, 63–76.
597. N. NADEZHKINA AND W. TAKAHASHI (2006), Strong convergence theorem by a hybrid method for nonexpansive mappings and Lipschitz-continuous monotone mappings, *SIAM J. Optim.* **16**, 1230–1241.
598. A. NAGURNEY (1999), *Network Economics: A Variational Inequality Approach*, 2nd edition, Kluwer, Dordrecht, The Netherlands.
599. N. M. NAM (2010), Coderivatives of normal cone mappings and applications to Lipschitzian stability, *Nonlinear Anal.* **73**, 2271–2282.
600. N. M. NAM (2015), Subdifferential formulas for a class of nonconvex infimal convolutions, *Optimization* **64**, 2113–2222.
601. N. M. NAM AND D. V. CUONG (2015), Generalized differentiation and characterizations for differentiability of infimal convolutions, *Set-Valued Var. Anal.* **23**, 333–353.
602. N. M. NAM AND N. D. HOANG (2013), A generalized Sylvester problem and a generalized Fermat-Torricelli problem, *J. Convex Anal.* **20**, 669–687.
603. N. M. NAM AND G. LAFFERRIERE (2013), Lipschitz properties of nonsmooth functions and set-valued mappings via generalized differentiation and applications, *Nonlinear Anal.* **89**, 110–120.
604. N. M. NAM AND C. ZĂLINESCU (2013), Variational analysis of directional minimal time functions and applications to location problems, *Set-Valued Var. Anal.* **21**, 405–430.
605. Y. NESTEROV (2005), Lexicographic differentiation of nonsmooth functions, *Math. Program.* **104**, 669–700.
606. L. W. NEUSTADT (1976), *Optimization: A Theory of Necessary Conditions*, Princeton University Press, Princeton, New Jersey.
607. K. F. NG AND X. Y. ZHENG (2001), Error bounds for lower semicontinuous functions in normed spaces, *SIAM J. Optim.* **12**, 1–17.
608. K. F. NG AND X. Y. ZHENG (2003), Global weak sharp minima on Banach spaces, *SIAM J. Control Optim.* **41**, 1868–1885.
609. H. V. NGAI, D. T. LUC AND M. THÉRA (2000), Approximate convex functions, *J. Nonlinear Convex Anal.* **1**, 155–176.
610. N. V. NGAI AND M. THÉRA (2001), Metric regularity, subdifferential calculus and applications, *Set-Valued Anal.* **9**, 187–216.

611. N. V. NGAI AND M. THÉRA (2002), A fuzzy necessary optimality condition for non-Lipschitz optimization in Asplund spaces, *SIAM J. Optim.* **12**, 656–668.
612. N. V. NGAI AND M. THÉRA (2004), Error bounds and implicit multifunction theorem in smooth Banach spaces and applications to optimization, *Set-Valued Anal.* **12**, 195–223.
613. N. V. NGAI AND M. THÉRA (2015), Directional metric regularity of multifunctions, *Math. Oper. Res.* **40**, 969–991.
614. T. T. A. NGHIA (2012), A nondegenerate fuzzy optimality condition for constrained optimization problems without qualification conditions, *Nonlinear Anal.* bf 75, 6379–6390.
615. S. NICKEL, J. PUERTO AND A. M. RODRIGUEZ-CHIA (2003), An approach to location models involving sets as existing facilities, *Math. Oper. Res.* **28**, 693–715.
616. F. NIELSEN AND R. NOCK (2009), Approximating smallest enclosing balls with applications to machine learning, *Int. J. Comput. Geom. Appl.* **19**, 389–414.
617. D. NOLL AND A. RONDEPIERRE (2016), On local convergence of the method of alternating projections, *Found. Comput. Math.* **16**, 425–455.
618. W. OETTLI (1982), Optimality conditions for programming problems involving multivalued mappings, in *Modern Applied Mathematics*, edited by B. H. Korte, pp. 195–226, North-Holland, Amsterdam.
619. J. V. OUTRATA (1990), On the numerical solution of a class of Stackelberg problems, *ZOR–Methods Models Oper. Res.* **34**, 255–277.
620. J. V. OUTRATA (1999), Optimality conditions for a class of mathematical programs with equilibrium constraints, *Math. Oper. Res.* **24**, 627–644.
621. J. V. OUTRATA (2006), Mathematical programs with equilibrium constraints: theory and numerical methods, in *Nonsmooth Mechanics of Solids*, edited by J. Haslinger and G. E. Stavroulakis, CISM Courses and Lecture Notes **485**, pp. 221–274, Springer, New York.
622. J. V. OUTRATA, F. C. FERRIS, M. ČERVINKA AND M. OUTRATA (2016), On Cournot-Nash-Walras equilibria and their computation, *Set-Valued Var. Anal* **24**, 387–402.
623. J. V. OUTRATA, J. JARUŠEK AND J. STARÁ (2011), On optimality conditions in control of elliptic variational inequalities, *Set-Valued Var. Anal.* **19**, 23–42.
624. J. V. OUTRATA, M. KOČVARA AND J. ZOWE (1998), *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints*, Kluwer, Dordrecht, The Netherlands.
625. J. V. OUTRATA AND H. RAMÍREZ C. (2011), On the Aubin property of critical points to perturbed second-order cone programs, *SIAM J. Optim.* **21**, 798–823; *Erratum* (with F. Opazo), *SIAM J. Optim.* **27** (2017), 2143–2151.
626. J. V. OUTRATA AND W. RÖMISCH (2005), On optimality conditions for some nonsmooth optimization problems over L^p spaces, *Optim. Theory Appl.* **126**, 1–28.

