10-1-2008

# Optimization of Delay-Differential Inclusions of Infinite Dimensions

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

Dong Wang
*Fayetteville State University, North Carolina*, dwang@uncfsu.edu

Lianwen Wang
*University of Central Missouri*, lwang@ucmo.edu

# OPTIMIZATION OF DELAY-DIFFERENTIAL INCLUSIONS IN INFINITE DIMENSIONS

BORIS S. MORDUKHOVICH, DONG WANG and LIANWEN WANG

WAYNE STATE UNIVERSITY

Detroit, MI 48202

Department of Mathematics

Research Report

2008 Series
#10

*This research was partly supported by the USA National Science Foundation*

# OPTIMIZATION OF DELAY-DIFFERENTIAL INCLUSIONS IN INFINITE DIMENSIONS

BORIS S. MORDUKHOVICH[1]
Department of Mathematics, Wayne State University
Detroit, MI, 48202, boris@math.wayne.edu

DONG WANG
Department of Mathematics and Computer Science, Fayetteville State University
Fayetteville, NC, 28301, dwang@uncfsu.edu

LIANWEN WANG
Department of Mathematics and Computer Science, University of Central Missouri
Warrensburg, MO, 64093, lwang@ucmo.edu

**Dedicated to the memory of Alex Rubinov**

**Abstract.** This paper concerns the study of dynamic optimization problems governed by delay-differential inclusions with finitely many equality and inequality endpoints constraints and multivalued initial conditions. We employ the method of discrete approximations and advanced tools of generalized differentiation in infinite-dimensional spaces to derive necessary optimality conditions in the extended Euler-Lagrange form.

**Key words.** dynamic optimization, variational analysis, delay-differential inclusions, functional endpoint constraints, multivalued initial conditions, discrete approximations, generalized differentiation, Banach and Asplund spaces, extended Euler-Lagrange conditions

**AMS subject classifications.** 49J52, 49J53, 49K25

**Abbreviated title.** Optimization of delay-differential inclusions

## 1 Introduction

The main objective of this paper is to study the *generalized Bolza problem* $(P)$ governed by *delay-differential inclusions* in *infinite dimensions* with finitely many *equality* and *inequality endpoint constraints* given by Lipschitzian functions and with *multivalued initial conditions.* The problem $(P)$ under consideration is formulated as follows.

Let $X$ be a Banach *state space*, let $[a,b] \subset I\!R$ be a fixed *time interval*, and let $x\colon [a-\Delta, b] \to X$ be a feasible trajectory of the constrained *delay-differential inclusion*

$$\dot{x}(t) \in F(x(t), x(t-\Delta), t) \quad a.e.\ t \in [a,b],\ x(a) = x_0 \in X, \tag{1.1}$$

$$x(t) \in C(t) \quad a.e.\ t \in [a-\Delta, a), \tag{1.2}$$

$$\varphi_i(x(b)) \le 0,\ i = 1, \dots, m, \tag{1.3}$$

$$\varphi_i(x(b)) = 0,\ i = m+1, \dots, m+r, \tag{1.4}$$

[1]Research of this author was partly supported by the US National Science Foundation under grants DMS-0304989 and DMS-0603846 and by the Australian Research Council under grant DP-0451168.

with a given *time delay* $\Delta > 0$, where $F\colon X \times X \times [a,b] \rightrightarrows X$ and $C\colon [a-\Delta, a] \rightrightarrows X$ are set-valued mappings defined the *system dynamics* and the *initial state conditions*, respectively, and where the functions $\varphi_i$, $i = 1, \ldots, m+r$, define the *endpoint constraints.*

By a *feasible arc* above we mean a mapping $x\colon [a-\Delta, b] \to X$ that is summable on $[a-\Delta, a]$, Fréchet differentiable for a.e. $t \in [a,b]$ satisfying the *Newton-Leibniz formula*

$$x(t) = x(a) + \int_a^t \dot{x}(s)\,ds \quad \text{for all} \quad t \in [a,b] \tag{1.5}$$

and all the constraints in (1.1)–(1.4), where the integral in (1.5) is taken in the *Bochner sense.* It is well known that for $X = I\!R^n$ the a.e. Fréchet differentiability and Newton-Leibnitz requirements on $x(t), a \le t \le b$, can be equivalently replaced by its *absolute continuity* in the standard sense. In fact, there is a full description of Banach spaces, where this equivalence holds true: they are spaces satisfying the so-called *Radon-Nikodým property* (RNP); see, e.g., [2]. The latter property is fulfilled, in particular, in any reflexive space.

