DigitalCommons @WayneState

Wayne State University

Mathematics Research Reports

Mathematics

4-1-2006

Variational Analysis of Evolution Inclusions

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

Recommended Citation

Mordukhovich, Boris S., "Variational Analysis of Evolution Inclusions" (2006). *Mathematics Research Reports.* Paper 34.
http://digitalcommons.wayne.edu/math_reports/34

This Technical Report is brought to you for free and open access by the Mathematics at DigitalCommons@WayneState. It has been accepted for inclusion in Mathematics Research Reports by an authorized administrator of DigitalCommons@WayneState.

VARIATIONAL ANALYSIS OF EVOLUTION INCLUSIONS

BORIS S. MORDUKHOVICH

WAYNE STATE UNIVERSITY

Detroit, MI 48202

Department of Mathematics
Research Report

2006 Series
#4

This research was partly supported by the National Science Foundation and the Australian Research Council.

VARIATIONAL ANALYSIS OF EVOLUTION INCLUSIONS[1]

B. S. MORDUKHOVICH [2]

Abstract. The paper is devoted to optimization problems of the Bolza and Mayer types for evolution systems governed by nonconvex Lipschitzian differential inclusions in Banach spaces under endpoint constraints described by finitely many equalities and inequalities with generally nonsmooth functions. We develop a variational analysis of such problems mainly based on their discrete approximations and the usage of advanced tools of generalized differentiation satisfying comprehensive calculus rules in the framework of Asplund (and hence any reflexive Banach) spaces. In this way we establish extended results on stability of discrete approximations (with the strong $W^{1,2}$-convergence of optimal solutions under consistent perturbations of endpoint constraints) and derive necessary optimality conditions for nonconvex discrete-time and continuous-time systems in the refined Euler-Lagrange and Weierstrass-Pontryagin forms accompanied by the appropriate transversality inclusions. In contrast to the case of geometric endpoint constraints in infinite dimensions, the necessary optimality conditions obtained in this paper do not impose any nonempty interiority/finite codimension/normal compactness assumptions. The approach and results developed in the paper make a bridge between optimal control/dynamic optimization and constrained mathematical programming problems in infinite-dimensional spaces.

Key words. variational analysis, dynamic optimization and optimal control, evolution and differential inclusions, Banach and Asplund spaces, discrete/finite-difference approximations, nondifferentiable programming, generalized differentiation, necessary optimality conditions

AMS subject classification. 49J53, 49J52, 49J24, 49M25, 90C30

1 Introduction

This paper concerns the study of dynamic optimization problems governed by constrained evolution systems in infinite-dimensional spaces. We pay the main attention to variational analysis of the following *generalized Bolza problem* (P) for differential inclusions in Banach spaces with endpoint constraints described by finitely many equalities and inequalities.

Let X be a Banach *state space* with the *initial state* $x_0 \in X$, and let $T := [a,b] \subset I\!R$ be a fixed *time interval.* Given a set-valued mapping $F: X \times T \rightrightarrows X$ and real-valued functions $\varphi_i: X \to I\!R$ as $i = 0,\ldots,m+r$ and $\vartheta: X \times X \times T \to I\!R$, consider the problem:

$$\text{(1.1)} \qquad \text{minimize } \ J[x] := \varphi_0\big(x(b)\big) + \int_a^b \vartheta\big(x(t),\dot{x}(t),t\big)\,dt$$

subject to *dynamic constraints* governed by the evolution/differential inclusion

$$\text{(1.2)} \qquad \dot{x}(t) \in F\big(x(t),t\big) \ \text{ a.e. } \ t \in [a,b] \ \text{ with } \ x(a) = x_0$$

with *functional endpoint constraints* of the inequality and equality types given by

$$\text{(1.3)} \qquad \varphi_i\big(x(b)\big) \le 0, \quad i = 1,\ldots,m,$$

$$\text{(1.4)} \qquad \varphi_i\big(x(b)\big) = 0, \quad i = m+1,\ldots,m+r.$$

Note that $\dot{x}(t)$ stands in (1.1) for the time derivative of $x(t)$ and that "a.e." (almost everywhere) signifies as usual that the inclusion holds up to the Lebesgue measure zero on $I\!R$. The initial state x_0 and the time interval T are fixed in problem (P) for simplicity; the methods developed in this paper

[1]Research was partially supported by the USA National Science Foundation under grant DMS-0304989 and by the Australian Research Council under grant DP-0451168.

[2]Department of Mathematics, Wayne State University, Detroit, Michigan 48202, USA; boris@math.wayne.edu

allow us to include $x(a)$ and $[a,b]$ in the optimization process and to derive necessary optimality conditions for these variable data.

Dynamic optimization problems for differential inclusions with the *finite-dimensional* state space $X = I\!R^n$ have been intensively studied over the years, especially during the last decade, mainly from the viewpoint of deriving necessary optimality conditions; see [3, 8, 12, 14, 17, 19] for various results, methods, and more references. Dynamic optimization problems governed by infinite-dimensional *evolution equations* have also been much investigated, motivating mainly by applications to optimal control of partial differential equations; see, e.g., the books [7, 10] and the references therein. To the best of our knowledge, deriving necessary optimality conditions in dynamic optimization problems for evolution systems governed by differential inclusions in *infinite-dimensional spaces* has not drawn attention in the literature till the very recent time.

In the book [14], the author developed the method of *discrete approximations* to study optimal control problems of minimizing the Bolza functional (1.1) over appropriate solutions to evolution systems governed by infinite-dimensional differential inclusions of type (1.2) with endpoint constrains given in the *geometric form*

$$x(b) \in \Omega \subset X \tag{1.5}$$

via closed subsets of Banach spaces satisfying certain requirements. The major assumption on Ω made in [14] is the so-called *sequential normal compactness* (SNC) property of Ω at the optimal endpoint $\bar{x}(b) \in \Omega$; see [13] for a comprehensive theory for this and related properties, which play a significant role in infinite-dimensional variational analysis and its applications. Loosely speaking, the SNC property means that Ω should be "sufficiently fat" around the reference point; e.g., it never holds for singletons unless X is finite-dimensional, where the SNC property is satisfied for every nonempty set. For *convex* sets in infinite-dimensional spaces, the SNC property automatically holds when $\operatorname{int}\Omega \neq \emptyset$. Furthermore, it happens to be closely related [14] to the so-called "finite-codimension" property of convex sets, which is known to be essential for the fulfillment of an appropriate counterpart of the Pontryagin maximum principle for infinite-dimensional systems of optimal control; see the books by Fattorini [7] and by Li and Yong [10] for the corresponding results, discussions, counterexamples, and more references.

In this paper we show that the dynamic optimization problem (P) formulated above, with the *functional* endpoint constraints (1.3) and (1.4) given by *finitely many* Lipschitz continuous functions on a broad class of Banach spaces (that particularly includes every reflexive space), admits necessary optimality conditions in the extended Euler-Lagrange form accompanied by the corresponding Weierstrass-Pontryagin/maximum and transversality relations with *no SNC* and similar assumptions imposed on the underlying endpoint constraint set. Moreover, the case of endpoint constraints (1.3) and (1.4) under consideration allows us to partly avoid some other rather restrictive assumptions (like "strong coderivative normality," which may not hold in infinite-dimensional spaces; see Sections 6, 7 for more details) imposed in [14] in the general case of geometric constraints (1.5). Our approach is based, in addition to [14], on certain delicate properties of appropriate *subdifferentials* of locally Lipschitzian functions on infinite-dimensional spaces, as well as on *dual/coderivative characterizations* of Lipschitzian and metric regularity properties of set-valued mappings.

The rest of the paper is organized as follows. In Section 2 we formulate the standing assumptions on the initial data of (P), make more precise the solution concept for the evolution inclusion (1.1) and the types of local minimizers to (P) under consideration, and also discuss the relaxation procedure used for some results and proofs in the paper. The main attention in this paper is paid to the so-called *intermediate local minimizers*, which occupy (strictly) an intermediate position between the classical weak and strong minima, being nevertheless closer to strong minimizers from the viewpoint of necessary optimality conditions for differential inclusions.

In Section 3 we construct a sequence of the *well-posed discrete approximations* (P_N) to the original Bolza problem (P), which take into account specific features of the functional endpoint constraints (1.3) and (1.4) involving *consistent perturbations* of these constraints in the discrete approximation procedure. Then we present a major result on the *strong stability* of discrete approximations that justifies the $W^{1,2}$-norm convergence of optimal solutions for (P_N) to the fixed local minimizer for the original problem (P).

Section 4 contains an overview of the basic tools of *generalized differentiation* needed to perform the subsequent variational analysis of the discrete-time and continuous-time evolution systems under consideration in infinite-dimensional spaces. Most of the material in this section is taken from the author's book [13], where the reader can find more results and commentaries in this direction and related topics.

Section 5 is devoted to deriving necessary optimality conditions for the constrained *discrete-time* problems arising from the discrete approximation procedure whose well-posedness and stability are justified in Section 3. These problems are reduced to (non-dynamic) constrained problems of mathematical programming in infinite dimensions, which happen to be *intrinsically nonsmooth* and involve finitely many functional and geometric constraints generated by those in (1.2)–(1.4) via the discrete approximation procedure. Variational analysis of such problems requires applications of the full power of the generalized differential calculus in infinite-dimensional spaces developed in [13].

In Section 6 we derive necessary optimality conditions of the extended *Euler-Lagrange* type for *relaxed* intermediate minimizers to the original Bolza problem (P) by passing to the limit from those obtained for discrete-time problems in Section 5. It worth emphasizing that the realization of the limiting procedure requires not only the strong convergence of optimal trajectories to discrete approximation problems established in Section 3 but also justifying an appropriate convergence of *adjoint trajectories* in necessary optimality conditions for the sequence of discrete-time inclusions. The latter becomes passible due to specific properties of the basic generalized differential constructions reviewed in Section 4, which include complete *dual characterizations* of Lipschitzian and metric regularity properties of set-valued mappings.

The concluding Section 7 concerns necessary optimality conditions for arbitrary (*non-relaxed*) intermediate minimizers to problem (P), considering for simplicity the Mayer form (P_M) with no integral term in (1.1), that are established in terms of the *extended Euler-Lagrange* inclusion accompanied by the *Weierstrass-Pontryagin/maximum* and transversality relations without imposing any SNC assumptions on the target/endpoint constraint set described by (1.3) and (1.4). The approach is based on an additional approximation procedure that allows us to reduce (P_M) to an unconstrained (while nonsmooth and nonconvex) Bolza problem of the type treated in Section 6, for which *any* intermediate local minimizer happens to be a relaxed one. The passage to the limit from the latter approximation is largely similar to that developed in Section 6, not requiring however any relaxation requirement due to the usage of *Ekeland's variational principle.*

Our notation is basically standard; cf. [13, 14]. 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

$$\operatorname*{Lim\,sup}_{x\to\bar{x}} F(x) := \Big\{x^*\in X^*\Big|\ \exists \text{ sequences } x_k\to\bar{x},\ x_k^*\xrightarrow{w^*}x^* \text{ with } x_k^*\in F(x_k) \text{ as } k\in I\!N:=\{1,2,\ldots\}\Big\}. \tag{1.6}$$

2 The Generalized Bolza Problem for Evolution Inclusions

Just for brevity and simplicity, we consider in this paper the Bolza problem (P) with *autonomous* (time-independent) data, i.e., when $\vartheta = v(x,v)$ in (1.1) and $F = F(x)$ in (1.2). The case of non-autonomous systems can be studied similarly to [14, Chapter 6] devoted to problems with geometric constraints of type (1.5). Let us start with the precise definition of solutions (trajectories, arcs) to the differential inclusion (1.2) following the book by Deimling [6].

Definition 2.1 (solutions to differential inclusions in infinite-dimensional spaces). *By a* SOLUTION *to inclusion* (1.2) *we understand a mapping* $x\colon T\to X$, *which is Fréchet differentiable* *for a.e. $t \in T$ satisfying* (1.2) *and the* NEWTON-LEIBNIZ FORMULA

$$x(t) = x_0 + \int_a^t \dot{x}(s)\,ds \quad \textit{for all} \;\; t \in T,$$

where the integral in taken in the BOCHNER SENSE.

