5-1-2006

Variational Analysis in Nonsmooth Optimization and Discrete Optimal Control

Boris S. Mordukhovich
Wayne State University, boris@math.wayne.edu

VARIATIONAL ANALYSIS IN NONSMOOTH OPTIMIZATION AND DISCRETE OPTIMAL CONTROL

BORIS S. MORDUKHOVICH

WAYNE STATE UNIVERSITY

Detroit, MI 48202

Department of Mathematics
Research Report

2006 Series
#5

VARIATIONAL ANALYSIS IN NONSMOOTH OPTIMIZATION AND DISCRETE OPTIMAL CONTROL

Boris S. Mordukhovich

Department of Mathematics, Wayne State University, Detroit, Michigan 48202
(boris@math.wayne.edu)

The paper is devoted to applications of modern methods of variational analysis to constrained optimization and control problems generally formulated in infinite-dimensional spaces. The main attention is paid to the study of problems with nonsmooth structures, which require the usage of advanced tools of generalized differentiation. In this way we derive new necessary optimality conditions in optimization problems with functional and operator constraints and then apply them to optimal control problems governed by discrete-time inclusions in infinite dimensions. The principal difference between finite-dimensional and infinite-dimensional frameworks of optimization and control consists of the "lack of compactness" in infinite dimensions, which leads to imposing certain "normal compactness" properties and developing their comprehensive calculus, together with appropriate calculus rules of generalized differentiation. On the other hand, one of the most important achievements of the paper consists of relaxing the latter assumptions for certain classes of optimization and control problems. In particular, we fully avoid the requirements of this type imposed on target endpoint sets in infinite-dimensional optimal control for discrete-time inclusions.

Key words. variational analysis; nonsmooth optimization and optimal control; discrete-time inclusions; generalized differentiation; infinite dimensions; necessary optimality conditions
MSC2000 subject classification: Primary: 49J53, 90C30
OR/MS subject classification: Primary: nondifferentiable programming, optimal control

1 Introduction

Variational analysis has been recognized as a rapidly growing and fruitful area in applied mathematics concerning mainly the study of optimization and equilibrium problems and also applying perturbation ideas and variational principles to a broad class of problems and situations, which may not be of a variational nature. The reader can find more information about basic principles and results of variational analysis in the now classical book by Rockafellar and Wets [20] devoted to variational analysis in finite dimensions and in the recent mutually complementary books by Borwein and Zhu [3] and by Mordukhovich [16, 17] dealing with variational analysis and its applications in both finite-dimensional and infinite-dimensional spaces. It is important to emphasize that basic techniques in variational analysis (particularly those related perturbations and the usage of modern variational principles) unavoidably lead to *nonsmooth structures* of functions, sets, and set-valued mappings, even for problems with initially smooth data. Thus appropriate tools of *generalized differentiation* lie at the very heart of modern variational analysis and its applications.

This paper concerns applications of advanced techniques of variational analysis and generalized differentiation to deriving *necessary optimality conditions* in three important classes of constrained optimization problems formulated generally in *infinite-dimensional* spaces. The *first class* consists of problems in *nondifferentiable programming* with *finitely many functional constraints* described by inequalities and inequalities with real-valued functions, along with *geometric constraints* given by closed subsets of *Asplund* (including any reflexive Banach) spaces. The main results are obtained via several versions of the *extended Lagrange*

principle expressed in terms of our basic normals, subgradients, and coderivatives satisfying comprehensive calculus rules ("full calculus").

The *second class* of optimization problems under consideration contains the so-called *operator constraints* given by mappings with values in infinite-dimensional spaces. Such constraints are typical in many infinite-dimensional optimization and equilibrium problems, particularly those arising in dynamic optimization and optimal control. Problems with operator constraints are significantly different from those with finitely many functional constraints; it is well known that a counterpart of the Lagrange multiplier rule *does not* hold even in smooth settings with no additional assumptions. We derive general results that ensure the fulfillment of first-order necessary optimality conditions in nonsmooth problems with operator constraints by imposing fairly weak assumptions of a "sequential normal compactness" (SNC) type, which are automatic in finite dimensions while being among the most essential ingredients of infinite-dimensional variational analysis. Furthermore, we show that the imposed assumptions always hold for a large class of *generalized Fredholm* mappings, which exhibit a kind of "finite codimension" behavior.

The *third class* of problems under consideration in this paper concerns *dynamic optimization* and deals with *discrete optimal control* in infinite dimensions. More precisely, we study problems of minimizing Mayer-type functionals on trajectories of *discrete-time inclusions*, as well as those arising from *discrete approximations* of optimal control systems governed by constrained differential/evolution inclusions in Banach spaces. Problems of these types can be reduced to mathematical programs in infinite-dimensional spaces with *many geometric constraints* and *operator constraints* of a special Fredholm kind. The results obtained in this direction are heavily based on comprehensive generalized differential and SNC calculi developed in the author's book [16]. At the same time, we are able to *significantly relax* SNC requirements in infinite-dimensional optimal control problems for discrete-time systems. In particular, this allows us to cover optimal control problems with *no* such assumptions imposed on *target/endpoint constraint* sets, e.g., the two-point constraint case that has always been an obstacle in infinite-dimensional optimal control.

The rest of the paper is organized as follows. In Section 2 we briefly overview the basic constructions of generalized differentiation and some of their properties widely used in formulations and proofs of the main results. Section 3 is devoted to the afore-mentioned sequential normal compactness properties and related issues. In particular, in this section we consider generalized Fredholm mappings and prove that they enjoy, in Banach and Asplund space frameworks, a certain underlying version of "partial sequential normal compactness" needed for deriving necessary optimality conditions in problems with operator constraints.

Section 4 concerns optimization problems with functional and geometric constraints in infinite dimensions, while in Section 5 we deal with optimization problems involving operator constraints. Finally, Section 6 is devoted to the study of optimal control problems governed by discrete-time and finite-difference inclusions in infinite-dimensional spaces.

Our notation is basically standard; see [16, 17]. Unless otherwise stated, all the spaces considered are Banach with the norm $\|\cdot\|$ and the canonical dual pairing $\langle\cdot,\cdot\rangle$ between the space in question, say X, and its topological dual X^* whose weak* topology is denoted by w^*. We use the symbols $I\!B$ and $I\!B^*$ to signify the closed unit balls of the space under consideration and its dual, respectively. Given a set-valued mapping $F\colon X \rightrightarrows X^*$, its *sequential Painlevé-Kuratowski upper/outer limit* at $\bar{x}$ is defined by

$$\underset{x\to\bar{x}}{\operatorname{Lim\,sup}} F(x) := \left\{ x^* \in X^* \mid \exists \text{ sequences } x_k \to \bar{x},\ x_k^* \xrightarrow{w^*} x^* \text{ with } x_k^* \in F(x_k) \text{ as } k \in I\!N \right\},$$

where $I\!N := \{1, 2, \ldots\}$. The symbols $x \xrightarrow{\Omega} \bar{x}$ and $x \xrightarrow{\varphi} \bar{x}$ signify, respectively, that $x \to \bar{x}$ with $x \in \Omega$ and that $x \to \bar{x}$ with $\varphi(x) \to \varphi(\bar{x})$ for sets $\Omega \subset X$ and extended-real-valued functions $\varphi \colon X \to \overline{I\!R} := [-\infty, \infty]$. We distinguish between single-valued and set-valued mappings by using the symbols $f \colon X \to Y$ and $F \colon X \rightrightarrows Y$, respectively. Recall that

$$\operatorname{gph} F := \{(x, y) \in X \times Y \mid y \in F(x)\}$$

stands for the *graph* of the mapping $F \colon X \rightrightarrows Y$.

2   Generalized Differentiation

We present here a brief overview of the basic generalized differential constructions in variational analysis and some of their properties widely used in what follows. The material is taken from the author's book [16], where the reader can find a comprehensive theory for these constructions with extensive discussions, references, and commentaries.

Given a nonempty *set* $\Omega \subset X$ and a point $\bar{x} \in \Omega$, the (basic, limiting) *normal cone* to Ω at $\bar{x}$ is defined by

$$N(\bar{x}; \Omega) := \operatorname*{Lim\,sup}_{\substack{x \to \bar{x} \\ \varepsilon \downarrow 0}} \widehat{N}_\varepsilon(x; \Omega), \tag{2.1}$$

where $\widehat{N}_\varepsilon(x; \Omega)$ stands for the set of ε-*normals* ($\varepsilon \geq 0$) to Ω at $x \in X$ given by

$$\widehat{N}_\varepsilon(x; \Omega) := \left\{ x^* \in X^* \;\middle|\; \operatorname*{lim\,sup}_{u \xrightarrow{\Omega} x} \frac{\langle x^*, u - x \rangle}{\|u - x\|} \leq \varepsilon \right\}, \quad x \in \Omega, \tag{2.2}$$

with $\widehat{N}_\varepsilon(x; \Omega) := \emptyset$ if $x \notin \Omega$. If the space X is *Asplund* (i.e., each of its separable subspace has a separable dual) and if the set Ω is locally *closed* around $\bar{x}$, then we can equivalently put $\varepsilon = 0$ in (2.1) and replace $\widehat{N}_\varepsilon$ by the generally smaller *prenormal* (or *Fréchet normal*) *cone* $\widehat{N}(x; \Omega) := \widehat{N}_0(x; \Omega)$ from (2.2) with $\varepsilon = 0$. Observe that the class of Asplund spaces is sufficiently large including every Banach space with a Fréchet differentiable renorm away from the origin (in particular, any reflexive space) and every space with a separable dual; see, e.g., the book by Phelps [19] and the references therein for the well-developed theory of Asplund spaces and some of its applications.

Given a *set-valued mapping* $F \colon X \rightrightarrows Y$ and a point $(\bar{x}, \bar{y}) \in \operatorname{gph} F$, consider two kinds of limiting coderivatives of F at $(\bar{x}, \bar{y})$: the *normal coderivative*

$$D_N^* F(\bar{x}, \bar{y})(\bar{y}^*) := \{x^* \in X^* \mid (x^*, -\bar{y}^*) \in N((\bar{x}, \bar{y}); \operatorname{gph} F)\}, \quad \bar{y}^* \in Y^*, \tag{2.3}$$

and the *mixed coderivative*

$$D_M^* F(\bar{x}, \bar{y})(\bar{y}^*) := \operatorname*{Lim\,sup}_{\substack{(x,y) \to (\bar{x}, \bar{y}) \\ y^* \to \bar{y}^* \\ \varepsilon \downarrow 0}} \widehat{D}_\varepsilon^* F(x, y)(y^*), \quad \bar{y}^* \in Y^*, \tag{2.4}$$

where the ε-*coderivative* $\widehat{D}_\varepsilon^* F(x, y)$ is defined by

$$\widehat{D}_\varepsilon^* F(x, y)(y^*) := \{x^* \in X^* \mid (x^*, -y^*) \in \widehat{N}_\varepsilon((x, y); \operatorname{gph} F)\}, \quad y^* \in Y^*, \quad \varepsilon \geq 0; \tag{2.5}$$

3

where the subindex "ε" is omitted if $\varepsilon = 0$. Note that we can equivalently replace $\widehat{D}_\varepsilon^*$ by $\widehat{D}^*$ in (2.4) and in the corresponding representation of D_N^* due to (2.1) if both spaces X and Y are Asplund and if the graph of F is locally closed around $(\bar{x}, \bar{y})$.