Given now the *endpoint* cost function $\varphi_0\colon X \to I\!R$ and the *integrand* $f\colon X \times X \times X \times [a,b] \to I\!R$, we consider the *Bolza functional*

$$J[x] := \varphi_0(x(b)) + \int_a^b f(x(t), x(t-\Delta), \dot{x}(t), t)\,dt \tag{1.6}$$

and formulate the *dynamic optimization/optimal control* problem $(P)$ as

$$\text{minimize} \quad J[x] \quad \text{subject to} \quad (1.1) - (1.4) \tag{1.7}$$

over feasible arcs $x\colon [a-\Delta, b] \to X$ assuming that $J[x] > -\infty$ for all the feasible arcs and there is at least one feasible $x(\cdot)$ with $J[x] < \infty$.

It has been well recognized that the generalized Bolza problem $(P)$ is a convenient model in dynamic optimization unifying many other problems of this kind and containing, in particular, conventional parameterized forms of optimal control problems governed by *controlled delay-differential equations* of the type

$$\dot{x}(t) = g(x(t), x(t-\Delta), u, t), \quad u \in \mathcal{U}, \quad a.e.\ t \in [a,b]. \tag{1.8}$$

Besides more generality and other advantages of model (1.1) in comparison with that for (1.8), the direct inclusion description (1.1) allows us to cover the *closed-loop* case $\mathcal{U} = \mathcal{U}(x)$ in (1.8), which is among the *most challenging* in control theory and the most important for various applications. Note also that the presence of the *set-valued* mapping $C(\cdot)$ defined on the initial time interval $[a-\Delta, a)$ in (1.2) is a *specific feature of delay-differential systems* providing an additional source for optimizing the cost functional (1.6) by a choice of the initial condition $x(t) \in C(t)$ on $[a-\Delta, a)$.

The problem $(P)$ under consideration has been studied in [12] in the case of *finite-dimensional state spaces* $X = I\!R^n$; see also the references therein for previous developments on finite-dimensional delay-differential inclusions as well as the books [8, 14] for more discussions and references on a variety of approaches and results on *nondelayed counterparts* of problem $(P)$ and related *finite-dimensional* control systems. On the other hand, there are recent developments in [8, 9] for *nondelayed* differential and evolution inclusions with *infinite-dimensional* state spaces and various types of endpoint constraints. Finally, in our recent paper [11] we consider a counterpart of problem $(P)$ in infinite dimensions with general endpoint constraints in the *geometric* form

$$x(b) \in \Omega \subset X \tag{1.9}$$

instead of the *functional* ones given by in (1.3) and (1.4).

The major and most restrictive assumption of [11] imposes the *sequential normal compactness* (SNC) property on the target set $\Omega$, which is automatic when the space $X$ is finite-dimensional while cannot be easily checked in infinite-dimensional settings. Roughly speaking, the SNC property means that a set should be "sufficiently fat" around the point in question; in particular, it is never satisfied for *singletons* in every infinite-dimensional space. This property is closely related to the so-called *finite-codimension* property of convex sets, which is essential for the fulfillment of the appropriate versions of the Pontryagin maximum principle for infinite-dimensional problems of optimal control; see, e.g., [3, 5, 8] for more discussions and references.

The main result of this paper justifies *extended Euler-Lagrange necessary optimality conditions* for the formulated Bolza problem $(P)$ that are of the same type as in [11] with an appropriate subdifferential counterpart of the transversality inclusion, but *without any SNC* assumptions on the set of endpoint constraints given by *finitely many Lipschitzian functions*. The results obtained are extensions on the case of *delay* systems under consideration of those established in [9] for nondelayed infinite-dimensional inclusions providing at the same time certain *improvements* of [9] even in the nondelayed setting. Indeed, in contrast to [9], we consider here *nonautonomous* systems and use for them *extended version* of the limiting normal cone and subdifferential to describe *adjoint inclusions* in the corresponding necessary optimality conditions.

In comparison with [11] we derive necessary optimality conditions not just for global solutions to $(P)$ but in the essentially more subtle and difficult setting of *relaxed intermediate local minimizers* introduced here for the delay-differential problems with multivalued initial conditions following the scheme of [6] in the case of nondelayed differential inclusions. The treatment of local minimizers of this type requires a more delicate variational analysis performed in this paper.

The driving force of our approach to obtain necessary optimality conditions for continuous-time systems is the *method of discrete approximations* developed in [6] for finite-dimensional nondelayed inclusions and then extended in [8, 9, 11, 12] to more general settings.

The rest of the paper is organized as follows. In Section 2 we formulate the *standing assumptions* and then define and discuss the notions of *intermediate* local minimizers and *relaxed* intermediate local minimizers for the delayed problem $(P)$ under consideration.