It is well known that for $X = \mathbb{R}^n$, $x(t)$ is a.e. differentiable on T and satisfies the Newton-Leibniz formula *if and only if* it is *absolutely continuous* on T in the standard sense. However, for infinite-dimensional spaces X even the Lipschitz continuity may not imply the a.e. differentiability. On the other hand, there is a *complete characterization* of Banach spaces X, where the absolute continuity of every $x\colon T \to X$ is *equivalent* to its a.e. differentiability and the fulfillment of the Newton-Leibniz formula: this is the class of spaces having the so-called *Radon-Nikodým property* (RNP), which is well investigated in the geometric theory of Banach spaces [4]. Observe, in particular, that every *reflexive* space enjoys the RNP.

Recall further that a Banach space X is *Asplund* if any of its separable subspaces has a separable dual. This is a major subclass of Banach spaces that particularly includes every space with a *Fréchet differentiable renorm* off the origin (i.e., every *reflexive* space), every space with a separable dual, etc.; see [4] for more details, characterizations, and references. There is a deep relationship between spaces having the RNP and Asplund spaces, which is used in what follows: *given a Banach space X, the dual space X^* has the RNP if and only if X is Asplund.*

It has been well recognized that differential inclusions (1.2), which are certainly of their own interest, provide a useful generalization of *control systems* governed by differential/evolution *equations* with control parameters:

$$\dot{x} = f(x,u), \quad u \in U, \tag{2.1}$$

where the control sets $U(\cdot)$ may also depend on time and *state* variables via $F(x,t) = f(x,U(x,t),t)$. In some cases, especially when the sets $F(\cdot)$ are convex, the differential inclusions (1.2) admit parametric representations of type (2.1), but in general they cannot be reduced to parametric control systems and should be studied for their own sake. Note also that the *ODE form* (2.1) in Banach spaces is strongly related to various control problems for evolution *partial differential equations* of parabolic and hyperbolic types, where solutions may be understood in some other appropriate senses; see, e.g., the books [7, 10, 14] for more discussions.

In what follows, we pay the main attention to the study of *intermediate local minimizers* for problem (P) introduced by the author in [12]. Recall that a *feasible arc* to (P) is a solution to the differential inclusion (1.2), in the sense of Definition 2.1, for which $J[x] < \infty$ in (1.1) and the endpoint constraints (1.3) and (1.4) are satisfied.

Definition 2.2 (intermediate local minimizers). *A feasible arc $\bar{x}(\cdot)$ is an* INTERMEDIATE LOCAL MINIMIZER *(i.l.m.) of rank $p \in [1,\infty)$ for (P) if there are numbers $\epsilon > 0$ and $\alpha \geq 0$ such that $J[\bar{x}] \leq J[x]$ for any feasible arcs to (P) satisfying the relationships*

$$\|x(t) - \bar{x}(t)\| < \epsilon \quad \textit{for all} \;\; t \in [a,b] \quad \textit{and} \tag{2.2}$$

$$\alpha \int_a^b \|\dot{x}(t) - \dot{\bar{x}}(t)\|^p\,dt < \epsilon. \tag{2.3}$$

In fact, relationships (2.2) and (2.3) mean that we consider a neighborhood of $\bar{x}(\cdot)$ in the Sobolev space $W^{1,p}([a,b];X)$ with the norm

$$\|x(\cdot)\|_{W^{1,p}} := \max_{t\in[a,b]} \|x(t)\| + \Big(\int_a^b \|\dot{x}(t)\|^p\,dt\Big)^{1/p},$$

where the norm on the right-hand side is taken in the space X. If there is only the requirement (2.2) in Definition 2.2, i.e., $\alpha = 0$ in (2.3), then we get the classical *strong* local minimum corresponding to a neighborhood of $\bar{x}(\cdot)$ in the norm topology of $C([a,b];X)$. If instead of (2.3) one puts the more restrictive requirement

$$\|\dot{x}(t)-\dot{\bar{x}}(t)\|<\epsilon \text{ a.e. } t\in[a,b],$$

then we have the classical *weak* local minimum in the framework of Definition 2.2. Thus the introduced notion of i.l.m. takes, for any $p\in[1,\infty)$, an *intermediate* position between the classical concepts of strong ($\alpha=0$) and weak ($p=\infty$) local minima, being indeed different from both classical notions; see various examples in [20, 14]. Clearly all the necessary conditions for i.l.m. automatically hold for strong (and hence for global) minimizers.

Considering the autonomous Bolza problem (P) in this paper, we impose the following *standing assumptions* on its initial data along a given intermediate local minimizer $\bar{x}(\cdot)$:

(H1) There are a open set $U\subset X$ and a number $\ell_F>0$ such that $\bar{x}(t)\in U$ for all $t\in[a,b]$, the sets $F(x)$ are nonempty and compact for all $x\in U$ and satisfy the inclusion

$$F(x)\subset F(u)+\ell_F\|x-u\|I\!B \text{ whenever } x,u\in U, \tag{2.4}$$

which implies the uniform boundedness of the sets $F(x)$ on U, i.e., the existence of some constant $\gamma>0$ such that

$$F(x)\subset\gamma I\!B \text{ for all } x\in U.$$

(H2) The integrand ϑ is Lipschitzian continuous on $U\times(\gamma I\!B)$.

(H3) The endpoint functions φ_i, $i=0,\ldots,m+r$, are locally Lipschitzian around $\bar{x}(b)$ with the common Lipschitz constant $\ell>0$.

Observe that (2.4) is equivalent to say that the set-valued mapping F is *locally Lipschitzian* around $\bar{x}(\cdot)$ with respect to the classical Hausdorff metric on the space of nonempty and compact subsets of X.

In what follows, along with the original problem (P), we consider its "relaxed" counterpart significantly used in some results and proofs of the paper. Roughly speaking, the relaxed problem is obtained from (P) by a *convexification* procedure with respect to the *velocity* variable. It follows the route of Bogolyubov and Young in the classical calculus of variations and of Gamkrelidze and Warga in optimal control; see the book [14] and the references therein for more details and commentaries.

To construct an appropriate relaxation of the Bolza problem (P) under consideration, we first consider the extended-real-valued function

$$\vartheta_F(x,v):=\vartheta(x,v)+\delta(v;F(x)),$$

where $\delta(\cdot;\Omega)$ is the *indicator function* of the set Ω equal to 0 on Ω and to ∞ out of it. Denote by

$$\widehat{\vartheta}_F(x,v):=(\vartheta_F)^{**}_v(x,v),\quad (x,v)\in X\times X,$$

the *biconjugate/bypolar* function to $\vartheta_F(x,\cdot)$, i.e., the greatest proper, convex, and lower semicontinuous (l.s.c.) function with respect to v, which is majorized by ϑ_F. Then the *relaxed problem* (R) to (P), or the *relaxation* of (P), is defined as follows:

$$\text{minimize } \widehat{J}[x]:=\varphi(x(b))+\int_a^b\widehat{\vartheta}_F(x(t),\dot{x}(t))\,dt \tag{2.5}$$

over a.e. differentiable arcs $x\colon[a,b]\to X$ that are Bochner integrable on $[a,b]$ together with $\vartheta_F(x(t),\dot{x}(t))$, satisfy the Newton-Leibniz formula and the endpoint constraints (1.3), (1.4).

Note that the feasibility requirement $\widehat{J}[x]<\infty$ in (2.5) is fulfilled only if $x(\cdot)$ is a solution (in the sense of Definition 2.1) to the *convexified differential inclusion*

$$\dot{x}(t)\in\operatorname{clco}F(x(t),\dot{x}(t)) \text{ a.e. } t\in[a,b] \text{ with } x(a)=x_0, \tag{2.6}$$

where "clco" stands for the convex closure of a set in X. Thus the relaxed problem (R) can be considered under explicit dynamic constraints given by the convexified differential inclusion (2.6). Any trajectory for (2.6) is called a *relaxed trajectory* for (1.2), in contrast to the *ordinary* (or *original*) trajectories for the latter inclusion.

There are deep relationships between relaxed and ordinary trajectories for differential inclusions, which reflect the fundamental *hidden convexity* inherent in continuous-time (nonatomic measure) dynamic systems defined by differential and integral operators. In particular, *any relaxed trajectory* of (1.2) under assumption (H1) can be *uniformly approximated* (in the $C([a,b];X)$-norm) by a sequence of ordinary trajectories; see, e.g., [6, 18]. We need the following version [5] of this approximation/density property involving not only differential inclusions but also minimizing functionals.

Lemma 2.3 (approximation property for the relaxed Bolza problem). *Let $x(\cdot)$ be a relaxed trajectory for the differential inclusion* (1.2) *with a separable state space X, where F and ϑ satisfy assumptions* (H1) *and* (H2), *respectively. Then there is sequence of the ordinary trajectories $x_k(\cdot)$ for* (1.2) *such that $x_k(\cdot) \to x(\cdot)$ in $C([a,b];X)$ as $k \to \infty$ and*

$$\liminf_{k\to\infty} \int_a^b \vartheta(x_k(t),\dot{x}_k(t))\,dt \le \int_a^b \widehat{\vartheta}_F(x(t),\dot{x}(t))\,dt.$$

Note that Theorem 2.3 does *not* assert that the approximating trajectories $x_k(\cdot)$ satisfy the endpoint constraints (1.3) and (1.4). Indeed, there are examples showing that the latter may not be possible and, moreover, the property of *relaxation stability*

$$\inf(P) = \inf(R) \tag{2.7}$$

is violated; in (2.7) the infima of the cost functionals (1.1) and (2.5) are taken over all the feasible arcs in (P) and (R), respectively.

An obvious sufficient condition for the relaxation stability is the *convexity* of the sets $F(x,t)$ and of the integrand ϑ in v. However, the relaxation stability goes *far beyond* the standard convexity due to the *hidden convexity* property of continuous-time differential systems. In particular, Theorem 2.3 ensures the relaxation stability of nonconvex problems (P) with no constraints on the endpoint $x(b)$. There are various efficient conditions for the relaxation stability of nonconvex problems with endpoint and other constraint; see [14, Subsection 6.1.2] with the commentaries therein for more details, discussions, and references.

A *local* version of the relaxation stability property (2.7) regarding intermediate minimizers for the Bolza problem (P) is postulated as follows.

Definition 2.4 (relaxed intermediate local minimizers). *A feasible arc $\bar{x}(\cdot)$ to the Bolza problem (P) is a* RELAXED INTERMEDIATE LOCAL MINIMIZER *(r.i.l.m.) of rank $p \in [1,\infty)$ for (P) if it is an intermediate local minimizer of this rank for the relaxed problem (R) providing the same value of the cost functionals: $J[\bar{x}] = \widehat{J}[\bar{x}]$.*

It is not hard to observe that, under the standing assumptions formulated above, the notions of intermediate local minima and relaxed intermediate local minima do not actually depend on rank p, i.e., they either hold or violate for all $p \in [1,\infty)$ simultaneously. In what follows we always take (unless otherwise stated in Section 7) $p = 2$ and $\alpha = 1$ in (2.3) for simplicity.

The principal method of our study in this paper involves *discrete approximations* of the original Bolza problem (P) for constrained continuous-time evolution inclusions by a family of dynamic optimization problems of the Bolza type governed by discrete-time inclusions with endpoint constraints. We show that this method generally leads to necessary optimality conditions for *relaxed* intermediate local minimizers of (P). Then an additional approximation procedure allows us to establish necessary optimality conditions for *arbitrary* (non-relaxed) intermediate local minimizers by reducing them to problems, which are *automatically* stable with respect to relaxation.

3 Stability of Discrete Approximations

In this section we present basic constructions of the method of discrete approximations in the theory of necessary optimality conditions for differential inclusions following the scheme of [12, 14] developed for the case of geometric constraints, with certain modifications required for the functional endpoint constraints (1.3) and (1.4) under consideration in infinite-dimensional spaces.