As follows (2.4) and (2.1), the only difference between the normal and mixed coderivatives is that the *norm* convergence of $y^* \to \bar{y}^*$ mixed with the *weak** sequential convergence of $x^* \xrightarrow{w^*} x^*$ are used in (2.4) instead of *both weak** sequential convergences $y^* \xrightarrow{w^*} \bar{y}^*$ and $x^* \xrightarrow{w^*} x^*$ in the limiting representation of D_N^*. Obviously $D_M^* F(\bar{x}, \bar{y})(y^*) \subset D_N^* F(\bar{x}, \bar{y})(y^*)$, where the equality holds if $\dim Y < \infty$, while this inclusion may be *strict* even for single-valued Lipschitzian mappings from $\mathbb{R}$ into Hilbert spaces with certain differentiability properties as in [16, Example 1.35]. In general, the equality

$$D_M^* F(\bar{x}, \bar{y})(y^*) = D_N^* F(\bar{x}, \bar{y})(y^*), \quad y^* \in Y^*,$$

is postulated in [16] as the *strong coderivative normality* of F at $(\bar{x}, \bar{y})$. This property holds for important classes of set-valued and single-valued mappings between infinite-dimensional spaces including convex-graph mappings, the so-called "strictly Lipschitzian" mappings (see below), etc., and it is preserved under various operations; see cases (a)–(i) summarized in [16, Proposition 4.9].

If $F = f \colon X \to Y$ is single-valued and *strictly differentiable* at $\bar{x}$ (which is automatic when f is C^1 around this point), then

$$D_M^* f(\bar{x})(y^*) = D_N^* f(\bar{x})(y^*) = \{\nabla f(\bar{x})^* y^*\}, \quad y^* \in Y^*, \tag{2.6}$$

via the *adjoint* derivative operator $\nabla f(\bar{x})^* \colon Y^* \to X^*$ in (2.6). In [16, 17], the reader can find equivalent analytic representations of both normal and mixed coderivatives and their efficient calculations for various classes of nonsmooth single-valued and set-valued mappings.

Let $\varphi \colon X \to \overline{\mathbb{R}}$ be an *extended-real-valued function* finite at $\bar{x}$. Then

$$\partial \varphi(\bar{x}) := \operatorname*{Lim\,sup}_{\substack{x \xrightarrow{\varphi} \bar{x} \\ \varepsilon \downarrow 0}} \widehat{\partial}_\varepsilon \varphi(x) \tag{2.7}$$

is the (basic, limiting) *subdifferential* of φ at $\bar{x}$, where

$$\widehat{\partial}_\varepsilon \varphi(x) := \left\{ x^* \in X^* \,\middle|\, \frac{\varphi(u) - \varphi(x) - \langle x^*, u - x \rangle}{\|u - x\|} \geq -\varepsilon \right\} \tag{2.8}$$

is the *ε-subdifferential* of φ at x, for each $\varepsilon \geq 0$. When $\varepsilon = 0$, the set $\widehat{\partial}\varphi(x) := \widehat{\partial}_0 \varphi(x)$ in (2.8) is known also as the *presubdifferential*, or the *Fréchet* (regular, viscosity) *subdifferential* of φ at x. If X is Asplund and if the function φ is *lower semicontinuous* (l.s.c.) around $\bar{x}$, the sets $\widehat{\partial}_\varepsilon \varphi(x)$ can be equivalently replaced by $\widehat{\partial}\varphi(x)$ in (2.7). Furthermore, the subdifferential (2.7) admits the geometric description

$$\partial \varphi(\bar{x}) = \left\{ x^* \in X^* \,\middle|\, (x^*, -1) \in N\big((\bar{x}, \varphi(\bar{x})); \operatorname{epi} \varphi\big) \right\}, \tag{2.9}$$

via the normal cone (2.1) to the *epigraph* $\operatorname{epi} \varphi := \{(x, \mu) \in X \times \mathbb{R} \,|\, \mu \geq \varphi(x)\}$ of φ.

On the other hand, the geometrically defined coderivatives (2.3) and (2.4) admit, in the case of single-valued *locally Lipschitzian* mappings $f \colon X \to Y$, the following representations

$$D_M^* f(\bar{x}) = \partial \langle y^*, f \rangle(\bar{x}), \quad D_N^* f(\bar{x})(y^*) = \partial \langle y^*, f \rangle(\bar{x}) \text{ as } y^* \in Y^* \tag{2.10}$$

4

via the basic subdifferential (2.7) of the scalarized function $\langle y^*, f \rangle(x) := \langle y^*, f(x) \rangle$. The first representation in (2.10) holds in arbitrary Banach spaces X and Y, while the second one requires in addition that X is Asplund (Y is still arbitrary Banach) and that f is *strictly Lipschitzian* at $\bar{x}$ in the sense that the sequence

$$\left\{ \frac{f(x_k + t_k v) - f(x_k)}{t_k} \right\}, \quad k \in I\!N,$$

contains a norm convergent subsequence whenever $x_k \to \bar{x}$ and v belongs to some neighborhood of the origin; see [16, Subsection 3.1.3] for characterizations, verifiable sufficient conditions, and applications. Note that the latter property is in fact equivalent to the "compactly Lipschitzian" property introduced by Thibault [21].

3 Sequential Normal Compactness and Fredholm Mappings

We start this section with recalling certain "sequential normal compactness" properties of sets and mappings that are automatic in finite dimensions while playing a crucial role in infinite-dimensional variational analysis and its applications. We refer the reader to the author's books [16, 17] for more details and historical comments and also to the papers by Ioffe [10, 11] and the books by Fattorini [6] and by Li and Yong [14] for related (somewhat different from each other) "finite codimension" properties and their significance in infinite-dimensional optimization and optimal control.

A subset $\Omega \subset X$ of a Banach space is *sequentially normally compact* (SNC) at $\bar{x} \in \Omega$ if for any sequences $\varepsilon_k \downarrow 0$, $x_k \overset{\Omega}{\to} \bar{x}$, and $x_k^* \in \widehat{N}_{\varepsilon_k}(x_k; \Omega)$ one has

$$x_k^* \overset{w^*}{\to} 0 \implies \|x_k^*\| \to 0 \quad \text{as} \quad k \to \infty.$$

In Asplund spaces, we can equivalently put $\varepsilon_k = 0$ in the above definition provided that Ω is locally closed around $\bar{x}$. Besides finite dimensions, the SNC property of Ω at $\bar{x}$ is automatic in any Banach space if the set Ω is "compactly epi-Lipschitzian" (CEL) around $\bar{x}$ in the sense of Borwein and Strójwas [2]. Note that in general the implication CEL$\Rightarrow$SNC is *strict* even for *convex cones* in nonseparable Asplund spaces; see Fabian and Mordukhovich [5] for a comprehensive study of the relationships between the SNC and CEL properties.

A set-valued mapping $F: X \rightrightarrows Y$ between Banach spaces is *SNC* at $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ if its graph enjoys this property at $(\bar{x}, \bar{y})$, which is in fact equivalent to require that for any sequences $(\varepsilon_k, x_k, y_k, x_k^*, y_k^*) \in I\!R_+ \times (\operatorname{gph} F) \times X^* \times Y^*$ satisfying

$$\varepsilon_k \downarrow 0, \quad (x_k, y_k) \to (\bar{x}, \bar{y}), \quad x_k^* \in \widehat{D}^*_{\varepsilon_k} F(x_k, y_k)(y_k^*), \quad \text{and} \quad (x_k^*, y_k^*) \overset{w^*}{\to} (0, 0) \qquad (3.1)$$

we have $\|(x_k^*, y_k^*)\| \to 0$ as $k \to \infty$, where one can equivalently put $\varepsilon_k = 0$ if both spaces X and Y are Asplund and if F is closed-graph around $(\bar{x}, \bar{y})$. A more subtle *partial SNC* (PSNC) property of F at $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ means that

$$\|x_k^*\| \to 0 \quad \text{provided that} \quad \|y_k^*\| \to 0 \quad \text{as} \quad k \to \infty$$

for any sequences $(\varepsilon_k, x_k, y_k, x_k^*, y_k^*) \in I\!R_+ \times (\operatorname{gph} F) \times X^* \times Y^*$ satisfying (3.1), with the similar simplification in the Asplund space setting. The PSNC property is significantly less restrictive than the SNC one and always holds, in particular, for mappings F having Aubin's

5

Lipschitz-like ("pseudo-Lipschitz") property around $(\bar{x}, \bar{y})$, in the sense [1] that there are neighborhoods U of $\bar{x}$ and V of $\bar{y}$ and a number $\ell \geq 0$ such that

$$F(x) \cap V \subset F(u) + \ell \|x - u\| I\!B \quad \text{whenever} \quad x, u \in U. \tag{3.2}$$

When $V = Y$, the latter property reduces to the classical (Hausdorff) Lipschitz continuity of F around $\bar{x}$. Moreover, the *simultaneous fulfillment* of the *PSNC property* of F at $(\bar{x}, \bar{y})$ and the *mixed coderivative condition*

$$D_M^* F(\bar{x}, \bar{y})(0) = \{0\} \tag{3.3}$$

is *necessary and sufficient* for F to be Lipschitz-like around $(\bar{x}, \bar{y})$; see [16, Theorem 4.10].

We refer the reader to [16, 17] for other efficient conditions implying the SNC/PSNC properties for specific classes of set-valued and single-valued mappings and to the well-developed *SNC calculus* ensuring the preservation of such properties under various operations; this seems to be the most important for applications. Note that the proofs of the major rules of SNC calculus in [16] are based on the *extremal principle*, which can be viewed as a local *variational* counterpart of the classical convex separation in nonconvex settings.

Next let us describe a general class of the mappings particularly important in applications to infinite-dimensional optimization and control problems, for which the PSNC property is satisfied. We first recall the following significant modification of the strictly Lipschitzian mappings (see Section 2) that is due to Ngai, Luc and Théra [18]: a mapping $f : X \to Y$ locally Lipschitzian around $\bar{x}$ is called *compactly strictly Lipschitzian* at $\bar{x}$ if, for each sequences $x_k \to \bar{x}$ and $h_k \to 0 \in X$ with $h_k \neq 0$, the sequence

$$\left\{ \frac{f(x_k + h_k) - f(x_k)}{\|h_k\|} \right\}, \quad k \in I\!N,$$

has a norm convergent subsequence.