Section 3 is devoted to the construction and justification of well-posed *discrete approximations* of intermediate local minimizers for problem $(P)$ with taking into account the Lipschitzian functional description of endpoint constraints in (1.3) and (1.4). Using further the possibility of *strong approximation* of feasible trajectories for $(P)$ by their discrete counterparts established in [11] and developing a certain relaxation procedure, we prove the $L^1/W^{1,1}$*-strong convergence* of optimal trajectories for discrete problems to the given relaxed intermediate local minimizer for the original problem $(P)$. This result requires appropriate *geometric assumptions* on the Banach state space $X$ in question that hold, in particular, when $X$ is *reflexive.*

In Section 4 we briefly overview the basic constructions of *dual-space generalized differentiation* (normals to sets, coderivatives of set-valued mappings, and subdifferentials of extended-real-valued functions) playing a fundamental role in the subsequent variational analysis and the derivation of necessary optimality conditions for discrete-time and continuous-time optimization problems.

Section 5 is devoted to deriving *necessary optimality conditions* for the discrete approximation problems constructed in Section 3, which are governed by *delay-difference inclusions* with Lipschitzian endpoint constraints in infinite-dimensional spaces. Our approach is based on reducing the *dynamic* discrete-time problems under consideration to the corresponding non-dynamic problems of *mathematical programming* that contain, along with Lipschitzian *functional constraints*, an increasing number of *geometric constraints* with possibly *empty interiors*. We obtain necessary optimality conditions for these problems by using advanced tools of variational analysis and generalized differential calculus in infinite dimensions. Finally, Section 6 presents the main result of the paper on the *Euler-Lagrange necessary optimality conditions* for relaxed intermediate local minimizers in the infinite-dimensional problem $(P)$ with Lipschitzian endpoint constraints *without SNC* assumptions on the initial data. These conditions are derived by passing to the limit from the "fuzzy" optimality conditions for the approximating delay-difference problems established in Section 5.

Our notation is basically standard; cf. [7, 8]. Unless otherwise stated, all the spaces considered are Banach with the norm $\|\cdot\|$ and the canonical pairing $\langle\cdot,\cdot\rangle$ between the space in question, say $X$, and its topological dual $X^*$ the weak* topology of which is denoted by $w^*$. We use the symbols $I\!B$ and $I\!B^*$ to signify the closed unit balls of the space in question 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

$$\mathop{\rm Lim\,sup}_{x\to\bar{x}} F(x) := \Big\{x^*\in X^*\Big|\ \exists \text{ sequences } x_k\to\bar{x},\ x_k^*\stackrel{w^*}{\to}x^* \text{ with } x_k^*\in F(x_k) \text{ as } k\in I\!N:=\{1,2,\ldots\}\Big\}. \tag{1.10}$$

## 2 Intermediate Minimizers and Relaxation

We begin this section with formulating the notion of *intermediate local minimizers* for problem $(P)$, which extends the original definition given in [6] (see also [8, Subsection 6.1.2]) from ordinary differential to delay-differential systems with multivalued initial conditions.

**Definition 2.1 (intermediate local minimizers for delay-differential systems).** *A feasible arc $\bar{x}\colon[a-\Delta,b]\to X$ is an* INTERMEDIATE LOCAL MINIMIZER *(i.l.m.) of rank $(r,p)\in[1,\infty)^2$ for $(P)$ if there are numbers $\varepsilon>0$, $\nu\ge 0$, and $\alpha\ge 0$ such that $J[\bar{x}]\le J[x]$ for all feasible arcs $x\colon[a-\Delta,b]\to X$ to $(P)$ satisfying the relationships*

$$\|x(t)-\bar{x}(t)\|<\varepsilon \ \textit{for all}\ t\in[a,b] \quad \textit{and} \tag{2.1}$$

$$\nu\int_{a-\Delta}^{a}\|x(t)-\bar{x}(t)\|^r\,dt+\alpha\int_a^b\|\dot{x}(t)-\dot{\bar{x}}(t)\|^p\,dt<\varepsilon. \tag{2.2}$$

Observe that relationships (2.1) and (2.2) mean that we consider in fact a neighborhood of $\bar{x}(t)$, $t\in[a-\Delta,b]$, in the Sobolev space $W^{1,p}([a,b];X)$ with the norm

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

on the main interval $[a,b]$ and in the classical Lebesgue space $L^r([a-\Delta,a];X)$ on the initial interval $[a-\Delta,a]$. The case of $\alpha=0$ for nondelayed systems ($\Delta=0$) with the only requirement (2.1) in Definition 2.1 clearly corresponds to the classical *strong* local minimum with respect to a neighborhood of $\bar{x}(\cdot)$ in the norm topology of $C([a,b];X)$. If instead of (2.2) with $\Delta=0$ we put the more restrictive $L^\infty$-norm requirement

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

we have the classical *weak* local minimum in the framework of Definition 2.1. Thus the notion introduced for the first time in Definition 2.1 for delay-differential systems with taking into account the *multivalued* initial condition (1.2) reduces to the notion of intermediate local minimizers given in [6] for ordinary differential inclusions and occupies, for any $p\in[1,\infty)$, an *intermediate* position between the classical concepts of strong and weak local minima. It has been well recognized that this notion is indeed *different* from both classical notions even for convex and autonomous nondelayed systems in finite dimensions; see [8] and the references therein. Of course, all the necessary conditions for intermediate minimizers automatically hold for strong (and hence for global) minimizers considered in [11] for the case of geometric endpoint constraints.