Since we use discrete approximations mostly from a "theoretical" viewpoint (as a vehicle to derive necessary optimality conditions), we use in what follows just the simplest finite-difference replacement of the derivative by the *uniform Euler scheme*:

$$\dot{x}(t) \approx \frac{x(t+h)-x(t)}{h}, \quad h \to 0.$$

To formalize this process, we take any natural number $N \in \mathbb{N}$ and consider the *discrete grid/mesh* on T defined by

$$T_N := \{a, a+h_N, \ldots, b-h_N, b\}, \quad h_N := (b-a)/N,$$

with the *stepsize of discretization* h_N and the *mesh points* $t_j := a + jh_N$ as $j = 0, \ldots, N$, where $t_0 = a$ and $t_N = b$. Then the differential inclusion (1.2) is replaced by a sequence of its *finite-difference/discrete approximations*

$$\text{(3.1)} \qquad x_N(t_{j+1}) \in x_N(t_j) + h_N F(x_N(t_j)), \; j = 0, \ldots, N-1, \quad x(t_0) = x_0.$$

Given a discrete trajectory $x_N(t_j)$ satisfying (3.1), we consider its *piecewise linear extension* $x_N(t)$ to the continuous-time interval $T = [a, b]$, i.e., the *Euler broken lines*. We also define the *piecewise constant extension* to T of the corresponding *discrete velocity* by

$$v_N(t) := \frac{x_N(t_{j+1}) - x_N(t_j)}{h_N}, \quad t \in [t_j, t_{j+1}), \; j = 0, \ldots, N-1.$$

It follows from the very definition of the Bochner integral that

$$x_N(t) = x_0 + \int_a^t v_N(s)\,ds \;\text{ for }\; t \in T.$$

The next result, which plays a significant role in the method of discrete approximations, establish the *strong* $W^{1,2}$*-norm approximation* of *any* trajectory for the differential inclusion (1.2) by extended trajectories of the sequence of discrete inclusions (3.1) under the general assumptions made in (H1). Note that the norm convergence in $W^{1,2}([a,b];X)$ implies the *uniform* convergence of the trajectories on $[a, b]$ and the *pointwise*, for a.e. $t \in [a, b]$, convergence of (some subsequence of) their *derivatives*. The latter is crucial for the purposes of this paper, especially in the case of *nonconvex*-valued differential inclusions. The proof of this result in given in [14, Theorem 6.4], which is an infinite-dimensional counterpart of the one in [12, Theorem 3.1].

Lemma 3.1 (strong $W^{1,2}$-approximation by discrete trajectories). *Let $\bar{x}(\cdot)$ be an arbitrary solution to the differential inclusion* (1.2) *under the assumptions in* (H1), *where X is a general Banach space. Then there is a sequence of solutions $\widehat{x}_N(t_j)$ to the discrete inclusions* (3.1) *such that their extensions $\widehat{x}_N(t)$, $a \le t \le b$, converge to $\bar{x}(t)$ strongly in the space $W^{1,2}([a,b];X)$ as $N \to \infty$.*

Observe that the proof of the above result given in [12, 14] is *constructive* and provides efficient estimates of the convergence rate being of certain independent interest for numerical analysis.

Now fix an *intermediate local minimizer* $\bar{x}(\cdot)$ for the Bolza problem (P) and construct a sequence of discrete approximation problems (P_N), $N \in \mathbb{N}$, admitting optimal solutions $\bar{x}_N(\cdot)$ whose extensions converge to $\bar{x}(\cdot)$ in the norm topology of $W^{1,2}([a,b];X)$ as $N \to \infty$.

To proceed, we take a sequence of the discrete trajectories $\widehat{x}_N(\cdot)$ approximating by Lemma 3.1 the given local minimizer $\bar{x}(\cdot)$ to (P) and denote

$$\text{(3.2)} \qquad \eta_N := \max_{j \in \{1,\ldots,N\}} \|\widehat{x}_N(t_j) - \bar{x}(t_j)\| \to 0 \;\text{ as }\; N \to \infty.$$

In [14, Subsection 6.1.1], the reader can find more information on computing and estimating η_N, which is not needed in what follows: it is sufficient to know that $\eta_N \to 0$ as $N \to \infty$.

Having $\epsilon > 0$ from relations (2.2) and (2.3) for the intermediate minimizer $\bar{x}(\cdot)$ with $p = 2$ and $\alpha = 1$, we always suppose that

$$\bar{x}(t) + \epsilon/2 \in U \quad \text{for all} \quad t \in [a, b],$$

where U is a neighborhood of $\bar{x}(\cdot)$ from (H1). Let $\ell > 0$ be the common Lipschitz constant of φ_i, $i = 1, \ldots, m+r$, from (H3). Construct problems (P_N), $N \in I\!N$, as follows: minimize

$$\tag{3.3} \begin{aligned} J_N[x_N] :&= \varphi_0\big(x_N(t_N)\big) + h_N \sum_{j=0}^{N-1} \vartheta\Big(x_N(t_j), \frac{x_N(t_{j+1}) - x_N(t_j)}{h_N}\Big) \\ &+ \sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} \Big\| \frac{x_N(t_{j+1}) - x_N(t_j)}{h_N} - \dot{\bar{x}}(t) \Big\|^2 dt \end{aligned}$$

over discrete trajectories $x_N = x_N(\cdot) = \big(x_0, x_N(t_1), \ldots, x_N(t_N)\big)$ for the difference inclusions (3.1) subject to the constraints

$$\tag{3.4} \varphi_i\big(x_N(t_N)\big) \le \ell\eta_N \quad \text{for} \quad i = 1, \ldots, m,$$

$$\tag{3.5} -\ell\eta_N \le \varphi_i\big(x_N(t_N)\big) \le \ell\eta_N \quad \text{for} \quad i = m+1, \ldots, m+r,$$

$$\tag{3.6} \|x_N(t_j) - \bar{x}(t_j)\| \le \frac{\epsilon}{2} \quad \text{for} \quad j = 1, \ldots, N, \quad \text{and}$$

$$\tag{3.7} \sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} \Big\| \frac{x_N(t_{j+1}) - x_N(t_j)}{h_N} - \dot{\bar{x}}(t) \Big\|^2 dt \le \frac{\epsilon}{2}.$$

Considering in the sequel (without mentioning any more) piecewise linear extension of $x_N(\cdot)$ to the whole interval $[a, b]$, we observe the relationships:

$$\tag{3.8} \begin{cases} x_N(t) = x_0 + \displaystyle\int_a^t \dot{x}_N(s)\, ds \quad \text{for all } t \in [a, b] \quad \text{and} \\ \dot{x}_N(t) = \dot{x}_N(t_j) \in F\big(x_N(t_j)\big), \quad t \in [t_j, t_{j+1}), \quad j = 0, \ldots, N-1. \end{cases}$$

In the next theorem, we establish that the given *relaxed* intermediate local minimizer (r.i.l.m.) $\bar{x}(\cdot)$ to (P) can be approximated by *optimal solutions* to (P_N) *strongly* in $W^{1,2}([a, b]; X)$; the latter implies the a.e. *pointwise* convergence of the derivatives significant for the main results of the paper. To justify such an approximation, we need to impose the *Asplund* structure on *both* the state space X and its dual X^*, which is particularly the case when X is *reflexive*. Note also there are *nonreflexive* (even separable) spaces for which both X and X^* are Asplund; see, e.g., [4].

Theorem 3.2 (strong convergence of discrete optimal solutions). *Let $\bar{x}(\cdot)$ be an r.i.l.m. for the Bolza problem (P) under the standing assumptions* (H1)–(H3) *in the Banach state space X, and let (P_N), $N \in I\!N$, be a sequence of discrete approximation problems built above. The following hold:*

(i) *Each (P_N) admits an optimal solution.*

(ii) *If in addition both X and X^* are Asplund, then any sequence $\{\bar{x}_N(\cdot)\}$ of optimal solutions to (P_N) converges to $\bar{x}(\cdot)$ strongly in $W^{1,2}([a, b]; X)$.*

Proof. To justify assertion (i), we first observe that the set of *feasible* solutions to each problem (P_N) is *nonempty* for all $N \in I\!N$ sufficiently large. Indeed, pick the discrete trajectory $\widehat{x}_N(\cdot)$ approximating the given local minimizer $\bar{x}(\cdot)$ by Lemma 3.1 and show that it satisfies all the constraints (3.4)–(3.7) for large N. By assumption (H3) we have

$$|\varphi_i(\widehat{x}_N(t_N)) - \varphi_i(\bar{x}(b))| \le \ell \|\widehat{x}(t_N) - \bar{x}(t_N)\| \le \ell\eta_N \quad \text{for all } i = 1,\ldots,m+r$$

due to (3.2). This implies the fulfillment of the endpoint constraints (3.4) and (3.5) for $\widehat{x}_N(\cdot)$, since those in (1.3) and (1.4) hold for $\bar{x}(\cdot)$. The fulfillment of (3.6) for $\widehat{x}_N(\cdot)$ follows directly from the construction of $\eta_N \to 0$ in (3.2). Further, it is easy to check that

$$\sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} \left\| \frac{\widehat{x}_N(t_{j+1}) - \widehat{x}_N(t_j)}{h_N} - \dot{\bar{x}}(t) \right\|^2 dt = \int_a^b \|\dot{\widehat{x}}_N(t) - \dot{\bar{x}}(t)\|^2\, dt =: \alpha_N$$

for the piecewise linear extension of $\widehat{x}_N(\cdot)$ to $[a,b]$. By the $W^{1,2}$-approximation in Lemma 3.1 we have that $\alpha_N \to 0$ as $N \to \infty$, which justifies the fulfillment of (3.7) for large N. The *existence of optimal solutions* to (P_N) follows now from the classical Weierstrass theorem due to the compactness and continuity assumptions made in (H1)–(H3).

Let us now prove the convergence assertion (ii) under the additional assumptions on the state space. Check first the *value convergence*

$$J_N[\bar{x}_N] \to J[\bar{x}] \quad \text{as} \quad N \to \infty \tag{3.9}$$

along a subsequence of $N \to \infty$. Considering the expression for $J_N[\widehat{x}_N]$ in (3.3) and using assumptions (H2) and (H3), we observe that (3.9) follows from

$$\begin{aligned} h_N \sum_{j=0}^{N-1} \vartheta\Big(\widehat{x}_N(t_j), \frac{\widehat{x}_N(t_{j+1}) - \widehat{x}_N(t_j)}{h_N}\Big) &= \sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} \vartheta(\widehat{x}_N(t_j), \dot{\widehat{x}}_N(t))\, dt \\ = \int_a^b \vartheta(\widehat{x}_N(t), \dot{\widehat{x}}_N(t))\, dt + O(h_N) &\to \int_a^b \vartheta(\bar{x}(t), \dot{\bar{x}}(t))\, dt \quad \text{as} \quad N \to \infty, \end{aligned}$$

which hold by Lemma 3.1 ensuring the a.e. convergence $\dot{\widehat{x}}_N(t) \to \dot{\bar{x}}(t)$ along a subsequence and by the Lebesgue dominated convergence theorem valid for the Bochner integral.

All the previous arguments did not use either the relaxation property of the intermediate minimizer, or the Asplund property of X and X^*. Now we are going to employ these properties to justify the relationship

$$\lim_{N\to\infty} \Big[\beta_N := \int_a^b \|\dot{\bar{x}}_N(t) - \dot{\bar{x}}(t)\|^2\, dt\Big] = 0 \tag{3.10}$$

for *every* sequence of optimal solutions $\bar{x}_N(\cdot)$ to (P_N).

Arguing by contradiction, pick a limiting point $\beta > 0$ of $\{\beta_N\}$ in (3.10) and suppose for simplicity that $\beta_N \to \beta$ for all $N \to \infty$. To proceed, observe that both spaces X and X^* *enjoy the RNP*. Indeed, the one for X^* is equivalent to the Asplund property of X, while the Asplund property of X^* ensures the RNP for X due to the latter fact and that of $X \subset X^{**}$. Taking into account (H1) and (3.8), we apply to the sequence $\{\dot{\bar{x}}_N(\cdot)\}$ the Dunford theorem [4, Theorem IV.1] on the *weak compactness* in $L^1([a,b];X)$, which allows us to find a subsequence of $\{\dot{\bar{x}}_N(\cdot)\}$ and a mapping $v(\cdot) \in L^1([a,b];X)$ such that

$$\dot{\bar{x}}_N(\cdot) \to v(\cdot) \quad \text{weakly in } L^1([a,b];X) \quad \text{as} \quad N \to \infty. \tag{3.11}$$

Using (3.8) and the compactness in $C([a,b];X)$ of solution sets for differential inclusions that holds under the assumptions made in (H1) (see, e.g., [18, Theorem 3.4.2]), we conclude that the sequence $\{\bar{x}_N(\cdot)\}$ contains a subsequence that converges to some $\widetilde{x} \in C([a,b];X)$ in the norm topology of the space $C([a,b];X)$. Passing to the limit in the first relationship of (3.8), with taking into account

(3.11) and the weak continuity of the Bochner integral as an operator from $L^1([a,b];X)$ to X, we arrive at the representation

$$\widetilde{x}(t) = x_0 + \int_a^t v(s)\,ds \quad \text{for all} \quad t \in [a,b],$$

which implies that $v(t) = \dot{\widetilde{x}}(t)$ for a.e. $t \in [a,b]$.