627. J. V. OUTRATA AND D. SUN (2008), On the coderivative of the projection operator onto the second-order cone, *Set-Valued Anal.* **16**, 999–1014.
628. M. OVEISIHA AND J. ZAFARANI (2012), Vector optimization problem and generalized convexity, *J. Global Optim.* **52**, 29–43.
629. D. PALLASCHKE AND S. ROLEWICZ (1998), *Foundations of Mathematical Optimization: Convex Analysis without Linearity*, Kluwer, Dordrecht, The Netherlands.
630. C. H. J. PANG (2011), Generalized differentiation with positively homogeneous maps: applications of set-valued analysis and metric regularity, *Math. Oper. Res.* **36**, 377–397.
631. C. H. J. PANG (2013), Characterizing generalized derivatives of set-valued maps: extending the tangential and normal approaches, *SIAM J. Optim.*, **51**, 145–171.
632. J.-S. PANG (1993), Convergence of splitting and Newton methods for complementarity problems: an application of some sensitivity results, *Math. Program.* **58**, 149–160.
633. T. PENNANEN (2002), Local convergence of the proximal point algorithm and multiplier methods without monotonicity, *Math. Oper. Res.* **27**, 170–191.
634. J.-P. PENOT (1974), Sous-diférentiels de fonctions numériques non convexes, *C. R. Acad. Sci. Paris* **278**, 1553–1555.
635. J.-P. PENOT (1998), Compactness properties, openness criteria and coderivatives, *Set-Valued Anal.* **6**, 363–380.
636. J.-P. PENOT (2001), Image space approach and subdifferentials of integral functionals, *Optimization* **60**, 69–87.
637. J.-P. PENOT (2013), *Calculus without Derivatives*, Springer, New York.
638. R. R. PHELPS (1993), *Convex Functions, Monotone Operators and Differentiability*, 2nd edition, Springer, Berlin.
639. T. POCK AND S. SABACH (2016), Inertial proximal alternating linearized minimization (iPALM) for nonconvex and nonsmooth problems, *SIAM J. Imaging Sci.* **9**, 1756–1787.
640. H. PHAN (2016), Linear convergence of the Douglas-Rachford method for two closed sets, *Optimization* **65**, 369–385.
641. R. A. POLIQUIN AND R. T. ROCKAFELLAR (1996), Prox-regular functions in variational analysis, *Trans. Amer. Math. Soc.* **348**, 1805–1838.
642. R. A. POLIQUIN AND R. T. ROCKAFELLAR (1998), Tilt stability of a local minimum, *SIAM J. Optim.* **8**, 287–299.
643. B. T. POLYAK (1979), *Sharp Minima*, Institute of Control Sciences Lecture Notes, Moscow; presented at the IIASA Workshop on Generalized Lagrangians and Their Applications, IIASA, Laxenburg, Austria.
644. B. T. POLYAK (1987), *Introduction to Optimization*, Optimization Software, New York.
645. L. S. PONTRYAGIN, V. G. BOLTYANSKII, R. V. GAMKRELIDZE AND E. F. MISHCHENKO (1962), *The Mathematical Theory of Optimal Processes*, Wiley, New York, 1962.

646. N. Popovici (2007), Explicitly quasiconvex set-valued optimization, *J. Global Optim.* **38**, 103–118.
647. D. Preiss (1990), Differentiability of Lipschitz functions on Banach spaces, *J. Funct. Anal.* **91**, 312–345.
648. B. N. Pshenichnyi (1976), Necessary conditions for an extremum for differential inclusions, *Kibernetika* **12**, 60–73 (in Russian).
649. B. N. Pshenichnyi (1980), *Convex Analysis and Extremal Problems*, Nauka, Moscow (in Russian).
650. H. Pühl and W. Schirotzek (2004), Linear semi-openness and the Lyusternik theorem, *European J. Oper. Res.* **157**, 16–27.
651. J.-H. Qiu (2014), A pre-order principle and set-valued Ekeland variational principle, *J. Math. Anal. Appl.* **419**, 904–937.
652. N. T. Qui (2014), Generalized differentiation of a class of normal cone operators, *J. Optim. Theory Appl.* **161**, 398–429.
653. N. T. Qui (2016), Coderivatives of implicit multifunctions and stability of variational systems, *J. Global Optim.* **65**, 615–635.
654. N. T. Qui and H. N. Tuan (2017), Stability of generalized equations under nonlinear perturbations, *Optim. Lett.*, DOI 10.1007/s11590-017-1147-4.
655. M. L. Radulescu and F. H. Clarke (1997), Geometric approximations of proximal normals, *J. Convex Anal.* **4**, 373–379.
656. S. Reich and A. Zaslavski (2014), *Genericity in nonlinear analysis*, Springer, New York.
657. J. Renegar (1995), Incorporating condition measures into the complexity theory of linear programming, *SIAM J. Optim.* **5**, 506–524.
658. S. M. Robinson (1972), Normed convex processes, *Trans. Amer. Math. Soc.* **174**, 127–140.
659. S. M. Robinson (1976), Regularity and stability for convex multivalued functions, *Math. Oper. Res.* **1**, 130–143.
660. S. M. Robinson (1976), Stability theory for systems of inequalities, II: differentiable nonlinear systems, *SIAM J. Numer. Anal.* **13**, 497–513.
661. S. M. Robinson (1979), Generalized equations and their solutions, I: basic theory, *Math. Program. Study* **10**, 128–141.
662. S. M. Robinson (1980), Strongly regular generalized equations, *Math. Oper. Res.* **5**, 43–62.
663. S. M. Robinson (1981), Some continuity properties of polyhedral multifunctions, *Math. Program. Study* **14**, 206–214.
664. S. M. Robinson (1991), An implicit function theorem for a class of nonsmooth functions, *Math. Oper. Res.* **16**, 292–309.
665. S. M. Robinson (1994), Newton's method for a class of nonsmooth functions, *Set-Valued Anal.* **2**, 291–305.
666. S. M. Robinson (2013), Equations on monotone graphs, *Math. Program.* **141** (2013), 49–101.
667. R. T. Rockafellar (1970), *Convex Analysis*, Princeton University Press, Princeton, New Jersey.

668. R. T. ROCKAFELLAR (1970), On the maximality of sums of nonlinear monotone operators, *Trans. Amer. Math. Soc.* **149**, 75–88.
669. R. T. ROCKAFELLAR (1979), Directional Lipschitzian functions and subdifferential calculus, *Proc. London Math. Soc.* **39**, 331–355.
670. R. T. ROCKAFELLAR (1980), Generalized directional derivatives and subgradients of nonconvex functions, *Canad. J. Math.* **32**, 157–180.
671. R. T. ROCKAFELLAR (1981), *The Theory of Subgradients and Its Applications to Problems of Optimization: Convex and Nonconvex Functions*, Helderman Verlag, Berlin.
672. R. T. ROCKAFELLAR (1981), Proximal subgradients, marginal values and augmented Lagrangians in nonconvex optimization, *Math. Oper. Res.* **6**, 424–436.
673. R. T. ROCKAFELLAR (1982), Favorable classes of Lipschitz continuous functions in subgradient optimization, in *Progress in Nondifferentiable Optimization*, edited by E. A. Nurminskii, pp. 125–143, IIASA, Laxenburg, Austria.
674. R. T. ROCKAFELLAR (1985), Lipschitzian properties of multifunctions, *Nonlinear Anal.* **9**, 867–885.
675. R. T. ROCKAFELLAR (1985), Extensions of subgradient calculus with applications to optimization, *Nonlinear Anal.* **9**, 665–698.
676. R. T. ROCKAFELLAR (1985), Maximal monotone relations and the second derivatives of nonsmooth functions, *Ann. Inst. H. Poincaré: Analyse Non Linéaire* **2**, 167–184.
677. R. T. ROCKAFELLAR (1989), Proto-differentiability of set-valued mappings and its applications in optimization, *Ann. Inst. H. Poincaré: Analyse Non Linéaire* **6**, 448–482.
678. R. T. ROCKAFELLAR AND R. J-B. WETS (1998), *Variational Analysis*, Springer, Berlin.
679. V. ROSHCHINA (2007), Relationships between upper exhausters and the basic subdifferential in variational analysis, *J. Math. Anal. Appl.* **334**, 261–272.
680. V. ROSHCHINA (2010), Mordukhovich subdifferential of pointwise minimum of approximate convex functions, *Optim. Methods Softw.* **25**, 129–141.
681. R. ROSSI AND G. SAVARÉ (2006), Gradient flows of nonconvex functionals in Hilbert spaces and applications, *ESAIM: Control Optim. Calc. Var.* **12**, 564–614.
682. A. RUSZCZYNSKI (2006), *Nonlinear Optimization*, Princeton University Press, Princeton, New Jersey.
683. N. SAGARA (2015), Cores and Weber sets for fuzzy extensions of cooperative games, *Fuzzy Sets Syst.* **272**, 102–114.
684. H. SCHEEL AND S. SCHOLTES (2000), Mathematical programs with complementarity constraints: stationarity, optimality, and sensitivity, *Math. Oper. Res.* **25**, 1–22.
685. W. SCHIROTZEK (2007), *Nonsmooth Analysis*, Springer, Berlin.
686. T. I. SEIDMAN (2010), Normal cones to infinite intersections, *Nonlinear Anal.* **72**, 3911–3917.