Let now $\bar{x}(\cdot)$ be an arbitrary *i.l.m.* for problem $(P)$. We impose the following *standing assumptions* on the the initial data of $(P)$ used throughout the whole paper:

**(H1)** The mapping $C\colon[a-\Delta,a]\rightrightarrows X$ is *compact-valued, uniformly bounded*

$$C(t)\subset M_C I\!B \ \text{ on } \ [a-\Delta,a] \ \text{ with some } \ M_C>0,$$

and *Hausdorff continuous* for a.e. $t\in[a-\Delta,a]$.

**(H2)** There are an open set $U\subset M_C I\!B$ and two positive numbers $L_F$ and $M_F$ such that $\bar{x}(t)\in U$ for any $t\in[a,b]$, the sets $F(x,y,t)$ are nonempty and *compact* for all $(x,y,t)\in U\times(M_C I\!B)\times[a,b]$, and the following inclusions

$$F(x,y,t)\subset M_F I\!B \quad \text{for all } (x,y,t)\in U\times(M_C I\!B)\times[a,b], \tag{2.3}$$

$$F(x_1,y_1,t)\subset F(x_2,y_2,t)+L_F(\|x_1-x_2\|+\|y_1-y_2\|)I\!B, \tag{2.4}$$

hold whenever $(x_1,y_1)$, $(x_2,y_2)\in U\times(M_C I\!B)$ and $t\in[a,b]$. Note that (2.3) means the *uniform boundedness* of $F(x,y,t)$ on $U\times(M_C I\!B)\times[a,b]$ while (2.4) signifies the local *Lipschitz continuity* of $F(\cdot,\cdot,t)$ around $(\bar{x}(t),\bar{x}(t-\Delta))$.

**(H3)** $F(x,y,\cdot)$ is *Hausdorff continuous* for a.e. $t\in[a,b]$ uniformly in $(x,y)\in U\times(M_C I\!B)$.

**(H4)** The endpoint cost function $\varphi_0$ and all the endpoint constraint functions $\varphi_i$, $i=1,\ldots,m+r$, are *locally Lipschitzian* around $\bar{x}(b)$ with the common Lipschitz constant $\ell\ge 0$.

**(H5)** The integrand $f(x,y,v,\cdot)$ is *continuous* for a.e. $t\in[a,b]$ and *bounded uniformly* with respect to $(x,y,v)\in U\times(M_C I\!B)\times(M_F I\!B)$; furthermore, there is $\mu>0$ such that $f(\cdot,\cdot,\cdot,t)$ is continuous on the set

$$A_\mu(t)=\big\{(x,y,v)\in U\times(M_C I\!B)\times(M_F+\mu)I\!B\big|\ v\in F(x,y,s) \ \text{ for some } \ s\in(t-\mu,t]\big\}$$

uniformly in $t\in[a,b]$.

It is easy to observe that the assumptions made allow us to conclude that the i.l.m. notion introduced in Definition 2.1 is *invariant* with respect to any $r, p \in [1, \infty)$. We use this in what follows.

To proceed further, along with the original problem $(P)$ consider its "relaxed" counterpart constructed in the way well understood in optimal control and variational analysis; see, e.g., the books [8, 13, 15]. Roughly speaking, the relaxed problem is obtained from $(P)$ by a *convexification* procedure with respect to the *velocity* variable. Let

$$f_F(x, y, v, t) := f(x, y, v, t) + \delta(v; F(x, y, t)),$$

where $\delta(\cdot; \Theta)$ stands for the *indicator function* of the set in question equal to 0 on $\Theta$ and to $\infty$ otherwise. Denote by $\widehat{f}_F(x, y, v, t)$ the *biconjugate* (second conjugate) function to $f_F$ in $v$, i.e.,

$$\widehat{f}_F(x, y, v, t) := (f_F)^{**}_v(x, y, v, t).$$

The *relaxed generalized Bolza problem* $(R)$ for the original problem $(P)$ governed by the delay-differential inclusions under consideration is defined as follows:

$$\text{minimize } \widehat{J}[x] := \varphi_0(x(b)) + \int_a^b \widehat{f}_F(x(t), x(t - \Delta), \dot{x}(t), t)\, dt \tag{2.5}$$

over feasible trajectories $x(t)$, $a - \Delta \leq t \leq b$, of the same class as for $(P)$ but to the *convexified* delay-differential inclusion

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

with the initial condition (1.2) and the endpoint constraints (1.3) and (1.4). As usual, the symbol "clco" in (2.6) stands for the *convex closure* of the set in question.