Furthermore, the classical Mazur *weak closure* theorem ensures that $\widetilde{x}(\cdot)$ is a solution to the *convexified* differential inclusion (2.6). By the structure of problems (P_N) and by the construction of $\widetilde{x}(\cdot)$, it is not hard to conclude that $\widetilde{x}(\cdot)$ satisfies the endpoint constraints (1.3) and (1.4) and that it belongs to the prescribed ϵ-neighborhood of $\bar{x}(\cdot)$ in the norm topology of $W^{1,2}([a,b];X)$. By passing to the limit in the obvious inequality

$$J_N[\bar{x}_N] \le J_N[\widehat{x}_N] \quad \text{for all large} \quad N \in I\!N,$$

with taking into account (3.9) and the lower semicontinuity of the *convexified* integrand $\widehat{\vartheta}_F(x,\cdot)$ from (2.5) in the *weak* topology of $L^2([a,b];X)$, we get

$$\widehat{J}[\widetilde{x}] = \varphi_0(\widetilde{x}(b)) + \int_a^b \widehat{\vartheta}_F(\widetilde{x}(t), \dot{\widetilde{x}}(t))\,dt + \beta \le J[\bar{x}].$$

Since $\beta > 0$ and $J[\bar{x}] = \widehat{J}[\bar{x}]$, the latter gives $\widehat{J}[\widetilde{x}] < \widehat{J}[\bar{x}]$, which contradicts the choice of $\bar{x}(\cdot)$ as a relaxed intermediate local minimizer for (P). Thus (3.10) holds, and so $\bar{x}_N(\cdot) \to \bar{x}(\cdot)$ as $N \to \infty$ strongly in $W^{1,2}([a,b];X)$. This completes the proof of the theorem. $\triangle$

The strong convergence result of Theorem 3.2 *makes a bridge* between the original continuous-time dynamic optimization problem (P) and its discrete-time counterparts (P_N), which allows us to derive necessary optimality conditions for (P) by passing to the limit from those for (P_N). The latter ones are *intrinsically nonsmooth* and require appropriate tools of generalized differentiation for their variational analysis.

4 Generalized Differentiation

In this section, we define the main constructions of generalized differentiation used in what follows. Since our major framework in this paper is the class *Asplund spaces*, we present simplified definitions and some properties held in this setting. All the material reviewed and employed below is taken from the author's book [13], where the reader can find more details and references.

We start with generalized normals to closed sets $\Omega \subset X$. Given $\bar{x} \in \Omega$, the (basic, limiting) *normal cone* to Ω at $\bar{x}$ is defined by

$$N(\bar{x};\Omega) := \operatorname*{Lim\,sup}_{x \to \bar{x}} \widehat{N}(x;\Omega), \tag{4.1}$$

where "Lim sup" stands for the *sequential* upper/outer limit (1.6) of the *Fréchet normal cone* (or the *prenormal cone*) to Ω at $x \in \Omega$ given by

$$\widehat{N}(x;\Omega) := \Big\{ x^* \in X^* \Big|\ \limsup_{u \xrightarrow{\Omega} x} \frac{\langle x^*, u - x\rangle}{\|u - x\|} \le 0 \Big\}, \tag{4.2}$$

where $x \xrightarrow{\Omega} \bar{x}$ signifies that $x \to \bar{x}$ with $x \in \Omega$, and where $\widehat{N}(x;\Omega) := \emptyset$ for $x \notin \Omega$.

Given a set-valued mapping $F\colon X \rightrightarrows Y$ of closed graph

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

define its *normal coderivative* and *Fréchet coderivative* at $(\bar{x},\bar{y}) \in \operatorname{gph} F$ by, respectively,

$$D^*F(\bar{x},\bar{y})(y^*) := \big\{ x^* \in X^* \big|\ (x^*, -y^*) \in N((\bar{x},\bar{y}); \operatorname{gph} F) \big\}, \tag{4.3}$$

(4.4) $$\widehat{D}^*F(\bar{x},\bar{y})(y^*) := \big\{x^* \in X^* \big| \, (x^*,-y^*) \in \widehat{N}\big((\bar{x},\bar{y}); \operatorname{gph} F\big)\big\}.$$

If $F = f\colon X \to Y$ is *strictly differentiable* at $\bar{x}$ (in particular, if $f \in C^1$), then

$$D^*f(\bar{x})(y^*) = \widehat{D}^*f(\bar{x})(y^*) = \big\{\nabla f(\bar{x})^*y^*\big\}, \quad y^* \in Y^*,$$

i.e., both coderivatives (4.3) and (4.4) are positively homogeneous extensions of the classical *adjoint* derivative operator to nonsmooth and set-valued mappings.

Finally, consider a function $\varphi\colon X \to I\!R$ *locally Lipschitzian* around $\bar{x}$; in this paper we do not use more general functions. Then the (basic, limiting) *subdifferential* of φ at $\bar{x}$ is defined by

(4.5) $$\partial\varphi(\bar{x}) := \operatorname*{Lim\,sup}_{x \to \bar{x}} \widehat{\partial}\varphi(x),$$

where the sequential outer limit (1.6) of the *Fréchet subdifferential* mapping $\widehat{\partial}\varphi(\cdot)$ is given by

(4.6) $$\widehat{\partial}\varphi(x) := \Big\{x^* \in X^* \Big| \, \frac{\varphi(u)-\varphi(x)-\langle x^*, u-x\rangle}{\|u-x\|} \ge 0\Big\}.$$

We are not going to review in this section appropriate properties of the generalized differential constructions (4.1)–(4.6) used in Sections 5–7: these properties will be invoked with the exact references to [13] in the corresponding places of the proofs in the subsequent sections. Just note here that our basic/limiting constructions (4.1), (4.3), and (4.5) enjoy *full calculus* in the framework of Asplund spaces, while the Fréchet-like ones (4.2), (4.4), and (4.6) satisfy certain rules of "fuzzy calculus." Both of these calculi are employed in what follows. The reader can find some additional and related material in the books by Rockafellar and Wets [16], Smirnov [17], and Vinter [19] (concerning exact/full calculus in finite dimensions) and in the book by Borwein and Zhu [1] on fuzzy calculus in infinite dimensions; see also the references therein.

5 Optimality Conditions for Discrete Inclusions

In this section we derive necessary optimality conditions for the sequence of discrete approximation problems (P_N) defined in (3.1) and (3.3)–(3.7). We only present results in the "fuzzy" form, which are more convenient to derive necessary conditions for the original problem (P) by the limiting procedure in Section 6. "Pointwise" necessary conditions for (P_N) and for related discrete-time problems (not used in this paper) can be found in [14, Subsection 6.1.4].

Observe first that each discrete optimization problem (P_N) can be equivalently written in a special form of *constrained mathematical programming (MP)* problem in infinite-dimensional spaces:

(5.1) $$\begin{cases} \text{minimize } \ \psi_0(z) \ \text{ subject to} \\ \psi_j(z) \le 0, \quad j = 1,\ldots,s, \\ f(z) = 0, \\ z \in \Theta_j \subset Z, \quad j = 1,\ldots,l, \end{cases}$$

where ψ_j are real-valued functions on some Banach space Z, where $f\colon Z \to E$ is a mapping between Banach spaces, and where $\Theta_j \subset Z$. To see this, we let

$$z := (x_1,\ldots,x_N,v_0,\ldots,v_{N-1}) \in Z := X^{2N}, \ E := X^N, \ s := N+2+m+2r, \ l := N-1$$

and rewrite (P_N) as an (MP) problem (5.1) with the following data:

(5.2) $$\psi_0(z) := \varphi_0(x_N) + h_N \sum_{j=0}^{N-1} \vartheta(x_j,v_j) + \sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} \|v_j - \dot{\bar{x}}(t)\|^2\,dt,$$

$$
(5.3)\quad \psi_j(z) := \begin{cases} \|x_{j-1} - \bar{x}(t_{j-1})\| - \epsilon/2 \ \text{ for } \ j = 1, \ldots, N+1, \\[2mm] \displaystyle\sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \|v_i - \dot{\bar{x}}(t)\|^2 \, dt - \epsilon/2 \ \text{ for } \ j = N+2, \\[2mm] \varphi_i(x_N) - \ell\eta_N, \ \text{ for } \ j = N+2+i, \quad i = 1, \ldots, m+r, \\[2mm] -\varphi_i(x_N) - \ell\eta_N, \ \text{ for } \ j = N+2+m+r+i, \quad i = m+1, \ldots, m+r, \end{cases}
$$

$$
(5.4)\qquad \begin{cases} f(z) = \big(f_0(z), \ldots, f_{N-1}(z)\big) \ \text{ with} \\ f_j(z) := x_{j+1} - x_j - h_N v_j, \quad j = 0, \ldots, N-1, \end{cases}
$$

$$
(5.5)\qquad \Theta_j := \Big\{ z \in X^{2N} \Big| \ v_j \in F(x_j) \Big\} \ \text{ for } \ j = 0, \ldots, N-1
$$

in terms of the initial data of problem (P_N).

The next theorem establishes necessary optimality conditions for each problem (P_N) in the *approximate/fuzzy* form of refined *Euler-Lagrange* and *transversality inclusions* expressed in terms of Fréchet-like normals and subgradients. The proof is based on applying the corresponding *fuzzy calculus* rules and *neighborhood criteria* for *metric regularity and Lipschitzian behavior* of mappings from [13]. Note that fuzzy calculus rules provide representations of Fréchet subgradients and normals of sums and intersections at the reference points via those at points that are arbitrarily close to the reference ones. *Just for notational simplicity*, we suppose in the formulation and proof of the next theorem that these *arbitrarily close points reduce to the reference points in question.* This convention does not restrict the generality from the viewpoint of our main goal to derive necessary optimality conditions in the continuous-time problem (P). Indeed, the possible difference between the mentioned points obviously disappears in the limiting procedure.