It is obvious that a compactly strictly Lipschitzian mapping is strictly Lipschitzian at the reference point. Moreover, for $\dim Y < \infty$ the above strict Lipschitzian notions agree and reduce to the standard local Lipschitz continuity. However, it is not the case when the space Y is infinite-dimensional being Asplund, in particular. Indeed, consider the mapping $f : c_0 \to c_0$ given by

$$f(x) := \{\sin x_k\} \quad \text{for} \quad x := \{x_k\},$$

where c_0 stands for the Asplund space of sequences $\{x_k\}$ with $x_k \to 0$ as $k \to \infty$ endowed with the supremum norm. It is not hard to check that the mapping f is strictly Lipschitzian but not compactly strictly Lipschitzian at the origin. Observe that f is compactly strictly Lipschitzian at $\bar{x}$ if it is strictly Fréchet differentiable at x with the *compact* derivative operator, or more generally: if f is a composition $f = g \circ f_0$, where g is strictly differentiable with the compact derivative while f_0 is locally Lipschitzian. Furthermore, the class of compactly strictly Lipschitzian mappings contains every mapping $f : X \to Y$ that is *uniformly directionally compact* around $\bar{x}$ as defined by Ioffe [11] in primal space terms: there is a norm compact set $Q \subset Y$ for which

$$f(x + th) \in f(x) + t\|h\|Q + t\,\eta(\|x - \bar{x}\|, t) I\!B$$

whenever $h \in X$ with $\|h\| \leq 1$ and x close to $\bar{x}$, with $\eta(\varepsilon, t) \to 0$ as $\varepsilon \downarrow 0$ and $t \downarrow 0$.

6

Definition 3.1 (generalized Fredholm mappings). *A mapping $f\colon X \to Y$ is* GENER-
ALIZED FREDHOLM *at $\bar{x}$ if there is $g\colon X \to Y$ compactly strictly Lipschitzian at $\bar{x}$ and such
that the difference $f - g$ is a linear bounded operator whose image is a closed subspace of
finite codimension in Y.*

The generalized Fredholm notion extends the "semi-Fredholm" one from Ioffe [11], where
g is assumed to be uniformly directionally compact; see also Ginsburg and Ioffe [7] for
the "Fredholm" predecessor of the latter definition with more discussions on the previous
developments and applications to optimal control. The following major result establishes
the *partial* sequential normal compactness of generalized Fredholm mappings and their
restrictions to closed SNC sets.

Theorem 3.2 (partial sequential compactness of generalized Fredholm mappings.)
Let $f\colon X \to Y$ be a generalized Fredholm at $\bar{x}$. The following assertions hold:
  (i) *The inverse set-valued mapping $f^{-1}\colon Y \rightrightarrows X$ is PSNC at $\big(f(\bar{x}),\bar{x}\big)$ for arbitrary
Banach spaces X and Y.*
  (ii) *Assume that both X and Y are Asplund, and let*

$$f_\Omega(x) := \begin{cases} f(x) & \text{if } x \in \Omega, \\ \emptyset & \text{if } x \notin \Omega \end{cases}$$

*be the restriction of f to a set $\Omega \subset X$. Then f_Ω^{-1} is PSNC at $\big(f(\bar{x}),\bar{x}\big)$ provided that Ω is
SNC at $\bar{x}$ and locally closed around this point.*

Proof. To prove (i), take sequences $\varepsilon_k \downarrow 0$, $x_k \xrightarrow{\Omega} \bar{x}$, $x_k^* \to 0$, and $y_k^* \xrightarrow{w^*} 0$ such that

$$x_k^* \in \widehat{D}_{\varepsilon_k}^* f(x_k)(y_k^*) \text{ for all } k \in I\!\!N.$$

To justify the PSNC property of f^{-1} at $\big(f(\bar{x}),\bar{x}\big)$, we need to show that $\|y_k^*\| \to 0$ as $k \to \infty$.

Denote by $A := f - g$ the linear bounded operator from X to Y whose image $Y_0 := AX$
is a closed subspace of finite codimension by Definition 3.1. The latter means that there is
a closed subspace $Y_1 \subset Y$ with $Y = Y_0 \bigoplus Y_1$ and $\operatorname{codim} Y_1 < \infty$. Due to the elementary
sum rule for ε-coderivatives in the setting $f = A + g$ under consideration, the desired PSNC
property reduces to justifying the implication

$$\Big[\varepsilon_k \downarrow 0, \ \|x_k^*\| \to 0, \ y_k^* \xrightarrow{w^*} 0\Big] \Longrightarrow \|y_k^*\| \to 0 \text{ as } k \to \infty \tag{3.4}$$

provided the fulfillment of the inclusions

$$x_k^* - A^* y_k^* \in \widehat{D}_{\varepsilon_k}^* g(x_k)(y_k^*) \text{ for all } k \in I\!\!N. \tag{3.5}$$

Since f is *compactly* strictly Lipschitzian at $\bar{x}$, inclusions (3.5) ensure, by the ε-coderivative
modification of the proof of Proposition 3.4 from Ngai, Luc and Théra [18], that $y_k^* \xrightarrow{w^*} 0$
implies the *strong* convergence

$$\|x_k^* - A^* y_k^*\| \to 0 \text{ and hence } \|A^* y_k^*\| \to 0 \text{ as } k \to \infty.$$

By $Y = Y_0 \bigoplus Y_1$ with $Y_0 = AX$, for each $k \in I\!\!N$ we find $y_{0k}^* \in Y_0$ and $y_{1k}^* \in Y_1^*$ such that

$$y_k^* = y_{0k}^* + y_{1k}^* \text{ and } A^* y_k^* = A^* y_{0k}^*.$$

Since $\dim Y_1^* < \infty$ and since A maps X *onto* Y_0, we get $\|y_{1k}^*\| \to 0$ and

$$\|A^* y_{0k}^*\| \geq \mu \|y_{0k}^*\| \quad \text{with some} \quad \mu > 0$$

by the classical *open mapping theorem*. Thus $\|y_{0k}^*\| \to 0$ and $\|y_k^*\| \to 0$ as $k \to \infty$, which justifies (3.4) and completes the proof of assertion (i) of the theorem.

To justify assertion (ii) under the Asplund space and SNC assumptions made, we use the simplified description of the PSNC property of f_Ω at $\bar{x}$, with $\varepsilon_k = 0$ in (3.1), and take sequences (x_k, x_k^*, y_k^*) satisfying

$$x_k^* \in \widehat{D}^*\big(A + g + \Delta(\cdot; \Omega)\big)(x_k)(y_k^*) \quad \text{for all} \quad k \in N,$$

where $\Delta(\cdot; \Omega)$ is the *indicator mapping* of the set Ω relative to the image space Y, which equals $0 \in Y$ of $x \in \Omega$ and $\emptyset$ otherwise. Our aim is to show that for all such (x_k, x_k^*, y_k^*) the implication (3.4) holds with $\varepsilon_k = 0$.

Employing the fuzzy sum rule for coderivatives in Asplund spaces as in [16, Theorem 3.10] and taking into account that

$$\widehat{D}^* \Delta(x; \Omega) = \widehat{N}(x; \Omega) \quad \text{whenever} \quad x \in \Omega,$$

we find $\widehat{x}_k \to \bar{x}$, $u_k \to \bar{x}$, $\widetilde{x}_k^* \to 0$, $\widetilde{y}_k^* \xrightarrow{w^*} 0$, $\widehat{y}_k^* \xrightarrow{w^*} 0$, and $\widehat{x}_k^* \in \widehat{D}^* g(\widehat{x}_k)(\widehat{y}_k^*)$ such that

$$\widetilde{x}_k^* - A^* \widetilde{y}_k^* - \widehat{x}_k^* \in \widehat{N}(u_k; \Omega) \quad \text{and} \quad \|\widetilde{y}_k^* - y_k^*\| \to 0 \quad \text{as} \quad k \to \infty.$$

As in the proof of assertion (i) above, conclude from $\widehat{x}_k^* \in \widehat{D}^* g(\widehat{x}_k)(\widehat{y}_k^*)$ and $\widehat{y}_k^* \xrightarrow{w^*} 0$ that, by the compact strict Lipschitzian property of g at $\bar{x}$, we have $\|\widehat{x}_k^*\| \to 0$ as $k \to \infty$. Furthermore, the strong convergence

$$\|\widetilde{x}_k^* - A^* \widetilde{y}_k^* - \widehat{x}_k^*\| \to 0 \quad \text{as} \quad k \to \infty,$$

follows from the SNC property of Ω at $\bar{x}$. We get therefore that $\|A^* \widetilde{y}_k^*\| \to 0$ and conclude, by the open mapping arguments as in the proof of (a), that

$$\|\widetilde{y}_k^*\| \to 0 \quad \text{as} \quad k \to \infty.$$

This implies the strong convergence $\|y_k^*\| \to 0$ and completes the proof of the theorem. $\triangle$

A "codirectional compactness" counterpart of Theorem 3.2(ii) in general Banach spaces was established by Ioffe [11, Theorem 6], with a different (significantly more involved) proof, for semi-Fredholm mappings f and CEL sets Ω.

4 Extended Lagrange Principle

In this section we study the following problem of *nondifferentiable programming* in infinite-dimensional spaces with finitely many functional constraints of the inequality and equality types, along with general geometric constraints:

$$\begin{cases} \text{minimize} \ \ \varphi_0(x) \ \ \text{subject to} \\[2mm] \varphi_i(x) \leq 0, \quad i = 1, \ldots, m, \\[2mm] \varphi_i(x) = 0, \quad i = m+1, \ldots, m+r, \\[2mm] x \in \Omega \subset X, \end{cases} \tag{4.1}$$

8

where $\varphi_i \colon X \to I\!R$ for $i = 0, \ldots, m + r$. Note that the single geometric constraint $x \in \Omega$ in (4.1) is considered just for convenience and simplicity; the results obtained below can be extended to the case of *finitely many* geometric constraints $x \in \Omega_j$, $j = 1, \ldots, l$, by reducing them to $x \in \Omega := \Omega_1 \cap \ldots \cap \Omega_l$ and employing the intersection formula for basic normals in Asplund spaces used in what follows; cf. [17, Section 5.1].

Define the classical *Lagrangian*

$$L(x, \lambda_0, \ldots, \lambda_{m+r}) := \lambda_0 \varphi_0(x) + \ldots + \lambda_{m+r} \varphi_{m+r}(x) \tag{4.2}$$

built upon the cost function and the functional (but not geometric) constraints.