Close relationships between the original and relaxed problems have been well understood in the calculus of variations and control theory for both differential and delay-differential systems; see the aforementioned books and the references therein. In fact, these relationships involving a certain *relaxation stability* reflect the deep *hidden convexity* property inherent in continuous-time (nonatomic measure) dynamic systems defined by differential and integral operators due to the fundamental *Lyapunov-Aumann convex theorem* and its extensions; see [8, 13, 15] for more details.

A *local* version of relaxation stability regarding *intermediate minimizers* for the delay-differential Bolza problem $(P)$ is postulated as follows and is studied in this paper.

**Definition 2.2 (relaxed intermediate local minimizers for delay-differential systems).** *A feasible arc $\bar{x}(\cdot)$ to the Bolza problem $(P)$ is a* RELAXED INTERMEDIATE LOCAL MINIMIZER *(r.i.l.m.) of rank $(r, p) \in [1, \infty)^2$ 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: $J[\bar{x}] = \widehat{J}[\bar{x}]$.*

Similarly to the i.l.m. case, we conclude and use in what follows that the notion of relaxed intermediate local minimizers do *not* actually depend on rank $(r, p) \in [1, \infty)^2$ under the assumptions made. Also we always take $\nu = \alpha = 1$ in (2.2) for simplicity.

## 3 Discrete Approximations

In this section we present basic constructions of the *method of discrete approximations* in the theory of necessary optimality conditions for delay-differential inclusions following the scheme of [11] developed there for the case of geometric constraints. Here we make important modifications required for *intermediate local minimizers* and *Lipschitzian* functional endpoint constraints (1.3) and (1.4) under consideration.

Let us first construct discrete approximations of the delay-differential inclusion (1.1) by replacing the time-derivative in (1.1) by the *uniform Euler finite difference*:

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

To formalize this procedure, for any natural number $N \in I\!N$ take $t_j := a + jh_N$ for $j = -N, \ldots, k$ and $t_{k+1} := b$, where $h_N := \Delta/N$ and $k \in I\!N$ is defined by

$$a + kh_N \le b < a + (k+1)h_N. \tag{3.1}$$

Note that $t_{-N} = a - \Delta$, $t_0 = a$, and $h_N \to 0$ as $N \to \infty$. Then the sequence of *delay-difference inclusions* approximating (1.1) is constructed as follows:

$$\begin{cases} x_N(t_{j+1}) \in x_N(t_j) + h_N F(x_N(t_j), x_N(t_j - \Delta), t_j), \; j = 0, \ldots, k, \; x(t_0) = x_0, \\ x_N(t_j) \in C(t_j) \quad \text{for } j = -N, \ldots, -1. \end{cases} \tag{3.2}$$

The collection of vectors $\{x_N(t_j) \mid j = -N, \ldots, k+1\}$ satisfying (3.2) is called a *discrete trajectory*. The corresponding collection

$$\Big\{ \frac{x_N(t_{j+1}) - x_N(t_j)}{h_N} \Big| \; j = 0, \ldots, k \Big\}$$

is called a *discrete velocity*. We also consider the *extended discrete velocities* defined by

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

It follows from the definition of the Bochner integral that the corresponding *extended discrete trajectories* are given by

$$x_N(t) = x(a) + \int_a^t v_N(s)ds, \quad t \in [a, b],$$

on the main interval $[a, b]$ and by

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

on the initial tail interval $[a - \Delta, a)$. Observe that $\dot{x}_N(t) = v_N(t)$ for a.e. $t \in [a, b]$.

The next result, which plays a significant role in the method of discrete approximations, establishes the *strong approximation* of *any feasible trajectory* $\bar{x}(\cdot)$ to the original delay-differential inclusion given in (1.1) and (1.2) by extended feasible trajectories to its delay-difference counterpart (3.2) in the following sense: the approximation/convergence in the $W^{1,1}([a,b]; X)$*-norm* on the main interval $[a,b]$ and the one in the $L^1([a-\Delta,a];X)$-norm on the initial interval $[a-\Delta,a]$. Note that the strong $W^{1,1}$-convergence of extended discrete trajectories on $[a,b]$ implies not only their *uniform convergence* on this interval but also the *a.e. pointwise* convergence of their *derivatives* on $[a,b]$ along some subsequence of $\{N\}$ as $N\to\infty$. A detailed proof of this result is given in [11, Theorem 2.1] with more discussions therein.