Theorem 5.1 (fuzzy Euler-Lagrange conditions for discrete approximations). *Let* $\bar{x}_N(\cdot) = \{\bar{x}_N(t_j) | \ j = 0, \ldots, N\}$ *be local optimal solutions to problems* (P_N) *as* $N \to \infty$ *under the assumptions* (H1)–(H3) *with the Asplund state space* X. *Consider the quantities*

$$
(5.6)\qquad \theta_{Nj} := 2 \int_{t_j}^{t_{j+1}} \Big\| \frac{\bar{x}_N(t_{j+1}) - \bar{x}_N(t_j)}{h_N} - \dot{\bar{x}}(t) \Big\| \, dt, \quad j = 0, \ldots, N-1.
$$

Then there exists a sequence $\varepsilon_N \downarrow 0$ *along some* $N \to \infty$, *and there are sequences of Lagrange multipliers* λ_{iN}, $i = 0, \ldots, m+r$, *and adjoint trajectories* $p_N(\cdot) = \{p_N(t_j) \in X^* | \ j = 0, \ldots, N\}$ *satisfying the following relationships:*

—the sign and nontriviality conditions

$$
(5.7)\qquad \lambda_{iN} \ge 0 \ \textit{ for all } \ i = 0, \ldots, m+r, \quad \sum_{i=0}^{m+r} \lambda_{iN} = 1;
$$

—the complementary slackness conditions

$$
(5.8)\qquad \lambda_{iN} \Big[\varphi_i(\bar{x}_N(t_N)) - \ell\eta_N \Big] = 0 \ \textit{ for } \ i = 1, \ldots, m;
$$

—the extended Euler-Lagrange inclusion in the approximate form

$$
(5.9)\qquad \begin{aligned} &\Big(\frac{p_N(t_{j+1}) - p_N(t_j)}{h_N}, p_N(t_{j+1}) - \lambda_{0N} \frac{\theta_{Nj}}{h_N} b^*_{Nj} \Big) \in \lambda_{0N} \widehat{\partial} \vartheta \Big(\bar{x}_N(t_j), \frac{\bar{x}_N(t_{j+1}) - \bar{x}_N(t_j)}{h_N} \Big) \\ &+ \widehat{N}\Big(\Big(\bar{x}_N(t_j), \frac{\bar{x}_N(t_{j+1}) - \bar{x}_N(t_j)}{h_N} \Big); \operatorname{gph} F \Big) + \varepsilon I\!B^* \ \textit{ with } \ b^*_{Nj} \in I\!B^*, \quad j = 0, \ldots, N-1; \end{aligned}
$$

—*the approximate transversality inclusion*

$$-p_N(t_N) \in \sum_{i=0}^{m} \lambda_{iN}\widehat{\partial}\varphi_i(\bar{x}_N(t_N)) + \sum_{i=m+1}^{m+r} \lambda_{iN}\Big[\widehat{\partial}\varphi_i(\bar{x}_N(t_N)) \bigcup \widehat{\partial}(-\varphi_i)(\bar{x}_N(t_N))\Big] \tag{5.10}$$
$$+\varepsilon I\!B^*.$$

Proof. Consider problem (P_N) for any fixed $N \in N$ in the equivalent (MP) form (5.1) with the initial data (5.2)–(5.5). Denote

$$\bar{z} := \big(\bar{x}_N(t_1), \ldots, \bar{x}_N(t_N), \bar{v}_N(t_0), \ldots, \bar{v}_N(t_{N-1})\big)$$

and take N so large that the $W^{1,2}$-constraints (3.6) and (3.7) for $\bar{x}_N(\cdot)$ hold with the strict inequality, which is possible by Theorem 3.2. Thus the latter constraints can be simply *ignored* in what follows.

To prove the theorem, it is sufficient to examine the following two mutually exclusive cases, which completely cover the situation.

Case 1. Assume that the operator constraint mapping $f\colon X^{2N} \to X^N$ in (5.1) and (5.4) is *metrically regular* at $\bar{z}$ relative to the set $\Theta := \Theta_0 \cap \ldots \cap \Theta_{N-1}$ in (5.5) in the sense that there is a constant $\mu > 0$ and a neighborhood U of $\bar{z}$ satisfying the distance estimate

$$\operatorname{dist}(z; S) \le \mu \|f(z) - f(\bar{z})\| \text{ for all } z \in \Theta \cap U, \text{ where } S := \big\{z \in \Theta \big|\ f(z) = f(\bar{z})\big\}.$$

Then applying Ioffe's *exact penalization theorem* (see [14, Theorem 5.16]) and taking into account the specific structures of the inequality constraint functions ψ_j in (5.3) for $j = N+2+i$ as $i = 1, \ldots, m+r$, we conclude that $\bar{z}$ is a local optimal solution to the *unconstrained penalized problem:*

$$\text{minimize} \quad \max\big\{\psi_0(z) - \psi_0(\bar{z}),\ \max_{i \in I(\bar{x}_N)} \varphi_i(x_N)\big\} + \mu\big(\|f(z)\| + \operatorname{dist}(z; \Theta)\big), \tag{5.11}$$

for all $\mu > 0$ sufficiently large, where

$$I(\bar{x}_N) := \big\{i \in \{1, \ldots, m\} \big|\ \varphi_i(\bar{x}_N) = \ell\eta_N\big\}$$
$$\bigcup \big\{i \in \{m+1, \ldots, m+r\} \big|\ \text{either } \varphi_i(\bar{x}_N) = \ell\eta_N \text{ or } -\varphi_i(\bar{x}_N) = \ell\eta_N\big\}.$$

Applying the generalized Fermat rule [13, Proposition 1.114] to the local optimal solution $\bar{z}$ for (5.11), we arrive at the subdifferential inclusion

$$0 \in \widehat{\partial}\Big[\max\big\{\psi_0(\cdot) - \psi_0(\bar{z}),\ \max_{i \in I(\bar{x}_N)} \varphi_i(\cdot)\big\} + \mu\|f(\cdot)\| + \mu \operatorname{dist}(\cdot; \Theta)\Big](\bar{z}). \tag{5.12}$$

Fix any $\varepsilon > 0$ and employ in the fuzzy sum rule for Fréchet subgradients from [13, Theorem 2.33(b)] in (5.12). It gives (remember out notational convention)

$$0 \in \widehat{\partial}\Big[\max\big\{\psi_0(\cdot) - \psi_0(\bar{z}),\ \max_{i \in I(\bar{x}_N)} \varphi_i(\cdot)\big\}\Big](\bar{z}) + \mu\,\widehat{\partial}\|f(\cdot)\|(\bar{z}) + \mu\,\widehat{\partial}\operatorname{dist}(\bar{z}; \Theta) + (\varepsilon/4)I\!B^*.$$

Computing now by [13, Proposition 1.85] the Fréchet subdifferential of the distance function $\operatorname{dist}(\bar{z}; \Theta)$ and using the simple chain rule for the composition $\|f(z)\| = (\phi \circ f)(z)$ with $\phi(y) := \|y\|$ and the smooth mapping f from (5.4), we get

$$0 \in \widehat{\partial}\Big[\max\big\{\psi_0(\cdot) - \psi_0(\bar{z}),\ \max_{i \in I(\bar{x}_N)} \varphi_i(\cdot)\big\}\Big](\bar{z}) + \sum_{j=0}^{N-1} \nabla f_j(\bar{z})^* e_j^* + \widehat{N}(\bar{z}; \Theta) + (\varepsilon/4)I\!B^*$$

for some $e_j^* \in X^*$ with, by the structure of f in (5.4),

$$\sum_{j=0}^{N-1} \nabla f_j(\bar{z})^* e_j^* = \big(-e_0^*, e_0^* - e_1^*, \ldots, e_{N-2}^* - e_{N-1}^*, e_{N-1}^*, -h_N e_0^*, \ldots, -h_N e_{N-1}^*\big). \tag{5.13}$$

To proceed further, we use the fuzzy intersection rule from [13, Lemma 3.1] ensuring that

$$\widehat{N}(\bar{z};\Theta) \subset \widehat{N}(\bar{z};\Theta_0) + \ldots + \widehat{N}(\bar{z};\Theta_{N-1}) + (\varepsilon/4)I\!B^*$$

and also employ the fuzzy rule for Fréchet subgradients of the maximum function (cf. [13, Theorem 3.46] and its proof) giving, by the structure of the index set $I(\bar{x}_N)$, the inclusion

$$\begin{aligned}\widehat{\partial}\Big[\max\big\{\psi_0(\cdot)-\psi_0(\bar{z}),\ \max_{i\in I(\bar{x}_N)}\varphi_i(\cdot)\big\}\Big](\bar{z}) \quad &\subset \sum_{i=0}^{m}\lambda_{iN}\widehat{\partial}\varphi_i(\bar{x}_N)\\ &+\sum_{i=m+1}^{m+r}\lambda_{iN}\Big[\widehat{\partial}\varphi_i(\bar{x}_N)\bigcup\widehat{\partial}(-\varphi_i)(\bar{x}_N)\Big]+(\varepsilon/4)I\!B^*,\end{aligned}$$

where the multipliers λ_{iN}, $i = 0,\ldots,m+r$, satisfy the sign, nontriviality, and complementary slackness conditions in (5.7) and (5.8).

Applying the afore-mentioned fuzzy sum rule to the cost function (5.2) and taking into account the classical relationship

$$\partial\|\cdot\|^2(x) \subset 2\|x\|I\!B^* \text{ for any } x \in X,$$

as well as the subdifferentiation formula under the integral sign in (5.2) well known from convex analysis, we have the inclusion

$$\widehat{\partial}\psi_0(\bar{z}) \quad \subset \widehat{\partial}\varphi_0(\bar{x}_N) + h_N\sum_{j=0}^{N-1}\Big[\widehat{\partial}\vartheta(\bar{x}_j,\bar{v}_j) + \big(0, 2\theta_{Nj}I\!B^*\big)\Big] + (\varepsilon/4)I\!B^*,$$

where θ_{Nj} are defined in (5.6). Finally, choose $p_N(t_0) \in X^*$ arbitrarily and let

$$p_N(t_j) := e^*_{j-1}, \quad j = 1,\ldots,N,$$

with e^*_j given in (5.13). Then taking into account the special separated structures of the sets Θ_j in (5.5), we arrive at the Euler-Lagrange and transversality inclusions (5.9) and (5.10) with $\varepsilon_N = \varepsilon$ by substituting the corresponding fuzzy relationships derived above into the generalized Fermat stationary condition (5.12). This completes the proof of the theorem in Case 1.

Case 2. It remains to consider the situation when the mapping f from (5.4) is *not metrically regular* at $\bar{z}$ relative to the set intersection $\Theta := \Theta_0 \cap \ldots \cap \Theta_{N-1}$. Let us show that this *never holds*, along some subsequences $\varepsilon_N \downarrow 0$ and $N \to \infty$, under the local Lipschitzian assumption imposed on F.

Indeed, in this case the *restriction* $f_\Theta\colon X^{2N} \to X^N$ of f to Θ defined by

$$f_\Theta(z) := \begin{cases} f(z) & \text{if } z \in \Theta,\\ \emptyset & \text{otherwise}\end{cases}$$

is obviously *not* metrically regular *around* $\bar{z}$ in the sense of [13, Definition 1.47]. Then the *characterization* of the latter property from [13, Theorem 4.5] allows us, for any fixed $\varepsilon > 0$, to find $z \in \bar{z} + \varepsilon I\!B$ and $e^* = (e^*_0,\ldots,e^*_{N-1}) \in (X^*)^N$ such that

$$\|e^*\| > 1 \text{ and } 0 \in \widehat{D}^* f_\Theta(z)(e^*)$$

via the Fréchet coderivative (4.4) of f_Θ. Employing now the coderivative sum rule from [13, Theorem 1.62] and the fuzzy intersection rule from [13, Lemma 3.1], we get

$$0 \in \sum_{j=0}^{N-1}\nabla f_j(z)^* e^*_j + \sum_{j=0}^{N-1}\widehat{N}(z_j;\Theta_j) + \varepsilon I\!B^*$$

with some $z_j \in \Theta_j \cap (z + \varepsilon I\!B)$. According to our notation agreement, we put $z_j = z = \bar{z}$ for simplicity. Thus there are $z_j^* \in \widehat{N}(\bar{z}; \Theta_j)$ satisfying

$$(5.14) \qquad -\sum_{j=0}^{N-1} z_j \in \sum_{j=0}^{N-1} \nabla f_j(\bar{z})^* e_j^* + \varepsilon I\!B^*.$$

Taking into account calculation (5.13) due to the form (5.4) of the mapping f and the specific structures of the sets Θ_j in (5.5), we find from (5.14) dual elements

$$(x_{ij}^*, v_{ij}^*) \in \widehat{N}\Big(\Big(\bar{x}_N(t_j), \frac{\bar{x}_N(t_{j+1}) - \bar{x}_N(t_j)}{h_N}\Big); \operatorname{gph} F_j\Big), \quad j = 0, \ldots, N-1,$$

satisfying the relationships

$$\begin{cases} -x_{jj}^* - e_{j-1}^* + e_j^* \in \varepsilon I\!B^*, \quad j = 0, \ldots, N-1, \\ -v_{jj}^* + h_N e_j^* \in \varepsilon I\!B^*, \quad j = 0, \ldots, N-1, \\ -e_{N-1}^* \in \varepsilon I\!B^*. \end{cases}$$

Define now the adjoint discrete trajectory $p_N(t_j)$, $j = 0, \ldots, N$, as in Case 1. Then the above relationships ensure that the pair $\big(\bar{x}_N(\cdot), p_N(\cdot)\big)$ satisfies the Euler-Lagrange inclusion (5.9) and the transversality inclusion (5.10) with

$$\lambda_{iN} = 0 \quad \text{for all} \quad i = 0, \ldots, m + r$$

and the following nontriviality condition:

$$(5.15) \qquad \|p_N(t_1)\| + \ldots + \|p_N(t_N)\| \geq 1 \quad \text{for all large} \quad N \in I\!N.$$

Let us show that condition (5.15) *contradicts* (5.9) and (5.10) with $\lambda_{iN} = 0$ due the *locally Lipschitzian* property of F assumed in the theorem.