It has been observed by Vladimir Tikhomirov in the 1960s (see, in particular, his book with Ioffe [12] and the recent one with Brinkhuis [4]) that necessary optimality conditions for various classes of extremal problems (arising, e.g., in nonlinear programming, approximation theory, the calculus of variations, optimal control, classical inequalities, mechanics, etc.) can be obtained via the so-called *Lagrange principle*: by deriving them as necessary conditions for minimizing the Lagrangian (4.2) involving the cost and functional constraint data, subject *only* to the remaining *geometric constraints*.

Of course, the validity of the Lagrange principle must be justified for each class of optimization problems under consideration. It has been done in the mentioned book [12] and related publications for some extremal and control problems with smooth, convex, and mixed "smooth-convex" structures typical in optimal control problems with smooth dynamics. More general nonsmooth developments of the Lagrange principle were later provided by many researchers; see particularly Hiriart-Urruty [8], Ioffe [10], Kruger [13], and Mordukhovich [15] for early results in this direction. We refer the reader to the author's book [17] for various results, discussions, and commentaries on first-order necessary optimality conditions on nonsmooth optimization.

In this section we present extended results of the Lagrange principle type for the general class of nondifferentiable programming problems (4.1) in infinite dimensions. Let us first establish in the following theorem certain *calculus rules* of independent interest that give efficient representations of basic normals (2.1) to the so-called *generalized epigraphs*

$$\mathcal{E}(f, \Theta, \Omega) := \big\{ (x, y) \in X \times Y \,\big|\, f(x) - y \in \Theta,\ x \in \Omega \big\}, \tag{4.3}$$

of the mapping $f \colon X \to Y$ with respect to the "ordering set" $\Theta \subset Y$ and the constraint set $\Omega \subset X$. As usual, $f_\Omega(x) := f(x) + \Delta(x; \Omega)$ signifies the restriction of f to the set Ω.

Theorem 4.1 (relationships for basic normals to generalized epigraphs). *Let $f \colon X \to Y$ be a mapping between Banach spaces, and let $\Omega \subset X$ and $\Theta \subset Y$ be nonempty sets with $\bar{x} \in \Omega$ and $f(\bar{x}) - \bar{y} \in \Theta$. The following assertions hold:*

    (i) *One always has the implication*

$$(x^*, y^*) \in N\big((\bar{x}, \bar{y}); \mathcal{E}(f, \Omega, \Theta)\big) \implies -y^* \in N\big(f(\bar{x}) - \bar{y}; \Theta\big).$$

    (ii) *Assume that X and Y are Asplund, that f is continuous around $\bar{x}$ relative to Ω, and that Ω and Θ are locally closed around $\bar{x}$ and $f(\bar{x}) - \bar{y}$, respectively. Then*

$$N\big((\bar{x}, \bar{y}); \mathcal{E}(f, \Omega, \Theta)\big) \subset \ \big\{ (x^*, y^*) \in X^* \times Y^* \,\big|\, x^* \in D_N^* f_\Omega(\bar{x})(y^*),$$

$$\tag{4.4}$$

$$-y^* \in N\big(f(\bar{x}) - \bar{y}; \Theta\big) \big\},$$

9

where the equality holds if f is locally Lipschitzian around $\bar{x}$ relative to Ω and if f_Ω is strongly coderivatively normal at this point.

(iii) *Assume that f is locally Lipschitzian around $\bar{x}$ relative to Ω in the Banach space framework. Then*

$$N\big((\bar{x},\bar{y});\mathcal{E}(f,\Omega,\Theta)\big) \supset \Big\{(x^*,y^*)\in X^*\times Y^* \,\Big|\, x^*\in\partial\langle y^*,f_\Omega\rangle(\bar{x}),$$

$$-y^*\in N\big(f(\bar{x})-\bar{y};\Theta\big)\Big\},\tag{4.5}$$

where the equality holds if in addition X and Y are Asplund, if Ω and Θ are locally closed around $\bar{x}$ and $f(\bar{x})-\bar{y}$, respectively, and if f_Ω is strongly coderivatively normal at $\bar{x}$.

Proof. Assertion (i) follows directly from definition (2.1) of basic normals in Banach spaces due to the structure of the set $\mathcal{E}(f,\Omega,\Theta)$ in (4.3). The proof of (ii) is more involved. First observe that $\mathcal{E}(f,\Omega,\Theta)$ admits the *inverse image* representation

$$\mathcal{E}(f,\Omega,\Theta) = g^{-1}(\Theta) \quad\text{with}\quad g(x,y):=f_\Omega(x)-y.\tag{4.6}$$

Thus we can apply [16, Theorem 3.10] ensuring an upper estimate of $N\big((\bar{x},\bar{y});g^{-1}(\Theta)\big)$ under the Asplund and closedness/continuity assumptions imposed in (ii) *provided that*

$$\ker\widetilde{D}^*_M g(\bar{x},\bar{y}) = \{0\} \quad\text{and}\quad g^{-1}\ \text{is PSNC at}\ \big(f(\bar{x})-\bar{y}),\bar{x},\bar{y}\big),\tag{4.7}$$

where $\widetilde{D}^*_M g(\bar{x},\bar{y})$ stands for the *reversed mixed coderivative* of g at $(\bar{x},\bar{y})$ defined in this case by the relation

$$\widetilde{D}^*_M g(\bar{x},\bar{y})(z^*) := \Big\{(x^*,y^*)\in X^*\times Y^* \,\Big|\, z^*\in -D^*_M g^{-1}\big(f(\bar{x})-\bar{y},\bar{x},\bar{y}\big)(-z^*)\Big\}\tag{4.8}$$

via the mixed coderivative (2.4) of the inverse mapping; see [16] for more details. Observe that the upper estimate of [16, Theorem 3.10] directly implies (4.4) due to the equivalence

$$(x^*,y^*)\in D^*_N g(\bar{x},\bar{y})(z^*) \iff x^*\in D^*_N f_\Omega(\bar{x})(z^*),\ y^*=-z^*,$$

which easily follows from the normal coderivative sum rule of [16, Theorem 1.62].

Let us now show that both conditions in (4.7) *are automatically satisfied* due the special structure of g in (4.6). First check the kernel condition in (4.7). To proceed, pick z^* with $0\in\widetilde{D}^*_M g(\bar{x},\bar{y})(z^*)$ and find $(x_k,y_k)\to(\bar{x},\bar{y})$ and $(u^*_k,v^*_k)\in\widehat{D}^*g(x_k,y_k)(z^*_k)$ such that

$$x_k\in\Omega,\quad \|(u^*_k,v^*_k)\|\to 0,\quad\text{and}\quad z^*_k\xrightarrow{w^*} z^*\ \text{as}\ k\to\infty.$$

Using the elementary sum rule for the Fréchet coderivative of g in (4.6), we get

$$\widehat{D}^*g(x_k,y_k)(z^*_k) = \Big(\widehat{D}^*f_\Omega(x_k)(z^*_k),0\Big) + (0,-z^*_k),$$

which implies the relationships

$$u^*_k\in\widehat{D}^*f_\Omega(x_k)(z^*_k),\quad v^*_k=-z^*_k,\quad\text{and hence}\ \|z^*_k\|\to 0=z^*.$$

This means that $\ker\widetilde{D}^*_M g(\bar{x},\bar{y}) = \{0\}$. To check further the PSNC condition in (4.7), we take $x_k\in\Omega$, $y_k\in Y$, and $(u^*_k,v^*_k,z^*_k)\in\widehat{N}\big((x_k,y_k,f(x_k)-y_k);\text{gph}\,g\big)$ with

$$(x_k,y_k,f(x_k)-y_k)\to(\bar{x},\bar{y},f(\bar{x})-\bar{y}),\quad \|(u^*_k,v^*_k)\|\to 0,\quad\text{and}\ z^*_k\xrightarrow{w^*} 0$$

10

and, arguing as above, get $\|z_k^*\| \to 0$ as $k \to \infty$. This precisely means, in the Asplund space setting, that g^{-1} is PSNC at $\big(f(\bar{x}) - \bar{y}, \bar{x}, \bar{y}\big)$, which therefore justifies (4.4).

Next we prove inclusion (4.5) in (iii) and then finally establish the equalities in both assertions (ii) and (iii) under the additional assumptions made. To proceed with (4.5), pick any $y^* \in -N(f(\bar{x}) - \bar{y}; \Theta)$ and $x^* \in \partial\langle y^*, f_\Omega\rangle(\bar{x})$. By definitions (2.1) and (2.7) of basic normals and subgradients in Banach spaces, find sequences $\varepsilon_{1k} \downarrow 0$, $\varepsilon_{2k} \downarrow 0$, $x_k \overset{\Omega}{\to} \bar{x}$, $y_k \to \bar{y}$, $x_k^* \overset{w^*}{\to} x^*$, and $y_k^* \overset{w^*}{\to} y^*$ as $k \to \infty$ such that

$$x_k^* \in \widehat{\partial}_{\varepsilon_{1k}}\langle y_k^*, f_\Omega\rangle(x_k) \quad \text{and} \quad -y_k^* \in \widehat{N}_{\varepsilon_{2k}}\big(f(x_k) - y_k; \Theta\big) \quad \text{for all} \ \ k \in I\!N.$$

It is easy to observe from definitions (2.2) and (2.8) of the corresponding ε-elements and from the assumed Lipschitz continuity of f_Ω around $\bar{x}$ with constant ℓ that

$$(x_k^*, y_k^*) \in \widehat{N}_{\varepsilon_k}\big((x_k, y_k); \mathcal{E}(f, \Omega, \Theta)\big) \quad \text{with} \ \ \varepsilon_k := \varepsilon_{1k} + (\ell + 1)\varepsilon_{2k} \downarrow 0 \ \ \text{as} \ \ k \to \infty.$$

This implies, by passing to the limit as $k \to \infty$, that $(x^*, y^*) \in N\big((\bar{x}, \bar{y}); \mathcal{E}(f, \Omega, \Theta)\big)$, which justifies inclusion (4.5) in the general Banach space setting.