687. F. SEVERI (1930), Su alcune questioni di topologia infinitesimale, *Ann. Soc. Polon. Math.* **9**, 97–108 (in Italian).
688. A. SHAPIRO (2009), Semi-infinite programming: duality, discretization and optimality conditions, *Optimization* **58**, 133–161.
689. S. SIMONS (2008), *From Hahn-Banach to Monotonicity*, 2nd edition, Springer, Berlin.
690. S. SIMONS AND X. WANG (2015), Ubiquitous subdifferentials, rL–density and maximal monotonicity, *Set-Valued Var. Anal.* **23**, 631–642.
691. M. SION (1958), On general minimax theorems, *Pac. J. Math.* **8**, 171–176.
692. M. SOLEIMANI-DAMANEH (2010), Nonsmooth optimization using Mordukhovich's subdifferential, *SIAM J. Control Optim.* **48**, 3403–3432.
693. W. SONG (2006), Calmness and error bounds for convex constraint systems, *SIAM J. Optim.* **17**, 353–371.
694. W. SONG AND R. ZANG (2006), Bounded linear regularity of convex sets in Banach spaces and its applications, *Math. Program.* **106**, 59–79.
695. A. SOUBEYRAN (2009), Variational rationality, a theory of individual stability and change: worthwhile and ambidextry behaviors, mimeo, GREQAM, Aix-Marseille University, France.
696. O. STEIN (2003), *Bilevel Strategies in Semi-Infinite Programming*, Kluwer, Boston, Massachusetts.
697. M. STUDNIARSKI AND D. E. WARD (1999), Weak sharp minima: characterizations and sufficient conditions, *SIAM J. Control Optim.* **38**, 219–236.
698. A. I. SUBBOTIN (1995), *Generalized Solutions of First-Order PDEs*, Birkhäuser, Boston, Massachusetts.
699. N. N. SUBBOTINA (1989), The maximum principle and the superdifferential of the value function, *Prob. Contr. Inform. Theory* **18**, 151–160.
700. D. SUN (2006), The strong second-order sufficient condition and constraint nondegeneracity in nonlinear semidefinite programming and their applications, *Math. Oper. Res.* **31**, 761–776.
701. A. TAA (2011), On subdifferential calculus for set-valued mappings and optimality conditions, *Nonlinear Anal.* **74**, 7312–7324.
702. S. TAGAWA (1978), *Optimierung mit Mengenwertigen Abbildugen*, Ph.D. dissertation, University of Mannheim, Germany (in German).
703. T. TANINO AND T. OGAWA (1984), An algorithm for solving two-level convex optimization problems, *Inter. J. Syst. Sci.* **15**, 163–174.
704. L. THIBAULT (1980), Subdifferentials of compactly Lipschitzian vector functions, *Ann. Mat. Pura Appl.* **125**, 157–192.
705. L. THIBAULT (1983), Tangent, cones and quasi-interiorly tangent cones to multifunctions, *Trans. Amer. Math. Soc.* **277**, 601–621.
706. L. THIBAULT (1991), On subdifferentials of optimal value functions, *SIAM J. Control Optim.* **29**, 1019–1036.
707. L. THIBAULT (1995), A note on the Zagrodny mean value theorem, *Optimization* **35**, 127–130.
708. L. THIBAULT (1997), Sequential convex subdifferential calculus and sequential Lagrange multipliers, *SIAM J. Control Optim.* **35**, 1434–1444.

709. L. THIBAULT (1997), On compactly Lipschitzian mappings, in *Recent Advances in Optimization*, edited by P. Gritzmann et al., Lecture Notes Econ. Math. Syst. **456**, pp. 356–364, Springer, Berlin.
710. L. THIBAULT AND D. ZAGRODNY (1995), Integration of subdifferentials of lower semicontinuous functions, *J. Math. Anal. Appl.* **189**, 22–58.
711. V. D. THINH AND T. D. CHUONG (2017), Directionally generalized differentiation for multifunctions and applications to set-valued programming problems, *Ann. Oper. Res.*, DOI 10.1007/s10479-017-2400-z.
712. L. Q. THUY AND N. T. TOAN (2016), Subgradients of the value function in a parametric convex optimal control problem, *J. Optim. Theory Appl.* **170**, 43–64.
713. N. T. TOAN AND J.-C. YAO (2014), Mordukhovich subgradients of the value function to a parametric discrete optimal control problem, *J. Global Optim.* **58**, 595–612.
714. N. T. Q. TRANG (2012), A note on an approximate mean value theorem for Fréchet subgradients, *Nonlinear Anal.* **75**, 380–383.
715. J. S. TREIMAN (1999), Lagrange multipliers for nonconvex generalized gradients with equality, inequality, and set constraints, *SIAM J. Control Optim.* **37**, 1313–1329.
716. J. S. TREIMAN (2002), The linear generalized gradient in infinite dimensions, *Nonlinear Anal.* **48**, 427–443.
717. N. V. TUYEN (2016), Some characterizations of solution sets of vector optimization problems with generalized order, *Acta Math. Vietnamica* **41**, 677–694.
718. N. V. TUYEN AND N. D. YEN (2012), On the concept of generalized order optimality, *Nonlinear Anal.* **75**, 1592–1601.
719. A. UDERZO (2009), On some regularity properties in variational analysis, *Set-Valued Var. Anal.* **17**, 409–430.
720. A. UDERZO (2010), Exact penalty functions and calmness for mathematical programming under nonlinear perturbations, *Nonlinear Anal.* **73**, 1596–1609.
721. A. UDERZO (2012), A metric version of Milyutin theorem, *Set-Valued Var. Anal.* **20**, 279–306.
722. A. UDERZO (2014), Localizing vector optimization problems with application to welfare economics, *Set-Valued Var. Anal.* **22**, 483–501.
723. A. UDERZO (2015), Convexity of the images of small balls through nonconvex multifunctions, lem Nonlinear Anal. **128**, 348–364.
724. A. UDERZO (2016), A strong metric regularity analysis of nonsmooth mappings via steepest displacement rate, *J. Optim. Theory Appl.* **171**, 573–599.
725. A. UDERZO (2017), An implicit multifunction theorem for the hemiregularity of mappings with applications to constrained optimization, to appear in *Pure Appl. Funct. Anal.*, arXiv:1703.10552.
726. C. URSESCU (1975), Multifunctions with closed convex graphs, *Czech. Math. J.* **25**, 438–441.
727. T. VALKONEN (2014), A primal-dual hybrid gradient method for nonlinear operators with applications to MRI, *Inverse Prob.* **30**, 055012.