To proceed, we observe that the Euler-Lagrange inclusion (5.9) with $\lambda_{0N} = 0$ can be equivalently written as

$$\frac{p_N(t_{j+1}) - p_N(t_j)}{h_N} \in \widehat{D}^* F\Big(\bar{x}_N(t_j), \frac{\bar{x}_N(t_{j+1}) - \bar{x}_N(t_j)}{h_N}\Big)\big(-p_N(t_{j+1})\big) + \varepsilon I\!B^*, \quad j = 0, \ldots, N-1.$$

Then the local Lipschitzian property of F with modulus ℓ_F yields, by the *neighborhood coderivative characterization* of [13, Theorem 4.7], that

$$\|x_j^*\| \leq \ell_F \|v_j^*\| \quad \text{whenever} \quad x_j^* \in \widehat{D}^* F_j(x_j, v_j)(v_j^*)$$

and (x_j, v_j) near $\big(\bar{x}_N(t_j), [\bar{x}_N(t_{j+1}) - \bar{x}_N(t_j)]/h_N\big)$. Thus

$$\|p_N(t_{N-1})\| \leq \|p_N(t_N)\|(1 + h_N \ell_F) + \varepsilon h_N$$

and then, as a discrete counterpart of the Gronwall lemma,

$$(5.16) \qquad \|p_N(t_j)\| \leq \exp\big(\ell_F(b-a)\big)\|p_N(t_N)\| + \varepsilon(b-a) \quad \text{for all} \quad j = 0, \ldots, N.$$

Finally, take a sequence $\nu_k \downarrow 0$ as $k \to \infty$ and choose numbers N_k and ε_k such that

$$N_k := [1/\nu_k] \quad \text{and} \quad \varepsilon_k \leq \nu_k^2 \quad \text{as} \quad k \in I\!N,$$

where $[\cdot]$ stands for the greatest integer less than or equal to the given real number. By (5.16) and by the transversality condition (5.10) with $\lambda_{iN} = 0$ along the chosen sequences of $\varepsilon_k = \varepsilon_{N_k} \downarrow 0$ and $N_k \to \infty$ as $k \to \infty$, we have the estimate

$$\sum_{j=1}^{N_k} \|p_{N_k}(t_j)\| \leq \nu_k \exp\big(\ell_F(b-a)\big) + \nu_k(b-a) \downarrow 0 \quad \text{as} \quad k \in I\!N,$$

which contradicts (5.15) and completes the proof of the theorem. $\triangle$

6 Extended Euler-Lagrange Conditions for Relaxed Minimizers

In this section we derive necessary optimality conditions in the refined forms of the extended Euler-Lagrange and transversality inclusions for *relaxed* intermediate local minimizers of the original problem (P). The proof is based on the passing to the limit from the necessary optimality conditions for discrete approximation problems obtained in Section 5 and on the usage of the *strong stability* of discrete approximations established in Section 3. A crucial part of the proof involves the justification of an appropriate convergence of *adjoint arcs*; the latter becomes possible due to the *coderivative characterization* of Lipschitzian set-valued mappings taken from [13].

Theorem 6.1 (extended Euler-Lagrange and transversality inclusions for relaxed intermediate minimizers). *Let $\bar{x}(\cdot)$ be a relaxed intermediate local minimizer for the Bolza problem (P) given in* (1.1)–(1.4) *under the standing assumptions of Section 2, where the spaces X and X^* are Asplund. Then there are nontrivial Lagrange multipliers $0 \neq (\lambda_0, \ldots, \lambda_{m+r}) \in I\!R^{m+r+1}$ and an absolutely continuous mapping $p\colon [a,b] \to X^*$ such that the following necessary conditions hold:*

—the sign conditions

$$\lambda_i \geq 0 \text{ for all } i = 0, \ldots, m+r, \tag{6.1}$$

—the complementary slackness conditions

$$\lambda_i \varphi_i(\bar{x}(b)) = 0 \text{ for } i = 1, \ldots, m, \tag{6.2}$$

—the extended Euler-Lagrange inclusion, for a.e., $t \in [a,b]$,

$$\dot{p}(t) \in \operatorname{clco}\Big\{u \in X^* \Big| \, (u, p(t)) \in \lambda_0 \partial \vartheta(\bar{x}(t), \dot{\bar{x}}(t)) + N\big((\bar{x}(t), \dot{\bar{x}}(t)); \operatorname{gph} F\big)\Big\}, \tag{6.3}$$

—and the transversality inclusion

$$-p(b)) \in \sum_{i=0}^{m} \lambda_i \partial \varphi_i(\bar{x}(b)) + \sum_{i=m+1}^{m+r} \lambda_i \Big[\partial \varphi_i(\bar{x}(b)) \bigcup \partial(-\varphi_i)(\bar{x}(b))\Big]. \tag{6.4}$$

Proof. Given the intermediate local minimizer $\bar{x}(\cdot)$ to (P), employ Theorem 3.2 that ensures the strong $W^{1,2}$-approximation of $\bar{x}(\cdot)$ by a sequence of optimal solutions $\bar{x}_N(\cdot)$ to problems (P_N). Applying now the necessary optimality condition of Theorem 5.1, we find sequence of multipliers λ_{iN}, $i = 0, \ldots, m+r$, and adjoint trajectories $p_N(\cdot)$ satisfying conditions (5.7)–(5.10). Without loss of generality, we can and do suppose that

$$\lambda_{iN} \to \lambda_i \text{ as } N \to \infty \text{ for all } i = 0, \ldots, m+r,$$

where the limiting multipliers λ_i, $i = 0, \ldots, m+r$, are not zero simultaneously and satisfy the sign condition (6.1). Moreover, we get the complementarity slackness conditions (6.2) by passing to the limit in (5.8) with $\eta_N \to 0$ as $N \to \infty$.

Let us next justify the possibility of passing to the limit in the approximate Euler-Lagrange (5.9) and transversality (5.10) inclusions for the discrete-time problems (P_N). Having θ_{Nj} defined in (5.6), consider the corresponding sequence of functions $\theta_N\colon [a,b] \to I\!R$ given by

$$\theta_N(t) := \frac{\theta_{Nj}}{h_N} b^*_{Nj} \text{ for } t \in [t_j, t_{j+1}), \; j = 0, \ldots, N-1.$$

It follows from the strong $W^{1,2}$-convergence of Theorem 3.2 that

$$\begin{aligned} \int_a^b \|\theta_N(t)\| \, dt &\leq \sum_{j=0}^{N-1} \theta_{Nj} \leq 2 \sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} \left\| \frac{\bar{x}_N(t_{j+1}) - \bar{x}_N(t_j)}{h_N} - \dot{\bar{x}}_N(t) \right\| dt \\ &= 2 \int_a^b \|\dot{\bar{x}}_N(t) - \dot{\bar{x}}(t)\| \, dt \to 0, \end{aligned}$$

which allows us to suppose without loss of generality that

$$\dot{\bar{x}}_N(t) \to \dot{\bar{x}}(t) \text{ and } \theta_N(t) \to 0 \text{ a.e. } t \in [a,b] \text{ as } N \to \infty.$$

Furthermore, we derive from the approximate Euler-Lagrange condition (5.9) that there are

$$e^*_{N_j}, \widetilde{e}^*_{Nj} \in I\!B^* \text{ and } (x^*_{Nj}, v^*_{Nj}) \in \widehat{\partial}\vartheta\Big(\bar{x}_N(t_j), \frac{\bar{x}_N(t_{j+1}) - \bar{x}_N(t_j)}{h_N}\Big), \quad j = 0, \ldots, N-1,$$

such that the discrete-time inclusions

$$\Big(\frac{p_N(t_{j+1}) - p_N(t_j)}{h_N} - \lambda_{0N} x^*_{Nj}\Big) + \varepsilon_N e^*_{N_j}$$

$$\in \widehat{D}^* F\Big(\bar{x}_N(t_j), \frac{\bar{x}_N(t_{j+1}) - \bar{x}_N(t_j)}{h_N}\Big)\Big(\lambda_N v^*_{Nj} + \lambda_{0N}\frac{\theta_{Nj}}{h_N} b^*_{Nj} - p_N(t_{j+1}) + \varepsilon_N \widetilde{e}^*_{N_j}\Big)$$

hold for all $j = 0, \ldots, N-1$ and all $N \in I\!N$. Observe that, due to [13, Proposition 1.85], the sequences $\{(x^*_{Nj}, v^*_{jN})\}$ are uniformly bounded for all $j = 0, \ldots, N-1$ by the Lipschitz constant of ϑ. Since the mapping F is locally Lipschitzian with constant ℓ_F, we get by the coderivative condition for the Lipschitz continuity from [13, Theorem 1.43] that

$$\Big\|\frac{p_N(t_{j+1}) - p_N(t_j)}{h_N} - \lambda_N x^*_{Nj} + \varepsilon_N e^*_{N_j}\Big\| \le \ell_F \Big\|\lambda_N v^*_{Nj} + \lambda_N \frac{\theta_{Nj}}{h_N} b^*_{Nj} - p_N(t_{j+1}) + \varepsilon_N \widetilde{e}^*_{N_j}\Big\|$$

for $j = 0, \ldots, N-1$. This allows us to conclude that the piecewise extensions $p_N(t)$, $a \le t \le b$, of the adjoint discrete arcs $p_N(\cdot)$ are uniformly bounded on $[a,b]$ with

$$\|\dot{p}_N(t)\| \le \alpha + \beta\|\theta_N(t)\| \text{ a.e. } t \in [a,b], \tag{6.5}$$

where the positive numbers α and β are independent of N. Using now the Dunford criterion for the weak compactness in $L^1([a,b]; X^*)$ from [4, Theorem IV.1] (note that both X and X^* enjoy the RNP due to their Asplund property assumed) and arguing similarly to the proof of Theorem 3.2 above, we find an absolute continuous mapping $p\colon [a,b] \to X^*$ satisfying the Newton-Leibniz formula and such that $\dot{p}_N(\cdot) \to \dot{p}(\cdot)$ as $N \to \infty$ (with no loss of generality) in the weak topology of $L^1([a,b]; X^*)$.

Next we conclude from the approximate transversality inclusion (5.10), the sign and nontriviality conditions in (5.7), and the local Lipschitz continuity of φ_i, $i = 0, \ldots, m+r$, with the common constant ℓ from (H3) that

$$\|p_N(b)\| \le \ell(m+2) + 1 \text{ for sufficiently large } N \in I\!N$$

due to the uniform boundedness of Fréchet subgradients of locally Lipschitzian functions by [13, Proposition 1.85]. Since X is Asplund, this implies the weak* sequential compactness of $\{p_N(b)\}$ in X^*. Thus, passing to the limit in (5.10) as $N \to \infty$ and using definition (4.5) of the basic subdifferential, we arrive at the transversality inclusion (6.4).