**Lemma 3.1 (strong approximation by discrete trajectories).** *Let $\bar{x}(\cdot)$ be a feasible trajectory to* (1.1) *and* (1.2) *under assumptions* (H1)–(H3)*, where $X$ is an arbitrary Banach space. Then there is a sequence of solutions $\{z_N(t_j)\mid j=-N,\ldots,k+1\}$ to the delay-difference inclusions* (3.2) *such that the extended discrete trajectories $z_N(t)$, $t\in[a-\Delta,b]$, converge to $\bar{x}(\cdot)$ strongly in $L^1$ on $[a-\Delta,a]$ and strongly in $W^{1,1}$ on $[a,b]$ as $N\to\infty$.*

From now on we fix an arbitrary *relaxed intermediate local minimizer* $\bar{x}(\cdot)$ for problem $(P)$ considering the case of $r=p=2$ and $\nu=\alpha=1$ in Definition 2.1 and Definition 2.2 without loss of generality. Having a positive number $\varepsilon$ from the latter definitions and an open set $U$ from the assumptions in (H2), we always suppose that

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

and take a sequence $\{z_N(t), a-\Delta\le t\le b\}$ of the extended trajectories for the delay-difference inclusions (3.2) approximating $\bar{x}(\cdot)$ in the sense of Lemma 3.1. Denoting

$$\eta_N:=\max_{t\in[a,b]}\|z_N(t)-\bar{x}(t)\|\to 0\quad\text{as}\quad N\to\infty, \tag{3.3}$$

construct the sequence of *discrete approximation problems* $(P_N)$ as follows:

$$\begin{aligned}\text{minimize } J_N[x_N]:=\varphi_0(x_N(b))&+\sum_{j=-N}^{-1}\int_{t_j}^{t_{j+1}}\|x_N(t_j)-\bar{x}(t)\|^2dt\\ &+h_N\sum_{j=0}^{k}f\Big(x_N(t_j),x_N(t_j-\Delta),\frac{x_N(t_{j+1})-x_N(t_j)}{h_N},t_j\Big)\\ &+\sum_{j=0}^{k}\int_{t_j}^{t_{j+1}}\Big\|\frac{x_N(t_{j+1})-x_N(t_j)}{h_N}-\dot{\bar{x}}(t)\Big\|^2dt\end{aligned} \tag{3.4}$$

subject to the constraints

$$x_N(t_{j+1})\in x_N(t_j)+h_NF(x_N(t_j),x_N(t_j-\Delta),t_j),\quad j=0,\ldots,k,\ x(t_0)=x_0, \tag{3.5}$$

$$x_N(t_j)\in C(t_j),\ j=-N,\ldots,-1, \tag{3.6}$$

$$\varphi_i(x_N(t_{k+1}))\le\ell\eta_N,\ \text{for } i=1,\ldots,m, \tag{3.7}$$

$$-\ell\eta_N\le\varphi_i(x_N(t_{k+1}))\le\ell\eta_N,\ \text{for } i=m+1,\ldots,m+r, \tag{3.8}$$

$$\|x_N(t_j)-\bar{x}(t_j)\|\le\frac{\varepsilon}{2},\quad j=1,\ldots,k+1, \tag{3.9}$$

$$\sum_{j=-N}^{-1}\int_{t_j}^{t_{j+1}}\|x_N(t_j)-\bar{x}(t)\|^2dt\le\frac{\varepsilon}{2}, \tag{3.10}$$

$$\sum_{j=0}^{k}\int_{t_j}^{t_{j+1}}\Big\|\frac{x_N(t_{j+1})-x_N(t_j)}{h_N}-\dot{\bar{x}}(t)\Big\|^2dt\le\frac{\varepsilon}{2}. \tag{3.11}$$

The next theorem justifies the *existence* of optimal solutions $\bar{x}_N(\cdot)$ to the discrete approximation problems $(P_N)$ and their *strong convergence* to the reference r.i.l.m. $\bar{x}(\cdot)$ for the original problem $(P)$. The strong convergence $\bar{x}_N(\cdot) \to \bar{x}(\cdot)$ is understood in the same sense as in Lemma 3.1, i.e., as the norm convergence in $L^1$ on the initial tail interval $[a-\Delta, a]$ and as the norm convergence in $W^{1,1}$ on the main interval $[a, b]$. In fact, under the assumptions made in (H1) and (H2), the strong convergence above can be equivalently replaced by that in the norm of $L^r$ on $[a-\Delta, a]$ and in the norm of $W^{1,p}$ on $[a, b]$ for any $r, p \geq 1$.