728. M. VALADIER (1969), Sous-différentiels d'une borne supérieure de fonctions convexes supérieure et d'une somme continue de fonctions convexes, *C.R. Acad. Sci. Paris Sér. A-B* **268**, 39–42 (in French).
729. R. B. VINTER (2000), *Optimal Control*, Birkhäuser, Boston, Massachusetts.
730. G. WACHSMUTH (2016), Towards M-stationarity for optimal control of the obstacle problem with control constraints, *SIAM J. Control Optim.* **54**, 964–986.
731. B. WANG AND X. YANG (2016), Weak differentiability with applications to variational analysis, *Set-Valued Var. Anal.* **24**, 299–321.
732. B. WANG AND D. WANG (2015), Generalized sequential normal compactness in Asplund spaces, *Applic. Anal.* **94**, 99–107.
733. J. J. WANG AND W. SONG (2014), Characterization of the strong metric subregularity of the Mordukhovich subdifferential on Asplund spaces, *Abstr. Appl. Anal.* 2014, 596582.
734. J. J. WANG AND W. SONG (2015), Characterization of quadratic growth of extended-real-valued functions, *J. Inequal. Appl.* 2016:29, DOI 10.1186/s13660-016-0977-4.
735. D. E. WARD AND G. M. LEE (2002), On relations between vector optimization problems and vector variational inequalities, *J. Optim. Theory Appl.* **113** (2002), 583–596.
736. J. WARGA (1976), Derivate containers, inverse functions, and controllability, in *Calculus of Variations and Control Theory*, edited by D. L. Russel, pp. 13–46, Academic Press, New York.
737. J. WARGA (1978), Controllability and a multiplier rule for nondifferentiable optimization problems, *SIAM J. Control Optim.* **16**, 803–812.
738. H. WOLKOWISZ, R. SAIGAL AND L. VANDENBERGHE, EDS. (2000), *Handbook of Semidefinite Programming: Theory, Algorithms and Applications*, Kluwer, Dordrecht, The Netherlands.
739. H. XU AND J. J. YE (2010), Necessary optimality conditions for two-stage stochastic programs with equilibrium constraints, *SIAM J. Optim.* **20**, 1685–1715.
740. X. XUE AND Y. ZHANG (2015), Coderivatives of gap function for Minty vector variational inequality, *J. Inequal. Appl.* 2015:285.
741. J.-C. YAO AND N. D. YEN (2009), Coderivative calculation related to a parametric affine variational inequality, I: basic calculation, *Acta Math. Vietnam.* **34**, 157–172.
742. J.-C. YAO AND N. D. YEN (2009), Coderivative calculation related to a parametric affine variational inequality, II: applications, *Pac. J. Optim.* **5**, 493–506.
743. J.-C. YAO, X. Y. ZHENG AND J. ZHU (2017), Stable minimizers of φ-regular functions, *SIAM J. Optim.* **27**, 1150–1170.
744. J. J. YE (1998), New uniform parametric error bounds, *J. Optim. Theory Appl.* **98**, 197–219.

745. J. J. YE (2000), Constraint qualifications and necessary optimality conditions for optimization problems with variational inequality constraints, *SIAM J. Optim.* **10**, 943–962.
746. J. J. YE AND X. Y. YE (1997), Necessary optimality conditions for optimization problems with variational inequality constraints, *Math. Oper. Res.* **22**, 977–997.
747. J. J. YE AND J. ZHOU (2017), Exact formulas for the proximal/regular/limiting normal cone of the second-order cone complementarity set, *Math. Program.*, DOI 10.1007/s10107-016-1027-1.
748. J. J. YE AND D. L. ZHU (1995), Optimality conditions for bilevel programming problems, *Optimization* **33**, 9–27.
749. J. J. YE AND D. L. ZHU (1997), A note on optimality conditions for bilevel programming problems, *Optimization* **39**, 361–366.
750. J. J. YE AND D. L. ZHU (2010), New necessary optimality conditions for bilevel programs by combining MPEC and the value function approach, *SIAM J. Optim.* **20**, 1885–1905.
751. N. D. YEN AND J.-C. YAO (2009), Pointbased sufficient conditions for metric regularity of implicit multifunctions, *Nonlinear Anal.* **70** (2009), 2806–2815.
752. N. D. YEN, J.-C. YAO AND B. T. KIEN (2008), Covering properties at positive-order rates of multifunctions and some related topics, *J. Math. Anal. Appl.* **338** (2008), 467–478.
753. L. C. YOUNG (1969), *Lectures on the Calculus of Variations and Optimal Control Theory*, Saunders, Philadelphia, Pennsylvania.
754. P. L. YU (1974), Cone convexity, cone extreme points, and nondominated solutions in decision problems with multiobjectives, *J. Optim. Theory Appl.* **14**, 319–377.
755. A. ZAFFARONI, Degrees of efficiency and degrees of minimality, *SIAM J. Control Optim.* **42**, 1071–1086.
756. D. ZAGRODNY (1988), Approximate mean value theorem for upper subderivatives, *Nonlinear Anal.* **12**, 1413–1428.
757. C. ZĂLINESCU (2002), *Convex Analysis in General Vector Spaces*, World Scientific, Singapore.
758. C. ZĂLINESCU (2015), On the use of the quasi-relative interior in optimization, *Optimization* **64**, 1795–1823.
759. E. H. ZARANTONELLO (1960), Solving functional equations by contractive averaging, University of Wisconsin, MRC Report 160, Madison, Wisconsin.
760. E. H. ZARANTONELLO (1971), Projections on convex sets in Hilbert space and spectral theory, I: projections on convex sets, *Contributions to Nonlinear Functional Analysis*, pp. 237–341, Academic Press, New York.
761. A. J. ZASLAVSKI (2010), Exact penalty in constrained optimization and the Mordukhovich basic subdifferential, in *Variational Analysis and Generalized Differentiation in Optimization and Control*, edited by R. S. Burachik and J.-C. Yao, pp. 223–232, Springer, New York.
762. A. J. ZASLAVSKI (2010), OPTIMIZATION ON METRIC AND NORMED SPACES, Springer, New York.