To establish the opposite inclusion in (4.5), and hence the equality statement in (iii), we invoke the *scalarization formula*

$$D_M^* f_\Omega(\bar{x})(y^*) = \partial\langle y^*, f_\Omega\rangle(\bar{x}) \quad \text{whenever} \ \ y^* \in Y^* \tag{4.9}$$

for the *mixed coderivative*, which is proved in [16, Theorem 1.90] for the general class of single-valued and locally Lipschitzian mappings on Banach spaces. Strictly speaking, the result of this theorem is formulated for locally Lipschitzian mappings relative to the whole space, i.e., with $\Omega = X$, but its proof holds with no change in the case of arbitrary closed sets considered in (4.9). Furthermore, we have

$$D_N^* f_\Omega(\bar{x})(y^*) = D_M^* f_\Omega(\bar{x})(y^*), \quad y^* \in Y^*,$$

under the coderivative normality requirement on f_Ω at $\bar{x}$. Thus the opposite inclusion in (4.5) follows from (4.4) under all the assumptions made in (iii). Moreover, in this case the right-hand sides in (4.5) and (4.4) are the same, and we simultaneously get the equality statement in (ii) and thus complete the proof of the theorem. $\qquad\qquad\qquad\triangle$

Observe that for our optimization problem (4.1) under consideration, the set

$$\mathcal{E}(\varphi_0, \ldots, \varphi_{m+r}, \Omega) := \Big\{ (x, \alpha_0, \ldots, \alpha_{m+r}) \in X \times I\!R^{m+r+1} \Big| \ x \in \Omega, \ \varphi_i(x) \le \alpha_i,$$

$$i = 0, \ldots, m; \ \varphi_i(x) = \alpha_i, \ i = m+1, \ldots, m+r \Big\}$$

corresponds to (4.3) with an arbitrary constraint set $\Omega \subset X$ and the other data given by $f = (\varphi_0, \ldots, \varphi_{m+r}) \colon X \to I\!R^{m+r+1}$ and $\Theta = I\!R_-^{m+1} \times \{0\} \subset I\!R^{m+r+1}$. We use the calculus results of Theorem 4.1 for the following necessary optimality conditions in (4.1) given in certain extended forms of the *Lagrange principle*, the proof of which is based on the *extremal principle* of variational analysis; see below. Along with the classical Lagrangian (4.2), consider the *extended Lagrangian*

$$L_\Omega(x; \lambda_0, \ldots, \lambda_{m+r}) := \lambda_0\varphi_0(x) + \ldots + \lambda_{m+r}\varphi_{m+r}(x) + \delta(x; \Omega) \tag{4.10}$$

involving also geometric constraints via the set *indicator function* $\delta(x; \Omega) := 0$ for $x \in \Omega$ and $\delta(x; \Omega) := \infty$ otherwise.

Theorem 4.2 (Lagrange principle for nondifferentiable programming). *Let $\bar{x}$ be a local optimal solution to problem (4.1), where the space X is Asplund and where the set Ω is locally closed around $\bar{x}$. The following assertions hold:*

(i) *Assume that the functions φ_i are lower semicontinuous around $\bar{x}$ relative to Ω for $i = 0, \ldots, m$ and continuous around $\bar{x}$ relative to Ω for $i = m+1, \ldots, m+r$. Then there are Lagrange multipliers $(\lambda_0, \ldots, \lambda_{m+r}) \in \mathbb{R}^{m+r+1}$, not all zero, such that*

$$(0, -\lambda_0, \ldots, -\lambda_{m+r}) \in N\big((\bar{x}, 0); \mathcal{E}(\varphi_0, \ldots, \varphi_{m+r}, \Omega)\big), \tag{4.11}$$

which implies the sign and complementary slackness conditions

$$\lambda_i \geq 0 \ \ for \ \ i = 0, \ldots, m, \quad \lambda_i \varphi_i(\bar{x}) = 0 \ \ for \ \ i = 1, \ldots, m, \tag{4.12}$$

If in addition the functions φ_i, $i = 0, \ldots, m$, are also continuous around $\bar{x}$ relative to Ω, then (4.11) yields the coderivative inclusion

$$0 \in D_N^*\big((\varphi_0, \ldots, \varphi_{m+r}) + \Delta(\cdot; \Omega)\big)(\bar{x})(\lambda_0, \ldots, \lambda_{m+r}). \tag{4.13}$$

(ii) *Assume that all the functions φ_i, $i = 0, \ldots, m+r$, are Lipschitz continuous around $\bar{x}$ relative to the set Ω. Then there are Lagrange multipliers $(\lambda_0, \ldots, \lambda_{m+r}) \neq 0$ satisfying (4.12) and the inclusion*

$$0 \in \partial L_\Omega(\cdot, \lambda_0, \ldots, \lambda_{m+r})(\bar{x}) \tag{4.14}$$

in terms of the basic subdifferential of the extended Lagrangian (4.10), which implies

$$0 \in \partial L(\cdot, \lambda_0, \ldots, \lambda_{m+r})(\bar{x}) + N(\bar{x}; \Omega) \tag{4.15}$$

via the classical Lagrangian (4.2). Moreover, in this case conditions (4.13) and (4.14) are equivalent, while (4.11) is equivalent to the simultaneous fulfillment of (4.13) and (4.14).

Proof. Suppose for simplicity that $\varphi_0(\bar{x}) = 0$. Since $\bar{x}$ is a local optimal solution to (4.1), there is a neighborhood U of $\bar{x}$ such that $\bar{x}$ provides the minimum to φ_0 over $x \in U$ subject to the constraints in (4.1). Consider the sets

$$\Omega_1 := \mathcal{E}(\varphi_0, \ldots, \varphi_{m+r}, \Omega) \ \text{ and } \ \Omega_2 := \text{cl}\, U \times \{0\}$$

in the Asplund space $X \times \mathbb{R}^{m+r+1}$ that are locally closed around $(\bar{x}, 0)$ due to the assumptions made on φ_i and Ω. Let us check that $(\bar{x}, 0)$ is an *extremal point* for the set system $\{\Omega_1, \Omega_2\}$ in the sense of [16, Definition 2.1], i.e., $(\bar{x}, 0) \in \Omega_1 \cap \Omega_2$ (which is obvious) and there is a sequence $\{a_k\} \subset X \times \mathbb{R}^{m+r+1}$ with $a_k \to 0$ as $k \to \infty$ such that

$$\big(\Omega_1 - a_k\big) \cap \Omega_2 = \emptyset \ \text{ for all } \ k \in \mathbb{N}.$$

Indeed, the latter holds by the local optimality of $\bar{x}$ in (4.1) for $a_k := (0, \nu_k, 0, \ldots, 0)$, where $\nu_k \uparrow 0 = \varphi_0(\bar{x})$ as $k \to \infty$. Note that Ω_2 is *SNC* at $(\bar{x}, 0)$, since $\bar{x} \in \text{int}\, U$ and the second set in the product therein is finite-dimensional. Applying now the *exact extremal principle* from [16, Theorem 2.22] to the system $\{\Omega_1, \Omega_2\}$ and taking into account the structure of Ω_2, we directly arrive at (4.11) with $(\lambda_0, \ldots, \lambda_{m+r}) \neq 0$. The sign and complementary slackness conditions in (4.12) follow from (4.11) by assertion (i) of Theorem 4.1 above due to the structure of $\Omega_1 = \mathcal{E}(\varphi_0, \ldots, \varphi_{m+r}, \Omega)$ corresponding to $\Theta = \mathbb{R}_-^{m+1} \times \{0\} \subset \mathbb{R}^{m+r+1}$

in (4.3). Furthermore, the coderivative condition (4.13) follows from inclusion (4.4) in assertion (ii) of Theorem 4.1 provided that the functions φ_i, $i = m + 1, \ldots, m + r$, are continuous around $\bar{x}$. This proves assertion (i) of the theorem.

To justify (ii), observe that the mapping

$$f_\Omega(x) := \big(\varphi_0(x), \ldots, \varphi_{m+r}(x)\big) + \Delta(x; \Omega)$$

is strongly coderivatively normal at $\bar{x}$, since its image space in finite-dimensional. Thus the extended Lagrangian inclusion (4.14) and the equivalence statements in (ii) follow from assertions (ii) and (iii) of Theorem 4.1, respectively. Finally, inclusion (4.15) follows from (4.14) by the semi-Lipschitzian sum rule of [16, Theorem 2.33(c)] applied to the extended Lagrangian (4.10) due to $\partial\delta(\bar{x}; \Omega) = N(\bar{x}; \Omega)$. $\triangle$

The following consequence of the Lagrange principle from Theorem 4.2 provides *separated* necessary conditions for local minimizers to problem (4.1) that are expressed, in contrast to the "condensed" Lagrangian forms (4.14) and (4.15), in terms of the corresponding subdifferentials of *each* function φ_i, $i = 0, \ldots, m + r$, separately.

Corollary 4.3 (separated necessary optimality conditions in nondifferentiable programming). *Let all the assumptions from assertion* (ii) *of Theorem 4.2 be fulfilled. Then there are nonnegative multipliers* $(\lambda_0, \ldots, \lambda_{m+r}) \neq 0$ *satisfying the complementary slackness conditions in* (4.12) *and such that*

$$0 \in \sum_{i=0}^{m} \lambda_i \partial\varphi_i(\bar{x}) + \sum_{i=m+1}^{m+r} \lambda_i\big[\partial\varphi_i(\bar{x}) \cup \partial(-\varphi_i)(\bar{x})\big] + N(\bar{x}; \Omega). \tag{4.16}$$

Proof. Inclusion (4.16) follows from (4.15) by the afore-mentioned sum rule from [16, Theorem 2.33(c)] due to the relationships

$$\partial(\lambda\varphi)(\bar{x}) = \lambda\partial\varphi(\bar{x}) \text{ for } \lambda \geq 0 \text{ and } \partial(\lambda\varphi)(\bar{x}) \subset |\lambda|\big[\partial\varphi(\bar{x}) \cup \partial(-\varphi)(\bar{x})\big] \text{ for any } \lambda \in \mathbb{R}.$$

Thus the results of the corollary follow from assertion (ii) of Theorem 4.2. $\triangle$

Note that *all* the multipliers λ_i in Corollary 4.3, including those corresponding to the equality constraints, are *nonnegative* that is different from conventional versions of the Lagrange multiplier rule, even for smooth functions. On the other hand, the *even subgradient set* $\partial\varphi(\bar{x}) \cup \partial(-\varphi)(\bar{x})$ does not reduce to the standard gradient when φ is smooth. Observe also that the even subgradient sets $\lambda_i\big[\partial\varphi_i(\bar{x}) \cup \partial(-\varphi_i)(\bar{x})\big]$ with $\lambda_i \geq 0$ in (4.16) can be replaced by the generally larger sets

$$\lambda_i\partial^0\varphi_i(\bar{x}) \text{ with } \lambda_i \in \mathbb{R} \text{ as } i = m + 1, \ldots, m + r$$

via the *symmetric subdifferential* of φ at $\bar{x}$ defined by

$$\partial^0\varphi(\bar{x}) := \partial\varphi(\bar{x}) \cup \big(-\partial(-\varphi)(\bar{x})\big),$$

which is a *two-sided* extension of the classical gradient to nonsmooth functions.

The next consequence of Theorem 4.2 gives a result of the *abstract minimim/maximum principle* type, which directly follows from the extended Lagrange principle under the *convexity* requirement on the geometric constraint set.