In contrast to Lemma 3.1 held in the general Banach state space $X$, the main part (ii) of Theorem 3.2 established below requires additional geometric assumptions imposed on the Banach space $X$ in question. Namely, we assume that *both spaces $X$ and $X^*$ are Asplund*, which automatically holds if $X$ is *reflexive*. Recall that a Banach space $X$ is *Asplund* if every separable subspace of $X$ has a separable dual. This is a broad class of Banach spaces well investigated in geometric theory and widely applied to many aspects of variational analysis and generalized differentiation; see the books [1, 2, 7, 8] for more details, numerous results, and discussions. Recall a remarkable fact from the geometric theory of Banach spaces: $X$ is Asplund *if and only if* the dual space $X^*$ has the Radon-Nikodým property.

Furthermore, part (ii) of the next theorem requires additional technical assumptions on the initial data in the case of *set-valued* initial conditions (1.2):

**(H6)** *either* the set $C(t)$ is a singleton $\{c(t)\}$ for a.e. $t \in [a-\Delta, a]$; *or* the set $C(t)$ is convex for a.e. $t \in [a-\Delta, a]$, the mapping $F(x, y, t)$ is linear in $y$ for a.e. $t \in [a, a+\Delta]$, and the function $f(x, y, v, t)$ is convex in $(y, v)$ for a.e. $t \in [a, a+\Delta]$.

**Theorem 3.2 (strong convergence of discrete optimal solutions).** *Let $\bar{x}(\cdot)$ be the given relaxed intermediate local minimizer for the original Bolza problem $(P)$ with the Banach state space $X$, let $\{(P_N)\}$ as $N \in I\!N$ be a sequence of discrete approximation problems constructed above, and let the standing assumptions* (H1)–(H5) *be satisfied. Then the following assertions hold:*

**(i)** *For all $N \in I\!N$ sufficiently large problem $(P_N)$ admits an optimal solution.*

**(ii)** *If in addition both spaces $X$ and $X^*$ are Asplund and* (H6) *holds, then any sequence $\{\bar{x}_N(\cdot)\}$ of optimal solutions to $(P_N)$ extended to the continuous-time interval $[a-\Delta, b]$ converges to $\bar{x}(\cdot)$ as $N \to \infty$ in the $L^1$-norm topology on $[a-\Delta, a]$ and in the $W^{1,1}$-norm topology on $[a, b]$.*

**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 $z_N(\cdot)$ approximating the given minimizer $\bar{x}(\cdot)$ by Lemma 2.1 and show that it satisfies all the constraints (3.7)–(3.11) for large $N$. By assumption (H4) we have

$$|\varphi_i(z_N(t_{k+1})) - \varphi_i(\bar{x}(b))| \leq \ell \|z_N(t_{k+1}) - \bar{x}(t_{k+1})\| \leq \ell \eta_N \text{ for all } i = 1, \ldots, m+r$$

due to (3.3). This implies the fulfillment of the endpoint constraints (3.7) and (3.8) for $z_N(\cdot)$, since those in (1.3) and (1.4) hold for $\bar{x}(\cdot)$. The fulfillment of (3.9) for $z_N(\cdot)$ follows directly from the construction of $\eta_N \to 0$ in (3.3). Further, it is easy to check that

$$\sum_{j=-N}^{-1} \int_{t_j}^{t_{j+1}} \|z_N(t_j) - \bar{x}(t)\|^2 dt = \int_{a-\Delta}^{a} \|z_N(t) - \bar{x}(t)\|^2 dt =: \alpha_N$$

for the piecewise linear extension of $z_N(\cdot)$ to $[a-\Delta, a)$ and

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

for the piecewise linear extension of $z_N(\cdot)$ to $[a,b]$. By the aforementioned equivalence between the $L^1/W^{1,1}$ and $L^2/W^{1,2}$ convergence in Lemma 2.1, we have that $\alpha_N \to 0$ and $\beta_N \to 0$ as $N\to\infty$, which justifies the fulfillment of (3.10) and (3.11) 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)–(H5).

To justify further assertion (ii) of the theorem on the *strong convergence* of discrete optimal trajectories, we observe first that

$$\limsup_{N\to\infty} J_N[\bar{x}_N]\le J[\bar{x}] \tag{3.12}$$

in any Banach spaces, which can be proved similarly to [8, Theorem 6.13] by using the Lebesgue dominated convergence theorem for the Bochner integral held due to (H5). Let us show that (3.12) implies the claimed strong convergence $\bar{x}_N(\cdot)\to\bar{x}(\cdot)$ as $N\to\infty$ under (H6) and the Asplund property of both spaces $X$ and $X^*$. This clearly follows from the relation

$$\lim_{N\to\infty}\left[\rho_N:=\int_{a-\Delta}^{a}\|\bar{x}_N(t)-\bar{x}(t)\|^2\,dt+\int_a^b\|\dot{\bar{x}}_N(t)-\dot{\bar{x}}(t)\|^2\,dt\right]=0, \tag{3.13}$$

which we now prove by contradiction under the additional assumptions imposed.