763. A. J. ZASLAVSKI (2012), Necessary optimality conditions for bilevel minimization problems, *Nonlinear Anal.* **75**, 1655–1678.
764. A. ZEMKOHO (2016), Solving ill-posed bilevel programs, *Set-Valued Var. Anal.* **24**, 423–448.
765. B. ZHANG, K. F. NG, X. Y. ZHENG AND Q. HE (2016), Hölder metric subregularity for multifunctions in $\mathcal{C}^2$ type Banach spaces, *Optimization* **65**, 1963–1982.
766. J. ZHANG, Y. LI AND L. ZHANG (2015), On the coderivative of the solution mapping to a second-order cone constrained parametric variational inequality, *J. Global Optim.* **61**, 379–396.
767. J. ZHANG, H. WANG AND Y. SUN (2015), A note on the optimality condition for a bilevel programming, *J. Inequal. Appl.* 361:2015, DOI 10.1186/s13660-015-0882-2.
768. R. ZHANG (1994), Problems of hierarchical optimization in finite dimensions, *SIAM J. Optim.* **4**, 521–536.
769. R. ZHANG (2003), Multistage bilevel programming problems, *Optimization* **52**, 605–616.
770. R. ZHANG (2005), Weakly upper Lipschitzian multifunctions and applications to parametric optimization, *Math. Program.* **102**, 153–166.
771. R. ZHANG AND J. S. TREIMAN (1995), Upper-Lipschitz multifunctions and inverse subdifferentials, *Nonlinear Anal.* **24**, 273–286.
772. X. Y. ZHENG (2016), Metric subregularity for a multifunction, *J. Math. Study* **49**, 379–392.
773. X. Y. ZHENG AND K. F. NG (2005), The Fermat rule for multifunctions in Banach spaces, *Math. Program.* **104**, 69–90.
774. X. Y. ZHENG AND K. F. NG (2006), The Lagrange multiplier rule for multifunctions in Banach spaces, *SIAM J. Optim.* **17**, 1154–1175.
775. X. Y. ZHENG AND K. F. NG (2009), Calmness for L-subsmooth multifunction in Banach spaces, *SIAM J. Optim.* **20**, 1648–1679.
776. X. Y. ZHENG AND K. F. NG (2010), Metric subregularity and calmness for nonconvex generalized equations in Banach spaces, *SIAM J. Optim.* **20**, 2119–2136.
777. X. Y. ZHENG AND K. F. NG (2011), A unified separation theorem for closed sets in a Banach spaces and optimality conditions for vector optimization, *SIAM J. Optim.* **21**, 886–911.
778. X. Y. ZHENG AND K. F. NG (2012), Subsmooth semi-infinite and infinite optimization problems, *Math. Program.* **134**, 365–393.
779. X. Y. ZHENG AND K. F. NG (2014), Metric subregularity of piecewise linear multifunctions and applications to piecewise linear multiobjective optimization, *SIAM J. Optim.* **24**, 154–174.
780. X. Y. ZHENG AND K. F. NG (2015), Hölder stable minimizers, tilt stability and Hölder metric regularity of subdifferentials, *SIAM J. Optim.* **25**, 416–438.
781. X. Y. ZHENG AND K. F. NG (2015), Hölder weak sharp minimizers and Hölder tilt-stability, *Nonlinear Anal.* **120**, 186–201.

782. X. Y. Zheng and X. Q. Yang (2007), Weak sharp minima for semi-infinite optimization problems with applications, *SIAM J. Optim.* **18**, 573–588.
783. X. Y. Zheng and X. Q. Yang (2007), Lagrange multipliers in nonsmooth semi-infinite optimization, *Math. Oper. Res.* **32**, 168–181.
784. J. Zhou, J. S. Chen and B. S. Mordukhovich (2015), Variational analysis of circular cone programs, *Optimization* **64**, 113–147.
785. J. Zhou, B. S. Mordukhovich and N. Xiu (2012), Complete characterizations of local weak sharp minima with applications to semi-infinite optimization and complementarity, *Nonlinear Anal.* **75**, 1700–1718.
786. Q. J. Zhu (1998), The equivalence of several basic theorems for subdifferentials, *Set-Valued Anal.* **6**, 171–185.
787. Q. J. Zhu (2004), Nonconvex separation theorem for multifunctions, subdifferential calculus and applications, *Set-Valued Anal.* **12**, 275–290.
788. S. K. Zhu (2016), Weak sharp efficiency in multiobjective optimization, *Optim. Lett.* **10**, 1287–1301.
789. S. E. Zhukovskiy (2015), On covering properties in variational analysis and optimization, *Set-Valued Var. Anal.* **23** (2015), 415–424.
790. S. E. Zhukovskiy (2016), Comparison of some types of locally covering mappings, *Fixed Point Theory* **17**, 215–222.

List of Statements

Chapter 1

B. S. Mordukhovich, *Variational Analysis and Applications*, Springer Monographs in Mathematics, https://doi.org/10.1007/978-3-319-92775-6