13

Corollary 4.4 (abstract minimum principle in nondifferentiable programming).
Let $\bar{x}$ be a local minimizer for problem (4.1). Suppose in addition to the assumptions of Theorem 4.2(ii) that the set Ω is convex. Then there are multipliers $(\lambda_0, \ldots, \lambda_{m+r}) \neq 0$, satisfying the sign and complementary slackness conditions in (4.12), and a basic subgradient of the classical Lagrangian

$$x^* \in \partial L(\cdot, \lambda_0, \ldots, \lambda_{m+r})(\bar{x})$$

such that the following minimum condition holds:

$$\langle x^*, \bar{x} \rangle = \min \{ \langle x^*, x \rangle \mid x \in \Omega \}.$$

Proof. This follows from condition (4.15) of Theorem 4.2(ii), since the basic normal cone (4.1) reduces, in the case of convex sets, to the classical normal cone of convex analysis

$$N(\bar{x}; \Omega) = \{ x^* \in X^* \mid \langle x^*, x - \bar{x} \rangle \leq 0 \ \text{ for all } \ x \in \Omega \}$$

intrinsically having an *extremal structure*. $\triangle$

5 Optimization Problems with Operator Constraints

In this section we study optimization problems that, in contrast to the nondifferentiable programming problem (4.1) from Section 4, have *infinitely many* equality constraints given by mappings/operators with values in infinite-dimensional spaces. Such constraints are known as *operator constraints*; they are typical, e.g., in optimal control problems governed by ordinary, or delay, or partial differential equations. The problem under consideration in this section is described as follows:

$$\text{minimize } \varphi_0(\bar{x}) \text{ subject to } \varphi_i(x) \leq 0, \ i = 1, \ldots, m, \ f(x) = 0, \ x \in \Omega, \qquad (5.1)$$

where $\varphi \colon X \to \mathbb{R}$, $\Omega \subset X$, and $f \colon X \to Y$ is a mapping between Banach spaces. This problem is formally more general that (4.1), while the results obtained and the methods employed in this section are different from those in Section 4.

Given a point $\bar{x}$ feasible to problem (5.1), define the the index set for *active inequality constraints* at $\bar{x}$ by

$$I(\bar{x}) := \{ i \in \{1, \ldots, m\} \mid \varphi_i(\bar{x}) = 0 \}.$$

Theorem 5.1 (necessary optimality conditions in problems with operator constraints). *Let $\bar{x}$ be local minimizer for (5.1), where the space X is Asplund, where the set Ω is locally closed around $\bar{x}$, and where the functions φ_i are locally Lipschitzian around $\bar{x}$ for $i \in I(\bar{x}) \cup \{0\}$ and upper semicontinuous at $\bar{x}$ for $i \in \{1, \ldots, m\} \setminus I(\bar{x})$. Assume that the mapping $f \colon X \to Y$ with values in Asplund spaces is Lipschitz continuous around $\bar{x}$ and the restriction f_Ω^{-1} is PSNC at $(f(\bar{x}), \bar{x})$. Then there are Lagrange multipliers $\lambda_i \geq 0$ for $i \in I(\bar{x}) \cup \{0\}$ and a linear functional $y^* \in Y^*$, not equal to zero simultaneously, such that*

$$0 \ \in \partial \Big(\sum_{i \in I(\bar{x}) \cup \{0\}} \lambda_i \varphi_i \Big)(\bar{x}) + D_N^* f(\bar{x})(y^*) + N(\bar{x}; \Omega)$$

$$\subset \sum_{i \in I(\bar{x}) \cup \{0\}} \lambda_i \partial \varphi_i(\bar{x}) + D_N^* f(\bar{x})(y^*) + N(\bar{x}; \Omega).$$

$$(5.2)$$

14

If in addition f is strictly Lipschitzian at $\bar{x}$, then

$$0 \in \partial\Big(\sum_{i\in I(\bar{x})\cup\{0\}} \lambda_i\varphi_i\Big)(\bar{x}) + \partial\langle y^*, f\rangle(\bar{x}) + N(\bar{x};\Omega)$$

$$\subset \sum_{i\in I(\bar{x})\cup\{0\}} \lambda_i\partial\varphi_i(\bar{x}) + \partial\langle y^*, f\rangle(\bar{x}) + N(\bar{x};\Omega). \tag{5.3}$$

Proof. Suppose first that the mapping f is *metrically regular* at $\bar{x}$ relative to Ω in the sense that there is $\mu > 0$ and a neighborhood U of $\bar{x}$ such that

$$\mathrm{dist}(x; S) \le \mu\,\|f(x) - f(\bar{x})\| \quad \text{for all} \quad x \in U \cap \Omega,$$

where $S := \{x \in \Omega|\ f(x) = f(\bar{x})\}$. Then, by Ioffe's *exact penalization* theorem from [9] (see [17, Theorem 5.16]), $\bar{x}$ is a local optimal solution to the the *unconstrained minimization problem* for all $\mu > 0$ sufficiently large:

$$\text{minimize} \quad \max\Big\{\varphi_0(x) - \varphi_0(\bar{x}),\ \max_{i\in I(\bar{x})} \varphi_i(x)\Big\} + \mu\big(\|f(x)\| + \mathrm{dist}(x;\Omega)\big). \tag{5.4}$$

Thus, by the generalized Fermat rule in (5.4, we have

$$0 \in \partial\Big(\max\big\{\varphi_0(\cdot) - \varphi_0(\bar{x}),\ \max_{i\in I(\bar{x})} \varphi_i(\cdot)\big\} + \mu\big(\|f(\cdot)\| + \mathrm{dist}(\cdot;\Omega)\big)\Big)(\bar{x}).$$

To transform the latter "condensed" subdifferential condition into the efficient form of the theorem under the assumptions made, we employ the sum, maximum, and chain rules for basic subgradients of functions in Asplund spaces taken from [16, Subsection 3.2.1] and the representation of basic subgradients of the distance function

$$\bigcup_{\lambda\geq 0} \partial\mathrm{dist}(\bar{x};\Omega) = N(\bar{x};\Omega), \quad \bar{x} \in \Omega,$$

valid in arbitrary Banach spaces; see [16, Theorem 1.97]. Thus we arrive at (5.2) with

$$\lambda_i \ge 0 \quad \text{and} \quad (\lambda_i) \ne 0 \quad \text{as} \quad i \in I(\bar{x}) \cup \{0\}.$$

The inclusions in (5.3) follow from (5.2) due to the scalarization formula for the normal coderivative of strictly Lipschitzian mappings given in (2.10); see [16, Theorem 3.28].

Suppose now that f is *not* metrically regular at $\bar{x}$ relative to Ω. One can easily check that in this case the mapping

$$f_\Omega(x) = f(x) + \Delta(x;\Omega), \quad x \in \Omega,$$

is not metrically regular *around* $\bar{x}$ in the conventional sense; see [16, Definition 1.47]. Thus, by the *characterization* of the latter notion from [16, Theorem 4.18], we have that

$$\text{either} \quad \ker \widetilde{D}_M^* f_\Omega(\bar{x}) \ne \{0\} \quad \text{or} \quad f_\Omega^{-1} \text{ is not PSNC at } (f(\bar{x}), \bar{x}),$$

where the reversed mixed coderivative is defined via the mixed one (2.4) in the way of (4.8). Since the latter is excluded by the assumption of the theorem, we find $y^* \ne 0$ satisfying $0 \in \widetilde{D}_M^* f_\Omega(\bar{x})(y^*)$ and thus

$$0 \in D_N^* f_\Omega(\bar{x})(y^*) = D_N^*\big(f + \Delta(\cdot;\Omega)\big)(\bar{x})(y^*) \subset D_N^* f(\bar{x})(y^*) + N(\bar{x};\Omega); \tag{5.5}$$

15

the set inclusion in (5.5) is by the coderivative sum rule from [16, Proposition 3.12], which holds due to the local Lipschitz continuity of f around $\bar{x}$. By (5.5), we conclude that (5.2) is fulfilled in this case with $y^* \neq 0$. As above, this implies (5.3) provided that f is strictly Lipschitzian at $\bar{x}$ and thus completes the proof of the theorem. $\triangle$

It is obvious that the PSNC assumption on f_Ω in the theorem automatically holds when Y is finite-dimensional. In this case Theorem 5.1 actually reduces to assertion (iii) of Theorem 4.2. The next corollary, ensuring the latter property in infinite dimensions, provides efficient conditions for the fulfillment of the optimality results of Theorem 5.1.

Corollary 5.2 (optimality conditions for problems with operator constraints of the generalized Fredholm type.) *Let $\bar{x}$ be a local solution to problem (5.1). Suppose that, in addition to the assumptions imposed on the functions φ_i, $i = 0, \ldots, m + r$, and on the spaces X and Y in Theorem 5.1, the set Ω is locally closed around $\bar{x}$ and SNC at this point, and the mapping $f: X \to Y$ is generalized Fredholm at $\bar{x}$. Then there are Lagrange multipliers $\lambda_i \geq 0$ for $i \in I(\bar{x}) \cup \{0\}$ and a linear functional $y^* \in Y^*$, not all zero, such that conditions (5.3) are satisfied.*

Proof. First observe from Definition 3.1 of generalized Fredholm mappings that f is strictly Lipschitzian at $\bar{x}$. Furthermore, Theorem 3.2(ii) ensures that f_Ω^{-1} is PSNC at $\big(f(\bar{x}), \bar{x}\big)$. Thus we arrive at all the requirements and conclusions of Theorem 5.1 in the case of operator constraints given by strictly Lipschitzian mappings. $\triangle$

6 Discrete Optimal Control

In the concluding section of this paper, we study *dynamic optimization* problems governed by *discrete-time* inclusions and also by those arising from discrete approximations of continuous-time problems of optimal control.

Let us first consider the general *Mayer problem* of *discrete optimal control* with endpoint constraints of inequality, equality, and geometric types in infinite-dimensional spaces:

$$
\left\{
\begin{array}{l}
\text{minimize } \varphi_0(x_0, x_K) \text{ subject to} \\[2mm]
x_{j+1} \in F_j(x_j) \text{ for } j = 0, \ldots, K - 1, \\[2mm]
\varphi_i(x_0, x_K) \leq 0 \text{ for } i = 1, \ldots, m, \\[2mm]
\varphi_i(x_0, x_K) = 0, \text{ for } i = m + 1, \ldots, m + r, \\[2mm]
(x_0, x_K) \in \Omega \subset X^2,
\end{array}
\right.
\tag{6.1}
$$

where $F_j: X \rightrightarrows X$, $\varphi_i: X^2 \to \mathbb{R}$, and $K \in \mathbb{N}$. Observe that the *discrete inclusion* model in (6.1) covers more conventional discrete control systems of the *parameterized type*

$$
x_{j+1} = f_j(x_j, u_j), \quad u_j \in U_j \text{ as } j = 0, \ldots, K - 1
$$

with *explicit* control variables u_j taking values in some admissible control regions U_j.