Considering now the approximate Euler-Lagrange inclusion (5.9), we equivalently rewrite it as

$$\begin{aligned}\dot{p}_N(t) \in \Big\{u \in X^* \Big| \quad & \big(u, p_N(t_{j+1}) - \lambda_{0N}\theta_N(t)\big) \in \lambda_{0N}\widehat{\partial}\vartheta(\bar{x}_N(t_j), \dot{\bar{x}}_N(t)) \\ & + \widehat{N}\big((\bar{x}_N(t_j), \dot{\bar{x}}_N(t)); \operatorname{gph} F\big) + \varepsilon_N I\!B^*\Big\}\end{aligned} \tag{6.6}$$

for $t \in [t_j, t_{j+1})$ with $j = 0, \ldots, N-1$. Obverse, by the weak continuity of the Bochner integral in the Newton-Leibniz formula and by $\dot{p}_N(\cdot) \to \dot{p}(\cdot)$ weakly in $L^1[a,b]; X^*)$, that the values $p_N(t)$ converge to $p(t)$ weakly in X^*. Furthermore, the Mazur weak closure theorem ensures that some sequence of *convex combinations* of $\{\dot{p}_N(\cdot)\}$ converges to $\dot{p}(\cdot)$ strongly in $L^1([a,b]; X^*)$ as $N \to \infty$, and hence its subsequence converges to $\dot{p}(t)$ *almost everywhere* on $[a,b]$. Passing finally to the limit in (6.6) and taking into account the established *pointwise* convergence together with (6.5), we arrive at the extended Euler-Lagrange inclusion (6.3) and complete the proof of the theorem. $\triangle$

Note that the results obtained in Theorem 6.1 are different from those derived in [14, Subsection 6.1.5] not only by the *absence* of any SNC-like assumptions on the target/constraint set but also by *not* imposing the "coderivative normality" property on F needed in [14] in similar settings. Observe also that the arguments developed above allow us to provide the correspondent improvements in the case of Lipschitzian endpoint constraints of the Euler-Lagrange type necessary optimality conditions derived in [15] for evolution models governed by *semilinear inclusions*

$$\dot{x}(t) \in Ax(t) + F(x(t), t), \tag{6.7}$$

where A is an *unbounded* infinitesimal generator of a *compact C_0-semigroup* on X, and where continuous solutions to (6.7) are understood in the *mild* sense.

7 Euler-Lagrange and Maximum Conditions with No Relaxation

The main objective of this section is to derive necessary optimality conditions for intermediate local minimizers $\bar{x}(\cdot)$ of evolution inclusions *without any relaxation*. We show that it can be done under certain more restrictive assumptions on the initial data in comparison with those in Theorem 6.1. For simplicity, we consider here the *Mayer version* (P_M) of problem (P) with $\vartheta = 0$ in (1.1). In this case, the *Euler-Lagrange inclusion* (6.3) admits the *coderivative form*

$$\dot{p}(t) \in \operatorname{clco} D^*F(\bar{x}(t), \dot{\bar{x}}(t))\big(-p(t)\big) \quad \text{a.e. } t \in [a, b], \tag{7.1}$$

which easily implies, due to the extremal property for coderivatives of convex-valued mappings given in [13, Theorem 1.34], the *Weierstrass-Pontryagin maximum condition*

$$\langle p(t), \dot{\bar{x}}(t)\rangle = \max_{v \in F(\bar{x}(t))} \langle p(t), v\rangle \quad \text{a.e. } t \in [a, b] \tag{7.2}$$

provided that the sets $F(x)$ are *convex* near $\bar{x}(t)$ for a.e. $t \in [a, b]$. Our goal is to justify the above Euler-Lagrange and Weierstrass-Pontryagin conditions, together with the other necessary optimality conditions of Theorem 6.1, for intermediate minimizers of the Mayer problem (P_M) subject to the Lipschitzian endpoint constraints (1.3) and (1.4), *without any convexity or relaxation* assumptions and with *no SNC-like* requirements imposed on the endpoint constraint set. To accomplish this goal, we employ a certain approximation technique involving *Ekeland's variational principle* combined with other advanced results of variational analysis and generalized differentiation, which allow us to reduce the constrained problem under consideration to an unconstrained (and thus *stable with respect to relaxation*) Bolza problem studied in Section 6. However, this requires additional assumptions on the initial data of (P_M) imposed in what follows.

Recall that a set-valued mapping $F\colon X \rightrightarrows Y$ is *strongly coderivatively normal* at $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ if its normal coderivative (4.3) admits the representation

$$\begin{aligned} D^*F(\bar{x}, \bar{y})(y^*) = \Big\{x^* \in X^* \Big| \; &\exists \text{ sequences } (x_k, y_k) \to (\bar{x}, \bar{y}), \; x_k^* \xrightarrow{w^*} x^*, \text{ and } y_k^* \to y^* \\ &\text{with } y_k \in F(x_k) \text{ and } x_k^* \in \widehat{D}^*F(x_k, y_k)(y_k^*) \text{ as } k \to \infty\Big\} =: D_M^*F(\bar{x}, \bar{y})(y^*), \end{aligned} \tag{7.3}$$

where $D_M^*F(\bar{x}, \bar{y})$ is called the *mixed coderivative* of F at $(\bar{x}, \bar{y})$. Observe that the only difference between the normal and mixed coderivatives of F at $(\bar{x}, \bar{y})$ is that the *mixed* weak* convergence of $x_k^* \xrightarrow{w^*} x^*$ and the norm convergence of $y_k^* \to y^*$ is used for $D_M^*F(\bar{x}, \bar{y})$ in (7.3), in contrast to the weak* convergence of *both* components $(x_k^*, y_k^*) \xrightarrow{w^*} (x^*, y^*)$ for $D^*F(\bar{x}, \bar{y})$ in (4.3) via (4.1). Besides the obvious case of $\dim Y < \infty$, the strong coderivative normality holds in many important infinite-dimensional settings, and the property is preserved under various compositions; see [13, Proposition 4.9] describing major classes of mappings satisfying this property.

A mapping $F\colon X \rightrightarrows Y$ is called *sequentially normally compact* (SNC) at $(\bar{x},\bar{y}) \in \operatorname{gph} F$ if for any sequences $(x_k, y_k) \overset{\operatorname{gph} F}{\to} (\bar{x},\bar{y})$ and $(x_k^*, y^*) \in \widehat{N}\big((x_k,y_k); \operatorname{gph} F\big)$ one has

$$(x_k^*, y_k^*) \overset{w^*}{\to} 0 \Longrightarrow \|(x_k^*, y_k^*)\| \to 0 \text{ as } k \to \infty.$$

As discussed in Section 1, this property is a far-going extension of the "finite-codimension" and other related properties of sets and mappings. It always holds in finite dimensions, while in reflexive spaces agrees with the"compactly epi-Lipschitzian" property by Borwein and Strójwas; see [13] for more details, discussions, and calculus.

Finally, recall that the given norm on a Banach space X is *Kadec* if the strong and weak convergences agree on the boundary of the unit sphere of X. It is well known that every reflexive space admits an equivalent Kadec norm.

Theorem 7.1 (Euler-Lagrange and Weierstrass-Pontryagin conditions for intermediate local minimizers with no relaxation). *Let $\bar{x}(\cdot)$ be an intermediate local minimizer for the Mayer problem (P_M) in (1.1)–(1.4) under the standing hypotheses (H1) and (H3) on F and φ_i. Assume in addition that:*

(a) *the state space X is separable and reflexive with the Kadec norm on it;*

(b) *the velocity mapping F is SNC at $\big(\bar{x}(t), \dot{\bar{x}}(t)\big)$ and strongly coderivatively normal with weakly closed graph around this point for a.e. $t \in [a,b]$.*

Then there are nontrivial Lagrange multipliers $0 \neq (\lambda_0, \ldots, \lambda_{m+r}) \in I\!R^{m+r+1}$ and an absolutely continuous mapping $p\colon [a,b] \to X^*$ satisfying the following relationships:

—the sign and complementarity slackness conditions in (6.1) and (6.2);

—the Euler-Lagrange inclusion (7.1), where the closure operation is redundant;

—the Weierstrass-Pontryagin maximum condition (7.2); and

—the transversality inclusion (6.4).

Proof. Denote

$$\varphi_0^+(x,\nu) := \max\big\{\varphi_0(x) - \nu, 0\big\}, \quad \varphi_i^+(x) := \max\big\{\varphi_i(x), 0\big\} \text{ for } i = 1, \ldots, m \tag{7.4}$$

and, following the *method of metric approximations* [11], consider the parametric cost functional

$$\theta_\nu[x] := \Big[(\varphi_0^+)^2(x(b),\nu) + \sum_{i=1}^{m} (\varphi_i^+)^2(x(b)) + \sum_{i=m+1}^{m+r} \varphi_i^2(x(b))\Big]^{1/2} \text{ as } \nu \in I\!R \tag{7.5}$$

over trajectories for the evolution inclusion (1.1) with *no endpoint constraints.* Since $\bar{x}(\cdot)$ is an *intermediate local minimizer* for (P_M) and due to the constructions in (7.4) and (7.5), we have

$$\theta_\nu[x] > 0 \text{ for any } \nu < \bar{\nu} := \varphi_0\big(\bar{x}(b)\big) \tag{7.6}$$

provided that $x(\cdot)$ is a trajectory for (1.2) belonging to the prescribed $W^{1,1}$-neighborhood of the given intermediate local minimizer and such that $x(t) \in U$ for all $t \in [a,b]$, where the open set $U \subset X$ is taken from the requirements in (H1) imposed on $\bar{x}(\cdot)$.

Form as in [2] the space $\mathcal{X}$ of all the trajectories $x(\cdot)$ for (1.2) satisfying *the only constraint* $x(t) \in \operatorname{cl} U$ as $t \in [a,b]$ with the metric

$$d[x,y] := \int_a^b \|\dot{x}(t) - \dot{y}(t)\|\, dt.$$

We can easily check, based on Definition 2.1 of solutions to the original differential inclusion and on standard properties of the Bochner integral, that the metric space $\mathcal{X}$ is *complete* and that the function $\theta_\nu[\cdot]$ is (Lipschitz) continuous on $\mathcal{X}$ for any $\nu \in I\!R$. It follows from the above constructions that for every $\varepsilon > 0$ there is $\nu_\varepsilon < \bar{\nu}$ such that $\nu_\varepsilon \to \bar{\nu}$ as $\varepsilon \downarrow 0$ and

$$0 \le \theta_\varepsilon[\bar{x}] < \varepsilon \le \inf_{x \in \mathcal{X}} \theta_\varepsilon[x] + \varepsilon \quad \text{with} \quad \theta_\varepsilon := \theta_{\nu_\varepsilon}.$$

Now applying the classical *Ekeland variational principle*, find an arc $x_\varepsilon(\cdot) \in \mathcal{X}$ satisfying

$$d[x_\varepsilon, \bar{x}] \le \sqrt{\varepsilon} \quad \text{and} \quad \theta_\varepsilon[x] + \sqrt{\varepsilon} d[x, x_\varepsilon] \ge \theta_\varepsilon[x_\varepsilon] \quad \text{for all} \quad x \in \mathcal{X}. \tag{7.7}$$

This distance estimate yields that $x_\varepsilon(t) \in U$ as $t \in [a, b]$ and that $x_\varepsilon(\cdot)$ belongs to the fixed $W^{1,1}$-neighborhood of the intermediate local minimizer $\bar{x}(\cdot)$ whenever $\varepsilon > 0$ is sufficiently small. Hence $\theta_\varepsilon[x_\varepsilon] > 0$ for such ε by (7.6).