Chapter 2

Chapter 4

4.8 Corollary:	Subdifferential quotient rules
4.9 Proposition:	Basic and singular subdifferentials of minimum functions
4.10 Theorem:	Subdifferentiation of maximum functions
4.11 Theorem:	Symmetric subdifferential mean value theorem for continuous functions
4.12 Corollary:	Symmetric subdifferential mean value theorem for Lipschitzian functions
4.13 Theorem:	Approximate mean value theorem for l.s.c. functions
4.14 Corollary:	Mean value inequalities
4.15 Theorem:	Subdifferential characterizations of local Lipschitz continuity
4.16 Corollary:	Subdifferential characterization of constancy for l.s.c. functions
4.17 Theorem:	Subdifferential characterizations of strict differentiability
4.18 Theorem:	Subdifferential characterization of monotonicity for l.s.c. functions
4.19 Theorem:	Subdifferential monotonicity and convexity for l.s.c. functions
4.20 Exercise:	Subdifferentials of marginal functions in infinite dimensions
4.21 Exercise:	Extended inner semicontinuity and inner semicompactness
4.22 Exercise:	Equality representations for subgradients of marginal functions
4.23 Exercise:	Regular subgradients of value functions for parametric nonlinear programs
4.24 Exercise:	Regular subgradients of value functions for nondifferentiable programs
4.25 Exercise:	Lipschitz continuity of marginal functions in finite and infinite dimensions
4.26 Exercise:	Subdifferentials of infimal convolutions in Asplund spaces
4.27 Exercise:	Subdifferentiation of marginal functions and convex infimal convolutions
4.28 Exercise:	Subgradients of compositions in infinite dimensions
4.29 Exercise:	Subdifferential product and quotient rules in infinite dimensions
4.30 Exercise:	Partial subgradients
4.31 Exercise:	Regular and limiting subgradients of minimum functions
4.32 Exercise:	Subgradients of maximum functions on Asplund spaces
4.33 Exercise:	Symmetric mean value theorem in Asplund spaces
4.34 Exercise:	Approximate mean value theorem and its applications
4.35 Exercise:	Approximate mean value theorem via basic subgradients
4.36 Exercise:	Relationships between basic subgradients and Clarke counterparts

Chapter 5

5.1 Definition:	Hypomonotonicity
5.2 Theorem:	Maximal monotonicity via regular coderivative and global hypomonotonicity
5.3 Lemma:	Semilocal monotonicity of set-valued mappings with convex domains
5.4 Theorem:	Maximal monotonicity via regular coderivative and semilocal hypomonotonicity
5.5 Example:	Semilocal monotonicity doesn't yield the convexity of the domain
5.6 Theorem:	Limiting coderivative characterizations of global maximal monotonicity
5.7 Remark:	Preservation of maximal monotonicity
5.8 Example:	Hypomonotonicity conditions are essential
5.9 Corollary:	Coderivative characterizations of strong global maximal monotonicity
5.10 Definition:	Locally monotone and strongly monotone operators
5.11 Definition:	Single-valued localizations
5.12 Definition:	Strong metric regularity
5.13 Theorem:	Strong local maximal monotonicity via Lipschitzian localization
5.14 Theorem:	Neighborhood coderivative characterization of strong local maximal monotonicity
5.15 Corollary:	Sufficient conditions for strong optimality conditions for metric regularity
5.16 Theorem:	Pointbased coderivative conditions for strong local maximal monotonicity
5.17 Remark:	Pointbased coderivative criteria for strong local maximal monotonicity
5.18 Exercise:	Hypomonotonicity of single-valued mappings
5.19 Exercise:	Hypomonotonicity of subgradient mappings
5.20 Exercise:	Calculus of hypomonotonicity
5.21 Exercise:	Global maximal monotonicity for mappings with convex domains
5.22 Exercise:	Coderivative characterizations of global maximal monotonicity

Chapter 7

Chapter 8

Chapter 9

List of Figures

B. S. Mordukhovich, *Variational Analysis and Applications*, Springer Monographs in Mathematics, https://doi.org/10.1007/978-3-319-92775-6

Glossary of Notation and Acronyms

Operations and Symbols

$:=$ and $=:$	equal by definition
$\equiv$	identically equal
$*$	indication of some dual/adjoint/polar operation
$\langle\cdot,\cdot\rangle$	canonical pairing between space X and its topological dual X^*
$\\|\cdot\\|$ and $\|\cdot\|$	norm and absolute value (for a real number), respectively
$[\alpha]_+$	$\max\{\alpha,0\}$ for a real number α
$x\to\bar{x}$	x converges to $\bar{x}$ strongly (by norm)
$x\xrightarrow{w^*}\bar{x}$	x converges to $\bar{x}$ weak* (in weak* topology of X^*)
w^*-lim	weak* topological/net limit
$\liminf$	lower limit for real numbers
$\limsup$	upper limit for real numbers
$\operatorname{Lim\,inf}$	inner/lower sequential limit for set-valued mappings
$\operatorname{Lim\,sup}$	outer/upper sequential limit for set-valued mappings
$\dim X$ and $\operatorname{codim} X$	dimension and codimension of X, respectively
$\Pi_{i\in I}X_i$	Cartesian product of X_i
$\prec$ and $\preceq$	preference relations: "less" and "less or equal", respectively
$\preceq_\Theta$	"less or equal" preference with respect to cone Θ
$\operatorname{haus}(\Omega_1,\Omega_2)$	Pompeiu-Hausdorff distance between sets
$\beta(\Omega_1,\Omega_2)$	Hausdorff semidistance between sets
$\varphi_1\oplus\varphi_2$	infimal convolution of two functions
$\operatorname{lip} F(\bar{x},\bar{y})$	exact Lipschitzian bound of F around $(\bar{x},\bar{y})$
$\operatorname{clm} F$	exact bound of calmness of F at $(\bar{x},\bar{y})$
$\operatorname{reg} F(\bar{x},\bar{y})$	exact metric regularity bound of F around $(\bar{x},\bar{y})$
$\operatorname{hemireg} F(\bar{x},\bar{y})$	exact bound of hemiregularity of F at $(\bar{x},\bar{y})$
$\operatorname{cov} F(\bar{x},\bar{y})$	exact covering/linear openness bound of F around $(\bar{x},\bar{y})$
$\operatorname{Min}\Xi=\operatorname{Min}_\Theta\Xi$	collections of Pareto minimal points of Ξ with respect to cone Θ
$\operatorname{wMin}\Xi=\operatorname{wMin}_\Theta\Xi$	collections of weak Pareto minimal points of Ξ with respect to Θ
$\triangle$	end of proof

Spaces

$\mathbb{R}:=(-\infty,\infty)$	real line
$\overline{\mathbb{R}}:=[-\infty,\infty]$	extended real line
$\mathbb{R}^n$	n-dimensional Euclidean space

B. S. Mordukhovich, *Variational Analysis and Applications*, Springer Monographs in Mathematics, https://doi.org/10.1007/978-3-319-92775-6