The next theorem provides necessary optimality conditions of the *discrete Euler-Lagrange type* for problem (6.1), where we impose the Lipschitzian assumptions on φ_i for simplicity.

16

Theorem 6.1 (extended Euler-Lagrange conditions for discrete optimal control).
Let $\{\bar{x}_j|\ j = 0, \ldots, K\}$ be a local optimal solution to the optimal control problem (6.1). Assume that the space X is Asplund, that the functions φ_i are locally Lipschitzian around $(\bar{x}_0, \bar{x}_K)$ for all $i = 0, \ldots, m+r$ while the set Ω is locally closed around this point, and that the graphs of F_j are locally closed around $(\bar{x}_j, \bar{x}_{j+1})$ for every $j = 0, \ldots, K-1$. Suppose also that all but one of the sets Ω and $\operatorname{gph} F_j$, $j = 0, \ldots, K-1$, are SNC at the points $(\bar{x}_0, \bar{x}_K)$ and $(\bar{x}_j, \bar{x}_{j+1})$, respectively. Then there are multipliers $(\lambda_0, \ldots, \lambda_{m+r})$ and an adjoint discrete trajectory $\{p_j \in X^*|\ j = 0, \ldots, K\}$ satisfying the relationships:

—*the Euler-Lagrange inclusion*

$$-p_j \in D_N^* F_j(\bar{x}_j, \bar{x}_{j+1})(-p_{j+1}) \ \text{ for } \ j = 0, \ldots, K-1, \tag{6.2}$$

—*the transversality inclusion*

$$(p_0, -p_K) \in \partial\Big(\sum_{i=0}^{m+r} \lambda_i \varphi_i \Big)(\bar{x}_0, \bar{x}_K) + N\big((\bar{x}_0, \bar{x}_K); \Omega\big), \tag{6.3}$$

—*the sign and complementary slackness conditions*

$$\lambda_i \geq 0 \ \text{ for } \ i = 0, \ldots, m, \quad \lambda_i \varphi_i(\bar{x}_0, \bar{x}_K) = 0 \ \text{ for } \ i = 1, \ldots, m, \tag{6.4}$$

—*and the nontriviality condition*

$$(\lambda_0, \ldots, \lambda_{m+r}, p_0, \ldots, p_K) \neq 0. \tag{6.5}$$

If in addition the set-valued mappings F_j are Lipschitz-like around $(\bar{x}_j, \bar{x}_{j+1})$ and strongly coderivatively normal at these points for all $j = 0, \ldots, K-1$, then we have the enhanced nontriviality condition

$$(\lambda_0, \ldots, \lambda_{m+r}, p_K) \neq 0. \tag{6.6}$$

Proof. Introduce the new variable

$$z := (x_0, \ldots, x_K) \in Z := X^{K+1}$$

and consider the nondifferentiable programming problem of type (4.1):

$$\begin{cases} \text{minimize } \phi_0(z) := \varphi_0(x_0, x_K) \text{ subject to} \\[2mm] \phi_i(z) := \varphi_i(x_0, x_K) \leq 0 \text{ for } i = 1, \ldots, m, \\[2mm] \phi_i(z) := \varphi_i(x_0, x_K) = 0 \text{ for } i = m+1, \ldots, m+r, \\[2mm] z \in \Lambda := \bigcap_{j=0}^{K} \Lambda_j \subset Z, \end{cases} \tag{6.7}$$

where the latter geometric constraints are defined by

$$\Lambda_j := \big\{ z \in Z \big|\ (x_j, x_{j+1}) \in \operatorname{gph} F_j \big\}, \ j = 0, \ldots, K-1,$$

$$\Lambda_K := \big\{ z \in Z \big|\ (x_0, x_K) \in \Omega \big\}. \tag{6.8}$$

17

Applying assertion (ii) of Theorem 4.2 to the optimal solution $\bar{z} := (\bar{x}_0, \ldots, \bar{x}_K)$ for problem (6.7), we find multipliers $(\lambda_0, \ldots, \lambda_{m+r}) \neq 0$ satisfying (6.4) and such that

$$0 \in \partial\Big(\sum_{i=0}^{m+r} \lambda_i \phi_i \Big)(\bar{z}) + N\Big(\bar{z}; \bigcap_{j=0}^{K} \Lambda_j\Big). \tag{6.9}$$

By the *intersection rule* for basic normals from [16, Corollary 3.37], we have

$$N\Big(\bar{z}; \bigcap_{j=0}^{K} \Lambda_j\Big) \subset N(\bar{z}; \Lambda_0) + \ldots + N(\bar{z}; \Lambda_K) \tag{6.10}$$

provided that *all but one* of the sets Λ_j are SNC at $\bar{z}$ and that the *qualification condition*

$$\big[z_0^* + \ldots + z_K^* = 0, \ z_j^* \in N(\bar{z}; \Lambda_j)\big] \Longrightarrow \big[z_j^* = 0, \ j = 0, \ldots, K\big] \tag{6.11}$$

is satisfied. It is easy to observe, by the structures of the sets Λ_j, that the above SNC conditions for Λ_j reduce to the SNC requirements on Ω and F_j imposed in the theorem. Furthermore, these structures yield that the generalized normals

$$z_j^* = (x_{0j}^*, \ldots, x_{Kj}^*) \in N(\bar{z}_j; \Lambda_j), \quad j = 0, \ldots, K, \tag{6.12}$$

admit the representations

$$x_{jj}^* \in D_N^* F_j(\bar{x}_j, \bar{x}_{j+1})(-x_{(j+1)j}^*) \text{ and } x_{ij}^* = 0 \text{ as } i \neq j, j+1 \text{ for } j = 0, \ldots, K-1; \tag{6.13}$$

$$(x_{0K}^*, x_{KK}^*) \in N\big((\bar{x}_0, \bar{x}_K); \Omega\big) \text{ and } x_{iK}^* = 0 \text{ as } i \neq 0, K.$$

It follows from (6.9) and (6.10) that, under the SNC and qualification conditions imposed, there are generalized normals $z_j^*, j = 0, \ldots, K$, from (6.12) such that

$$-\sum_{j=0}^{K} z_j^* \in \partial\Big(\sum_{i=0}^{m+r} \lambda_i \phi_i \Big)(\bar{z}). \tag{6.14}$$

Taking into account the structures of ϕ_i in (6.7) and representations of z_j^* in (6.13), we deduce from (6.14) that

$$(-x_{00}^*, -x_{K(K-1)}^*) \in \partial\Big(\sum_{i=0}^{m+r} \lambda_i \varphi_i \Big)(\bar{x}_0, \bar{x}_K) + N\big((\bar{x}_0, \bar{x}_K); \Omega\big), \tag{6.15}$$

$$-x_{jj}^* = x_{(j+1)j}^* \text{ for } j = 0, \ldots, K-1.$$

Introducing now the *adjoint trajectory* $p := (p_0, \ldots, p_K) \in (X^*)^{K+1}$ by

$$p_j := x_{j(j-1)}^* \text{ for } j = 1, \ldots, K \text{ and } p_0 := -x_{00}^*, \tag{6.16}$$

we get from (6.13) and (6.15) both Euler-Lagrange and transversality inclusions of the theorem, where $(\lambda_0, \ldots, \lambda_{m+r}) \neq 0$ in the nontriviality condition (6.5).

Consider next the case when the qualification requirement (6.11) is *not fulfilled*. This means that there are $(z_0^*, \ldots, z_K^*) \neq 0$ such that

$$z_j^* \in N(\bar{x}; \Lambda_j) \text{ as } j = 0, \ldots, K \text{ with } z_0^* + \ldots + z_K^* = 0. \tag{6.17}$$

18

Equality (6.17) obviously implies that $(z_0^*, \ldots, z_{K-1}^*) \neq 0$. Then defining $(p_0, \ldots, p_K)$ by (6.16), we deduce from (6.13) and (6.17) that there is an adjoint discrete trajectory $p = (p_0, \ldots, p_K) \neq 0$ satisfying all the conclusions of the theorem with $(p_0, \ldots, p_K) \neq 0$ in the nontriviality condition (6.5).

Finally, let us show that the relationships (6.2)–(6.4) hold with the *enhanced* transversality condition (6.6) under the additional assumptions imposed in the theorem. It will follow from the implication

$$p_K \implies [p_j = 0 \text{ for all } j = 0, \ldots, K-1] \tag{6.18}$$

that is fulfilled along the trajectory $p = (p_0, \ldots, p_K)$ of the Euler-Lagrange inclusion (6.2) provided that the mappings F_j are Lipschitz-like around $(\bar{x}_j, \bar{x}_{j+1})$ and strongly coderivatively normal at these points for all $j = 0, \ldots, K-1$. It is sufficient to justify that $p_{K-1} = 0$ whenever $p_K = 0$ under the above assumptions imposed at the point $(\bar{x}_{K-1}, \bar{x}_K)$; the other relationships in (6.18) are proved similarly.

To proceed, we get $-p_{K-1} \in D_N^* F_{K-1}(\bar{x}_{K-1}, \bar{x}_K)(0)$ from the Euler-Lagrange inclusion (6.2) as $j = K-1$ and $p_K = 0$. Since F_{K-1} is strongly coderivatively normal at $(\bar{x}_{K-1}, \bar{x}_K)$,

$$D_N^* F_{K-1}(\bar{x}_{K-1}, \bar{x}_K)(0) = D_M^* F_{K-1}(\bar{x}_{K-1}, \bar{x}_K)(0),$$

while $D_M^* F_{K-1}(\bar{x}_{K-1}, \bar{x}_K)(0) = \{0\}$ by the coderivative characterization (3.3) of Lipschitz-like mappings. Thus $p_{K-1} = 0$, which completes the proof of the theorem. $\triangle$

As a consequence of Theorem 6.1, we derive next necessary optimality conditions for the discrete optimal control problem (6.1) in the form of the *discrete maximum principle*. It is well known that maximum-type conditions require certain *convexity* assumptions, which are generally unavoidable even for systems with smooth dynamics; cf. the abstract framework of constrained optimization in Corollary 4.4. The following corollary lists the precise requirements that ensure the fulfillment of the discrete maximum principle in the constrained Mayer problem (6.1) for discrete-time evolution inclusions.

Recall [20] that a set-valued mapping $F \colon X \rightrightarrows Y$ between Banach spaces is *inner semicontinuous* at $\bar{x}$ if for every $\bar{y} \in F(\bar{x})$ and every sequence $x_k \to \bar{x}$ with $F(x_k) \neq \emptyset$ there is a sequence of $y_k \in F(x_k)$ converging to $\bar{y}$ as $k \to \infty$.