Given positive numbers ε and $\eta > 0$, we form the Bolza functional

$$J_{\varepsilon,\eta}[x] := \theta_\varepsilon[x] + \sqrt{\varepsilon} d[x, x_\varepsilon] + \eta \sqrt{1 + \ell_F^2} \int_a^b \operatorname{dist}\big((x(t), \dot{x}(t)); \operatorname{gph} F\big)\, dt$$

and show, following the proof of Claim in [14, Theorem 6.27], that there is a number $\eta \ge 1$ such that for every $\varepsilon \in (0, 1/\eta)$ the arc $x_\varepsilon(\cdot)$ built above is an *intermediate local minimizer* for this functional over all absolutely continuous mappings $x\colon [a, b] \to X$, *not necessarily trajectories* for (1.1), satisfying the constraints

$$x(a) = x_0 \quad \text{and} \quad x(t) \in U \quad \text{for} \quad t \in [a, b],$$

where the one $x(t) \in U$ can be *ignored* from the viewpoint of necessary optimality conditions, since the set U is open in X. Taking into account the structures of $\theta_\varepsilon[\cdot]$ and $d[\cdot, \cdot]$, we conclude that $x_\varepsilon(\cdot)$ is an intermediate minimizers for the following *unconstrained* Bolza problem with *Lipschitzian* data:

$$\text{minimize} \quad \varphi_\varepsilon\big(x(b)\big) + \int_a^b \vartheta_\varepsilon\big(x(t), \dot{x}(t), t\big)\, dt \tag{7.8}$$

over absolutely continuous arcs $x\colon [a, b] \to X$ satisfying $x(a) = x_0$ and lying in a $W^{1,1}$-neighborhood of $\bar{x}(\cdot)$, where the functions $\varphi_\varepsilon\colon X \to I\!R$ and $\theta_\varepsilon\colon X \times X \times [a, b] \to I\!R$ are given by

$$\varphi_\varepsilon(x) := \Big[(\varphi_0^+)^2(x, \nu_\varepsilon) + \sum_{i=1}^m \big(\varphi_i^+\big)^2(x) + \sum_{i=m+1}^{m+r} \varphi_i^2(x)\Big]^{1/2}, \tag{7.9}$$

$$\vartheta_\varepsilon(x, v, t) := \eta \sqrt{1 + \ell_F^2} \operatorname{dist}\big((x, v); \operatorname{gph} F\big) + \sqrt{\varepsilon} \|v - \dot{x}_\varepsilon(t)\|. \tag{7.10}$$

To apply the results of Theorem 6.1 to the case of problem (7.8), we first note that *every intermediate local minimizer* for the unconstrained problem (7.8) provides a *relaxed* intermediate local minimum for this problem. It follows from the *relaxation stability* of unconstrained Bolza problems with finite integrands, which is ensured by an appropriate infinite-dimensional extension of the classical Bogolyubov theorem valid under the assumptions made; see Lemma 2.3 above and its "intermediate" local counterpart given in [9, Theorem 4] whose proof holds in the infinite-dimensional setting under consideration. Furthermore, observe that, although Theorem 6.1 is presented for autonomous problems, its results hold true with no change for the case of *summable* integrands as in (7.10); it can be justified similarly to the proof of [14, Theorem 6.22] given for problems with geometric endpoint constraints. Finally, it follows from the proof of Theorem 6.1 that the compactness of the velocity sets assumed in (H1) is, in fact, *not needed* for the unconstrained and $W^{1,1}$-bounded framework of the Bolza problem (7.8).

Applying the optimality conditions of Theorem 6.1 to problem (7.8) with the initial data (7.9) and (7.10), for all small $\varepsilon > 0$, we find an absolutely continuous adjoint arc $p_\varepsilon\colon [a, b] \to X^*$ satisfying

$$\dot{p}_\varepsilon(t) \in \operatorname{co}\Big\{u \in X^* \Big|\ \big(u, p_\varepsilon(t)\big) \in \mu\, \partial \operatorname{dist}\big((x_\varepsilon(t), \dot{x}_\varepsilon(t)); \operatorname{gph} F\big) + \sqrt{\varepsilon}\big(0, I\!B^*\big)\Big\} \tag{7.11}$$

for a.e. $t \in [a,b]$ with $\mu := \eta\sqrt{1+\ell_F^2}$ and

$$(7.12)\qquad -p_\varepsilon(b) \in \partial\Big[(\varphi_0^+)^2(\cdot,\nu_\varepsilon) + \sum_{i=1}^{m}(\varphi_i^+)^2(\cdot) + \sum_{i=m+1}^{m+r}\varphi_i^2(\cdot)\Big]^{1/2}(x_\varepsilon(b)).$$

Note that the last term on the right-hand side of (7.11) appears due to employing the sum rule from [13, Theorem 2.33(c)] to the integrand (7.10) and using the well-known subdifferential formula for the norm function. The other difference between (7.10) and (6.3) is that (7.11) *does not contain the closure operation* as in (6.3). The norm-closure operation can be omitted in (7.11), since the basic subdifferential sets for Lipschitzian functions are weak compact in reflexive spaces (which are weakly compactly generated) by [13, Theorem 3.59(i)], and hence the right-hand side of (7.11) is closed in the norm topology of the dual space X^*.

To deal further with (7.11), fix $t \in [a,b]$ and consider the two possible cases for the location of $\big(x_\varepsilon(t),\dot x_\varepsilon(t)\big)$ relative to the graph of the velocity mapping $F(\cdot)$:

$$\textbf{(i)}\quad \big(x_\varepsilon(t),\dot x_\varepsilon(t)\big) \in \operatorname{gph} F \quad\text{and}\quad \textbf{(ii)}\quad \big(x_\varepsilon(t),\dot x_\varepsilon(t)\big) \notin \operatorname{gph} F.$$

In case (i) we use [13, Theorem 1.97] on basic subgradients of the distance function at set points, which gives the *approximate adjoint inclusion*

$$(7.13)\qquad \dot p_\varepsilon(t) \in \operatorname{co}\Big\{u \in X^*\Big|\ \big(u,p_\varepsilon(t)\big) \in N\big((x_\varepsilon(t),\dot x_\varepsilon(t));\operatorname{gph} F\big) + \sqrt{\varepsilon}(0,I\!B^*)\Big\}.$$

The out-of-set case (ii) is more involved and requires the *Kadec norm* structure of X together with the *weak closedness* assumption on the graph of F. Then we, by [13, Theorem 1.105], the relationship

$$\partial\operatorname{dist}\big((x_\varepsilon(t),\dot x_\varepsilon(t));\operatorname{gph} F\big) \subset \bigcup_{(x,v)\in\Pi\big((x_\varepsilon(t),\dot x_\varepsilon(t));\operatorname{gph} F\big)} N\big((x,v);\operatorname{gph} F\big)$$

via the *projection operator* $\Pi(\cdot;\operatorname{gph} F)$ at the reference point. Taking into account the a.e. pointwise convergence $\big(x_\varepsilon(t),\dot x_\varepsilon(t)\big) \to \big(\bar x(t),\dot{\bar x}(t)\big)$ as $\varepsilon \downarrow 0$ that follows from (7.7), we come up to a modified inclusion (7.13) with the replacement of $\big(x_\varepsilon(t),\dot x_\varepsilon(t)\big)$ by some sequence $(\widetilde x_\varepsilon,\widetilde v_\varepsilon) \xrightarrow{\operatorname{gph} F} \big(\bar x(t),\dot{\bar x}(t)\big)$ as $\varepsilon \downarrow 0$, while we keep the form (7.13) for simplicity.

Consider next the transversality condition (7.12) with φ_i^+ defined in (7.4). Employing the sum and chain rules [13, Subsection 3.2.1] for basic subgradients in (7.12) and taking into account relationships (7.5) and (7.6) with $\nu_\varepsilon \uparrow \bar\nu$ as $\varepsilon \downarrow 0$, we have

$$(7.14)\qquad -p_\varepsilon(b) \in \sum_{i=0}^{m}\lambda_{i\varepsilon}\partial\varphi_i\big(x_\varepsilon(b)\big) + \sum_{i=m+1}^{m+r}\lambda_{i\varepsilon}\Big[\partial\varphi_i\big(x_\varepsilon(b)\big)\bigcup\partial\big(-\varphi_i\big)\big(x_\varepsilon(b)\big)\Big],$$

where the multipliers $\lambda_{i\varepsilon}$ satisfy the conditions

$$(7.15)\qquad \lambda_{i\varepsilon} \ge 0 \ \text{ for all } \ i = 0,\ldots,m+r,\quad \sum_{i=0}^{m+r}\lambda_{i\varepsilon}^2 = 1 \ \text{ as } \ \varepsilon \downarrow 0.$$

By (7.15), we suppose without loss of generality that $\lambda_{i\varepsilon} \to \lambda_i$ as $\varepsilon \downarrow 0$ for $i = 0,\ldots,m+r$, where the limiting multipliers λ_i are not zero simultaneously and satisfy the sign and complementary slackness conditions in (6.1) and (6.2). Furthermore, it follows from (7.14) and (7.15) that the family $\{p_\varepsilon(b)\}_{\varepsilon>0}$ is uniformly bounded in X^* for ε sufficiently small. To proceed similarly to the proof of Theorem 6.1, we observe that the strong coderivative normality assumption on F allows us, by (7.13), to use the *mixed coderivative characterization* of the Lipschitz property of F from [13, Corollary 4.11] and thus to find an absolutely continuous arc $p\colon [a,b] \to X^*$ such that $\dot p_\varepsilon(\cdot) \to \dot p(\cdot)$ weakly in $L^1\big([a,b];X^*\big)$ and $p_\varepsilon(t) \to p(t)$ weakly in X^* as $\varepsilon \downarrow 0$ for each $t \in [a,b]$.

To complete the proof of the Euler-Lagrange and transversality inclusions of the theorem, we pass to the limit in (7.13) and (7.14) as $\varepsilon \downarrow 0$ by using the Mazur theorem on the strong convergence of convex combinations for $\{\dot{p}_\varepsilon(\cdot)\}$. To accomplish this limiting procedure and to arrive at the desired inclusions (7.1) and (6.4), we use the *closed-graph* property of the basic normal cone in (7.13) and the basic subdifferential in (7.14). This follows from [13, Theorem 3.60] due to the SNC assumption on F and the Lipschitz continuity of φ_i in the reflexive state space X. Observe that the closedness operation in (7.1) is redundant, similarly to (7.13), due to the *uniform boundedness* of $\{(\dot{p}_\varepsilon(\cdot), p_\varepsilon(\cdot))\}$ in $X^* \times X^*$ and the arguments above involving now [13, Theorem 3.59(ii)].

The given proof justifies the extended Euler-Lagrange and transversality conditions in the theorem for arbitrary intermediate local minimizers to problem (P_M) with *no relaxation.* In the general nonconvex setting the Euler-Lagrange inclusion (7.1) does not automatically imply the maximum condition (7.2). To establish the latter condition supplementing the other necessary conditions of the theorem, we follow the proof of [19, Theorem 7.4.1] given for a Mayer problem of the type (P_M) involving nonconvex differential inclusions in finite-dimensional spaces; it holds with minor changes in infinite-dimensions under the assumptions imposed. The proof of the latter theorem is based on reducing the constrained Mayer problem for nonconvex differential inclusions to an unconstrained Bolza (finite Lagrangian) problem, which in turn is reduced to a problem of optimal control with *smooth dynamics* and *nonsmooth endpoint constraints* first treated in [11] via the nonconvex normal cone (4.1) and the corresponding subdifferential (4.5) introduced therein to describe the appropriate transversality conditions in the maximum principle. $\triangle$

References

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

[2] F. H. Clarke, *Necessary conditions for a general control problem*, in Calculus of Variations and Control Theory, D. L. Russel, ed., Academic Press, New York, 1976, pp. 257–278.

[3] F. H. Clarke, *Necessary conditions in dynamic optimization*, Mem. Amer. Math. Soc. **173** (2005), No. 816.

[4] J. Diestel and J. J. Uhl, Jr., *Vector Measures*, American Mathematical Society, Providence, RI, 1977.

[5] F. S. De Blasi, G. Pianigiani, and A. A. Tolstonogov, *A Bogolyubov type theorem with a nonconvex constraint in Banach spaces*, SIAM J. Control Optim. **43** (2004), 466–476.

[6] K. Deimling, *Multivalued Differential Equations*, De Gruyter, Berlin, 1992.

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

[8] A. D. Ioffe, *Euler-Lagrange and Hamiltonian formalisms in dynamic optimization*, Trans. Amer. Math. Soc. **349** (1997), 2871–2900.

[9] A. D. Ioffe and R. T. Rockafellar, *The Euler and Weierstrass conditions for nonsmooth variational problems*, Calc. Var. Partial Diff. Eq. **4** (1996), 59–87.

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

[11] B. S. Mordukhovich, *Maximum principle in problems of time optimal control with nonsmooth constraints*, J. Appl. Math. Mech. **40** (1976), 960–969.

[12] B. S. Mordukhovich, *Discrete approximations and refined Euler-Lagrange conditions for nonconvex differential inclusions*, SIAM J. Control Optim. **33** (1995), 882–915.

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

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

[15] B. S. Mordukhovich and D. Wang, *Optimal control of semilinear evolution inclusions via discrete approximations*, Control Cybernet. **34** (2005), 849–870.

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

[17] G. V. Smirnov, *Introduction to the Theory of Differential Inclusions*, American Mathematical Society, Providence, RI, 2001.

[18] A. A. Tolstonogov, *Differential Inclusions in a Banach Space*, Kluwer, Dordrecht, The Netherlands, 2000.

[19] R. B. Vinter, *Optimal Control*, Birkhäuser, Boston, 2000.

[20] R. B. Vinter and P. D. Woodford, *On the occurrence of intermediate local minimizers that are not strong local minimizers*, Systems Cont. Lett. **31** (1997), 235–342.