$\mathbb{R}^n_+$ and $\mathbb{R}^n_-$	nonnegative and nonpositive orthant of $\mathbb{R}^n$, respectively
$\mathbb{R}^T$	space of $\lambda = (\lambda_t \mid t \in T)$ with $\lambda_t \in \mathbb{R}$, $t \in T$
$\mathbb{R}^{(T)}$	subspace of $\mathbb{R}^T$ with $\lambda_t \neq 0$ for finitely many $t \in T$
$\mathbb{R}^{(T)}_+$	positive (nonnegative) cone in $\mathbb{R}^{(T)}$
$\operatorname{supp}\lambda$	support $\{t \in T \mid \lambda_t \neq 0\}$ for $\lambda \in \mathbb{R}^{(T)}$
$l^\infty(T)$	bounded functions on T with norm $\|p\|_\infty := \sup\{\|p(t)\| \mid t \in T\}$
$l^\infty_+(T)$	positive (nonnegative) cone in $l^\infty(T)$.
$\mathcal{C}(T)$	continuous functions on compact T with $\|p\| := \max\{\|p(t)\| \mid t \in T\}$
$\mathcal{C}_+(T)$	positive (nonnegative) cone in $\mathcal{C}(T)$
$ba(T)$	space of bounded and additive measures on T
$ba_+(T)$	collection of nonnegative bounded and additive measures on T
$rba(T)$	space of regular finite Borel measures on T
$rba_+(T)$	collection of nonnegative regular finite Borel measures on T
$\mathcal{C}^1$	class of functions locally continuously differentiable
$\mathcal{C}^2$	class of functions locally twice continuously differentiable
$\mathcal{C}^{1,1}$	subclass of $\mathcal{C}^1$ with locally Lipschitzian derivatives
c	space of real number sequences with supremum norm
c_0	subspace of c with sequences converging to zero

Sets

$\emptyset$	empty set
$\mathbb{N}$	set of natural numbers
$\mathcal{N}$	notation for nets
$x \xrightarrow{\Omega} \bar{x}$	x converges to $\bar{x}$ with $x \in \Omega$
$B_r(x)$	ball centered at x with radius r
$\mathbb{B}$ and $\mathbb{B}^*$	closed unit balls of space and dual space in question
S_X and S^*	unit spheres in X and in dual space in question, respectively
$\operatorname{int}\Omega$ and $\operatorname{ri}\Omega$	interior and relative interior of Ω, respectively
$\operatorname{qri}\Omega$ and $\operatorname{iri}\Omega$	quasi-relative and intrinsic relative interiors of convex set Ω, respectively
$\operatorname{core}\Omega$	algebraic core of convex set Ω
$\operatorname{cl}\Omega$ and $\operatorname{cl}^*\Omega$	closure and weak* topological closure of Ω, respectively
$\operatorname{bd}\Omega$	set boundary
$\operatorname{co}\Omega$ and $\operatorname{clco}\Omega$	convex hull and closed convex hull of Ω, respectively
$\operatorname{cone}\Omega$	conic hull (convex conic hull in Chapters 7,8) of Ω
Ω^+	positive polar cone to convex cone Ω
$\operatorname{aff}\Omega$ and $\overline{\operatorname{aff}}\,\Omega$	affine hull and closed affine hull of Ω, respectively
$\operatorname{proj}_x\Omega$ and $\operatorname{proj}_X\Omega$	x-projection of sets in product spaces
$\Pi(x;\Omega)$ and $\Pi_\Omega(\bar{x})$	Euclidean projector of x to Ω
$N(\bar{x};\Omega)$	(basic, limiting) normal cone to Ω at $\bar{x}$
$\widehat{N}(\bar{x};\Omega)$	prenormal cone or regular normal cone to Ω at $\bar{x}$
$\widehat{N}_\varepsilon(\bar{x};\Omega)$	set of ε-normals to Ω at $\bar{x}$
$\overline{N}(\bar{x};\Omega)$	convexified or Clarke normal cone to Ω at $\bar{x}$
$T(\bar{x};\Omega)$	contingent cone to Ω at $\bar{x}$
$T_W(\bar{x};\Omega)$	weak contingent cone to Ω at $\bar{x}$
$\widehat{T}(\bar{x};\Omega)$	regular tangent cone to Ω at $\bar{x}$

Functions

$\delta(\cdot;\Omega)$	indicator function of Ω
χ_Ω	indicator function of Ω
$\sigma(\cdot;\Omega)$ or σ_Ω	support function of Ω
$\operatorname{dist}(\cdot;\Omega)$ or d_Ω	distance function for Ω
δ_t	Dirac function/measure with support at t

p_F	Minkowski gauge associate with set F
$\tau_F(\cdot;\Omega)$	minimal time function associated with dynamic F and target Ω
$\mathrm{dom}\,\varphi$	domain of $\varphi\colon X \to \overline{\mathbb{R}}$
$\mathrm{epi}\,\varphi$, $\mathrm{hypo}\,\varphi$, and $\mathrm{gph}\,\varphi$	epigraph, hypergraph, and graph of φ, respectively
φ_λ	Moreau envelope of φ with rate $\lambda > 0$
φ^* and φ^{**}	Fenchel conjugate and biconjugate of φ, respectively
$x \stackrel{\varphi}{\to} \bar{x}$	$x \to \bar{x}$ with $\varphi(x) \to \varphi(\bar{x})$
$x \stackrel{\varphi+}{\to} \bar{x}$	$x \to \bar{x}$ with $\varphi(x) \to \varphi(\bar{x})$ and $\varphi(x) \ge \varphi(\bar{x})$
$\varphi'(\bar{x})$ or $\nabla\varphi(\bar{x})$	(Fréchet) derivative/gradient of φ at $\bar{x}$
$\partial\varphi(\bar{x})$	(basic/limiting) subdifferential of φ at $\bar{x}$
$\partial^+\varphi(\bar{x})$	upper subdifferential of φ at $\bar{x}$
$\partial^0\varphi(\bar{x})$	symmetric subdifferential of φ at $\bar{x}$
$\partial_{\ge}\varphi(\bar{x})$	right-sided subdifferential of φ at $\bar{x}$
$\partial^\infty\varphi(\bar{x})$	singular subdifferential of φ at $\bar{x}$
$\partial^{\infty,+}\varphi(\bar{x})$	upper singular subdifferential of φ at $\bar{x}$
$\partial^{\infty,0}\varphi(\bar{x})$	symmetric singular subdifferential of φ at $\bar{x}$
$\widehat{\partial}\varphi(\bar{x})$	presubdifferential or regular/Fréchet subdifferential of φ at $\bar{x}$
$\widehat{\partial}_\varepsilon\varphi(\bar{x})$	ε-subdifferential of φ at $\bar{x}$
$\widehat{\partial}^+\varphi(\bar{x})$	upper regular subdifferential of φ at $\bar{x}$
$\overline{\partial}\varphi(\bar{x})$	generalized gradient or convexified/Clarke subdifferential of φ at $\bar{x}$
$\Lambda^0\varphi(\bar{x})$	Warga derivate container of φ at $\bar{x}$
$\partial_P\varphi(\bar{x})$	proximal subdifferential of φ at $\bar{x}$
$\widehat{\partial}_{H(s)}(\bar{x})$	s-Hölder subdifferential of φ at $\bar{x}$
$\widehat{\partial}^+_{H(s)}(\bar{x})$	upper s-Hölder subdifferential of φ at $\bar{x}$
$\partial_{H(s)}(\bar{x})$	limiting s-Hölder subdifferential of φ at $\bar{x}$
$d\varphi(\bar{x})$	Gâteaux derivative of φ at $\bar{x}$
$\varphi'(\bar{x};w)$	directional derivative of φ at $\bar{x}$ in direction w
$d\varphi(\bar{x};w)$	contingent derivative of φ at $\bar{x}$ in direction w
$\varphi^0(\bar{x};w)$	generalized directional derivative of φ at $\bar{x}$ in direction w
$\nabla^2\varphi(\bar{x})$	classical Hessian of φ at $\bar{x}$
$\partial^2\varphi$, $\partial^2_N\varphi$, and $\partial^2_M\varphi$	second-order subdifferentials (generalized Hessians) of φ