Corollary 6.2 (discrete maximum principle for constrained evolution inclusions).
Let, in addition to the assumptions imposed in the first part of Theorem 6.1, the mappings F_j be inner semicontinuous at $(\bar{x}_j, \bar{x}_{j+1})$ and convex-valued around these points for all $j = 0, \ldots, K-1$. Then there are multipliers $(\lambda_0, \ldots, \lambda_{m+r})$ and an adjoint discrete trajectory $(p_0, \ldots, p_K)$ satisfying relationships (6.2)–(6.5) and the maximum condition

$$\langle p_{j+1}, \bar{x}_{j+1} \rangle = \max_{y \in F(\bar{x}_j)} \langle p_{j+1}, y \rangle \text{ for all } j = 0, \ldots, K-1. \tag{6.19}$$

Proof. It follows from [16, Theorem 1.34] that, under the convexity and inner semicontinuity assumptions imposed on F_j, one has

$$D_N^* F_j(\bar{x}_j, \bar{x}_{j+1})(p) \neq \emptyset \implies \left[\langle p, \bar{x}_{j+1} \rangle = \min_{y \in F_j(\bar{x}_j)} \langle p, y \rangle \right], \quad j = 0, \ldots, K-1.$$

Thus the discrete maximum condition (6.19) readily follows from the Euler-Lagrange inclusion (6.2) under the additional assumptions of the corollary. $\triangle$

Finally, let us consider the following Mayer-type problem for *parametric discrete-time evolution inclusions*:

$$\left\{ \begin{array}{l} \text{minimize } \varphi_0(x_0, x_K) \text{ subject to} \\[2mm] x_{j+1} \in x_j + hF_j(x_j) \text{ for } j = 0, \ldots, K-1, \\[2mm] \varphi_i(x_0, x_K) \leq 0 \text{ for } i = 1, \ldots, m, \\[2mm] \varphi_i(x_0, x_K) = 0 \text{ for } i = m+1, \ldots, m+r, \\[2mm] (x_0, x_K) \in \Omega \subset X^2 \end{array} \right. \tag{6.20}$$

with general endpoint constraints, where F_j and φ_i are as in (6.1), while $h > 0$ is a parameter. Problems of type (6.20) particularly arise from *discrete/finite-difference approximations* of constrained optimal control systems governed by *differential inclusions*:

$$\text{minimize } \varphi\big(x(a), x(b)\big) \text{ subject to } \dot{x}(t) \in F\big(x(t), t\big) \text{ a.e. } t \in [a, b]$$

under the endpoint constraints; see [17, Chapter 6] for more details and references.

The next result, which is a consequence of Theorem 6.1 and certain rules of generalized differential and SNC calculi, provides necessary optimality conditions of the extended Euler-Lagrange type for problem (6.20) in Asplund spaces under minimal assumptions on the initial data; see the discussions below.

Theorem 6.3 (extended Euler-Lagrange conditions for discrete approximations). *Let $\{\bar{x}_j \mid j = 0, \ldots, K\}$ be a local optimal solution to problem (6.20) with the Asplund state space X. Assume that Ω and $\mathrm{gph}\, F_j$ are locally closed while φ_i are locally Lipschitzian around $(\bar{x}_0, \bar{x}_K)$ and $\big(\bar{x}_j, (\bar{x}_{j+1} - \bar{x}_j)/h\big)$, respectively, for every $i = 0, \ldots, m+r$ and $j = 0, \ldots, K-1$. Assume also that all but one of the sets Ω and $\mathrm{gph}\, F_j$ are SNC at the corresponding points $(\bar{x}_0, \bar{x}_K)$ and $\big(\bar{x}_j, (\bar{x}_{j+1} - \bar{x}_j)/h\big)$, $j = 0, \ldots, K-1$. Then there exist multipliers $(\lambda_0, \ldots, \lambda_{m+r}) \in \mathbb{R}^{m+r+1}$ and an adjoint discrete trajectory $\{p_j \in X^* \mid j = 0, \ldots, K\}$, not simultaneously zero, satisfying the extended Euler-Lagrange inclusion*

$$\frac{p_{j+1} - p_j}{h} \in D_N^* F_j\Big(\bar{x}_j, \frac{\bar{x}_{j+1} - \bar{x}_j}{h}\Big)\big(-p_{j+1}\big), \quad j = 0, \ldots, K-1, \tag{6.21}$$

accompanied by the transversality inclusion (6.3) together with the sign and complementary slackness conditions (6.4).

Proof. Problem (6.20) can be considered as a particular case of the general discrete-time Mayer problem (6.1) governed by the discrete inclusions

$$x_{j+1} \in G_j(x_j) \text{ with } G_j(x) := x + hF_j(x), \quad j = 0, \ldots, K-1, \tag{6.22}$$

subject to the above endpoint constraints. Applying Theorem 6.1 to this form of (6.20), we need to present the assumptions and necessary optimality conditions therein in terms of the initial data of (6.20). First check that, for each $j = 0, \ldots, K-1$, the SNC property of G_j at $(\bar{x}_j, \bar{x}_{j+1})$ is equivalent to this property of F_j $\big(\bar{x}_j, (\bar{x}_{j+1} - \bar{x}_j)/h\big)$. In fact, this equivalence follows directly from the *SNC calculus* result of [16, Theorem 1.70] applied to the mapping addition in (6.22).

Furthermore, applying the Euler-Lagrange inclusion (6.2) to G_j and using the *coderivative sum rule* from [16, Theorem 1.62(ii)] with taking into account the obvious relationship $D_N^*(hF)(\cdot) = hD_N^*F(\cdot)$ held for any mapping F and number $h > 0$, we readily deduce (6.21) from (6.2) and thus complete the proof of the theorem. $\triangle$

The results and methods developed in this section can be similarly applied to the so-called *Bolza-type problems* for discrete-time inclusions and discrete approximations. In the latter case the cost functional is given by

$$\text{minimize} \quad \varphi_0(x_0, x_K) + h \sum_{j=0}^{K-1} \vartheta_j\left(x_j, \frac{x_{j+1} - x_j}{h}\right), \qquad (6.23)$$

which comes, in particular, from the classical Euler finite-difference replacement of the time-derivative in the Bolza functional

$$\text{minimize} \quad \varphi_0\big(x(a), x(b)\big) + \int_a^b \vartheta\big(x(t), \dot{x})(t), t\big)\, dt$$

for differential inclusions; see [17, Chapter 6] for more details.

In [17, Subsection 6.1.4], we studied the Bolza problem for discrete approximations of differential inclusions in Asplund spaces with geometric endpoint constraints by using its reduction to infinite-dimensional problems of mathematical programming with *operator constraints of Fredholm type*. Such an approach allowed us to establish extended Euler-Lagrange conditions in the form of Theorem 6.3 but under more restrictive assumptions requiring that *all* the sets Ω and gph F_j for $j = 0, \ldots, K - 1$ are *SNC* at the corresponding points. In this paper we achieve, using another approach, significant improvements of the previous results requiring the SNC property of *all but one* of the above sets. This particularly allows us to *fully avoid* any *SNC/finite codimension* type assumptions on the *constraint/target* set Ω providing that all the mapping F_j, $j = 0, \ldots, K - 1$, are SNC at the corresponding points. It covers, e.g., the case of two-point constraints with fixed endpoints x_0 and x_K, which has always been troublesome in infinite-dimensional optimal control; cf. the books by Fattorini [6] and Li and Yong [14], and Mordukhovich [17].

References

[1] Aubin, J.-P. 1984. Lipschitz behavior of solutions to convex minimization problems. *Math. Oper. Res.* **9** 87–111.

[2] Borwein, J. M., H. M. Strójwas. 1985. Tangential approximations. *Nonlinear Anal.* **9** 1347–1366.

[3] Borwein, J. M., Q. J. Zhu. 2005. *Techniques of Variational Analysis.* CMS Books in Mathematics. Springer, New York.

[4] Brinkhuis, J., V. M. Tikhomirov. 2005. *Optimization: Insights and Applications.* Princeton University Press, Princeton, NJ.

[5] Fabian, M., B. S. Mordukhovich. 2003. Sequential normal compactness versus topological normal compactness in variational analysis. *Nonlinear Anal.* **54** 1057–1067.

[6] Fattorini, H. O. 1999. *Infinite Dimensional Optimization and Control Theory*. Cambridge University Press. Cambridge, UK.

[7] Ginsburg, B., A. D. Ioffe. 1996. The maximum principle in optimal control of systems governed by semilinear equations. B. S. Mordukhovich, H. J. Sussmann, eds. *Nonsmooth Analysis and Geometric Methods in Deterministic Optimal Control*. Springer, New York, 81–110.

[8] Hiriart-Urruty, J.-B. 1979. Refinements of necessary optimality conditions in nondifferentiable programming I. *Appl. Math. Optim.* **5** 63–82.

[9] Ioffe, A. D. 1979. Necessary and sufficient conditions for a local minimum, I: A reduction theorem and first order conditions. *SIAM J. Control Optim.* **17** 245–250.

[10] Ioffe, A. D. 1984. Necessary conditions in nonsmooth optimization. *Math. Oper. Res.* **9** 159–188.

[11] Ioffe, A. D. 1997. Directional compactness and nonsmooth semi-Fredholm mappings. *Nonlinear Anal.* **29** 201–219.

[12] Ioffe, A. D., V. M. Tikhomirov. 1979. *Theory of Extremal Problems*. North-Holland, Amsterdam.

[13] Kruger, A. Y. 1985. Generalized differentials of nonsmooth functions and necessary conditions for an extremum. *Siberian Math. J.* **26** 370–379.

[14] Li, X., J. Yong. 1995. *Optimal Control Theory for Infinite-Dimensional Systems*. Birkhäuser, Boston, MA.

[15] Mordukhovich, B. S. 1985. On necessary conditions for an extremum in nonsmooth optimization. *Soviet Math. Dokl.* **32** 215–220.

[16] Mordukhovich, B. S. 2006. *Variational Analysis and Generalized Differentiation, I: Basic Theory*. Grundlehren Series (Fundamental Principles of Mathematical Sciences) **330**. Springer, Berlin.

[17] Mordukhovich, B. S. 2006. *Variational Analysis and Generalized Differentiation, II: Applications*. Grundlehren Series (Fundamental Principles of Mathematical Sciences) **331**. Springer, Berlin.

[18] Ngai, H. V., D. T. Luc, M. Théra. 2002. Extensions of Fréchet ε-subdifferential calculus and applications. *J. Math. Anal. Appl.* **268** 266–290.

[19] Phelps, R. R. 1993. *Convex Functions, Monotone Operators and Differentiability*, 2nd edition. Lecture Notes in Mathematics **1364**. Springer, Berlin.

[20] Rockafellar, R. T., R. J.-B. Wets. 1998. *Variational Analysis*. Grundlehren Series (Fundamental Principles of Mathematical Sciences) **317**. Springer, Berlin.

[21] Thibault, L. 1980. Subdifferentials of compactly Lipschitzian vector functions. *Ann. Mat. Pura Appl.* **125** 157–192.