Mappings

$f\colon X \to Y$	single-valued mappings from X to Y
$F\colon X \rightrightarrows Y$	set-valued mappings from X to Y
$\mathrm{dom}\,F$	domain of F
$\mathrm{rge}\,F$	range of F
$\mathrm{gph}\,F$	graph of F
$\ker F$	kernel of F
$\|F\|$	norm of positive homogeneous mappings
$F^{-1}\colon Y \rightrightarrows X$	inverse mapping to $F\colon X \rightrightarrows Y$
$F(\Omega)$ and $F^{-1}(\Omega)$	image and inverse image/preimage of Ω under F
$F \circ G$	composition of mappings
$F \stackrel{h}{\circ} G$	h-composition of mappings
$\Delta(\cdot;\Omega)$	set indicator mapping
Ω_ρ	set enlargement mapping
$\mathrm{epi}_{\,\Theta} F$	epigraph of $F\colon X \rightrightarrows Y$ with respect to ordering cone $\Theta \subset Y$
$\mathcal{E}_{F,\Theta}$	epigraphical multifunction for $F\colon X \rightrightarrows Y$ with respect to cone $\Theta \subset Y$
$\mathcal{E}(f,\Theta,\Omega)$	generalized epigraph of $f\colon X \to Y$ with respect to $\Theta \subset Y$ and $\Omega \subset X$
$\nabla f(\bar{x})$	Jacobian or derivative of $f\colon X \to Y$

$DF(\bar{x},\bar{y})$	graphical/contingent derivative of F at $(\bar{x},\bar{y}) \in \text{gph}\, F$
$D^*F(\bar{x},\bar{y})$	(basic) coderivative of F at $(\bar{x},\bar{y}) \in \text{gph}\, F$
$D^*_N F(\bar{x},\bar{y})$	normal coderivative of F at $(\bar{x},\bar{y}) \in \text{gph}\, F$
$D^*_M F(\bar{x},\bar{y})$ and $\widetilde{D}^*_M F(\bar{x},\bar{y})$	mixed and reversed mixed coderivative of F at $(\bar{x},\bar{y})$, respectively
$\widehat{D}^*F(\bar{x},\bar{y})$	precoderivative or regular coderivative of F at $(\bar{x},\bar{y})$
$\widehat{D}^*_\varepsilon F(\bar{x},\bar{y})$	ε-coderivative of F at $(\bar{x},\bar{y})$
$\widehat{\partial}_\Theta F$	regular subdifferential of $F\colon X \rightrightarrows Y$ with respect to ordering cone $\Theta \subset Y$
$\partial_\Theta F$	basic subdifferential of $F\colon X \rightrightarrows Y$ with respect to cone Θ
$\partial^\infty_\Theta F$	singular subdifferential of $F\colon X \rightrightarrows Y$ with respect to cone Θ
$\widehat{D}^*_\Theta f$	regular Θ-coderivative of $f\colon X \to Y$ with respect to ordering cone $\Theta \subset Y$
$D^*_{N,\Theta} f$	sequential normal Θ-coderivative of $f\colon X \to Y$ with respect to cone Θ
$\widetilde{D}^*_{N,\Theta} f$	topological normal Θ-coderivative of $f\colon X \to Y$ with respect to cone Θ
$\breve{D}^*_N f$	cluster normal Θ-coderivative of $f\colon X \to Y$ with respect to cone Θ
$\overline{\partial} f(\bar{x})$	(Clarke) generalized Jacobian of $f\colon \mathbb{R}^n \to \mathbb{R}^m$ at $\bar{x}$

Acronyms

AMVT	approximate mean value theorem
CEL	compactly epi-Lipschitzian (sets)
CEP	conic extremal principle
CHIP	conic hull intersection property
CRCQ	constant rank constraint qualification
CQC	closedness qualification condition
CQs	constraint qualifications
DC	difference of convex (functions, programs)
EMFCQ	extended Mangasarian-Fromovitz constraint qualification
EPEC	equilibrium programming with equilibrium constraints
FMCQ	Farkas-Minkowski constraint qualification
GE	generalized equation
GSIP	generalized semi-infinite programming
KKT	Karush-Kuhn-Tucker (conditions)
LCTV	locally convex topological vector (spaces)
LFM	local Farkas-Minkowski (property)
LICQ	linear independence constraint qualification
l.s.c.	lower semicontinuous (functions)
MFCQ	Mangasarian-Fromovitz constraint qualification
MMA	method of metric approximations
MMFCQ	marginal Mangasarian-Fromovitz constraint qualification
MOEC	multiobjective optimization with equilibrium constraints
MPEC	mathematical programming with equilibrium constraint
NCC	normal closedness condition
NDQ	net demand qualification (condition)
NDSNQ	net demand strong qualification (condition)
NDSQ	net demand strict qualification (condition)
NDWC	net demand weak qualification (condition)
NFMCQ	nonlinear Farkas-Minkowski qualification condition
NLP	nonlinear programming
NQC	normal qualification condition

ODE	ordinary differential equation
PCS	parametric constraint systems
PDE	partial differential equation
PMFCQ	perturbed Mangasarian-Fromovitz constraint qualification
PSNC	partially sequentially normally compact (sets and mappings)
PVS	parametric variational systems
SC	subdifferential closedness condition
SCQ	Slater constraint qualification
SDP	semidefinite programming
SIP	semi-infinite programming
$\mathcal{SL}$	semi-Lipschitzian (sums)
SNC	sequentially normally compact
SNEC	sequentially normally epicompact (sets)
SQC	subdifferential qualification condition
SQP	sequential quadratic programming
SSC	strong Slater condition
RCQ	Robinson constraint qualification
u.s.c.	upper semicontinuous (functions)
VFCQ	value function constraint qualification

Subject Index

B. S. Mordukhovich, *Variational Analysis and Applications*, Springer Monographs in Mathematics, https://doi.org/10.1007/978-3-319-92775-6