Supposing that (3.13) does not hold, we get a number $\rho>0$ such that $\rho_N\to\rho$ as $N\to\infty$ with no loss of generality. Observe, by the discussions above, that both spaces $X$ and $X^*$ have the *Radon-Nikodým property.* Thus, applying the *Dunford weak compactness theorem* given, e.g., in [2, Theorem IV.I], we find $\widetilde{x}(\cdot)\in L^1([a-\Delta,a];X)$ and $v(\cdot)\in L^1([a,b];X)$ such that

$$\bar{x}_N(\cdot)\to\widetilde{x}(\cdot)\ \text{ weakly in }\ L^1([a-\Delta,a];X)\ \text{ and }\ \dot{\bar{x}}_N(\cdot)\to v(\cdot)\ \text{ weakly in } L^1([a,b];X) \tag{3.14}$$

as $N\to\infty$. It follows from [13, Theorem 3.4.2] that the sequence $\{\bar{x}_N(t), a\le t\le b\}$ is *relatively compact* in the *norm topology* of the space $C([a,b];X)$. Taking into account the *weak continuity* of the Bochner integral as an operator from $L^1([a,b];X)$ into $X$ and passing to the limit in the Newton-Leibniz formula for $\bar{x}_N(t)$, $a\le t\le b$, as $N\to\infty$ we conclude that $\widetilde{x}(\cdot)\in C([a,b];X)$ on $[a,b]$ and that $v(t)=\dot{\widetilde{x}}(t)$ for a.e. $t\in[a,b]$.

Let us show next that the limiting function $\widetilde{x}(t)$, $a-\Delta\le t\le b$, satisfies all the constraints in (1.2)–(1.4) and, furthermore, belongs to the prescribed neighborhood of the intermediate local minimizer $\bar{x}(\cdot)$ defined by (2.1) and (2.2) with $r=p=2$ and $\nu=\alpha=1$.

To check this for (1.2) on the initial interval $[a-\Delta,a]$, we employ to $\bar{x}_N(\cdot)$ on $[a-\Delta,a]$ the classical *Mazur theorem*, which ensures by the first relation in (3.14) the $L^1([a-\Delta,a];X)$-*norm convergence* to $\widetilde{x}(\cdot)$ of a sequence of *convex combinations* of $\bar{x}_N(\cdot)$. Since the latter convergence implies the a.e. *pointwise* on $[a-\Delta,a]$ convergence of a subsequence of these convex combinations and since the sets $C(t)$ are assumed to be Hausdorff continuous in (H1) and convex in (H6) for a.e. $t\in[a-\Delta,a]$, we conclude that $\widetilde{x}(\cdot)$ satisfies (1.1) by passing to the limit in (3.6) as $N\to\infty$. The fulfillment of the endpoint constraints (1.3) and (1.4) for $\widetilde{x}(\cdot)$ follows by passing to the limit in (3.7) and (3.8) for $\bar{x}_N(\cdot)$ with $t_{k+1} = b$ therein by taking into account the norm convergence $x_N(b) \to \widetilde{x}(b)$, the continuity of the endpoint functions $\varphi_i$, and the convergence $\eta_N \to 0$ as $N \to \infty$.

By passing to the limit in (3.9), we justify the intermediate minimum relation (2.1) for $\widetilde{x}(\cdot)$ since $\bar{x}_N(\cdot) \to \widetilde{x}(\cdot)$ in the norm topology of $C([a,b];X)$. To get the integral intermediate minimum relation (2.2) for $\widetilde{x}(\cdot)$, we pass to the limit in (3.10) and (3.11) as $N \to \infty$ by using subsequently the weak convergence in (3.14), the Mazur theorem for $\{\bar{x}_N(\cdot)\}$ in $L^1([a-\Delta,a];X)$ and for $\{\dot{\bar{x}}_N(\cdot)\}$ in $L^1([a,b];X)$, and the *weak lower semicontinuity* of the integral functionals

$$\int_{a-\Delta}^{a} \|\cdot - \bar{x}(t)\|^2\,dt \quad\text{and}\quad \int_a^b \|\cdot - \dot{\bar{x}}(t)\|^2\,dt$$

in the aforementioned spaces, respectively.

By using similar arguments, the structures of the cost functionals in (1.6) and (3.4), the additional assumptions on $F$ and $f$ together with the imposed standing assumptions, and the upper estimate (3.12) established above, we conclude by the construction of the relaxed problem $(R)$ in Section 2 that $\widetilde{x}(\cdot)$ is a *feasible arc* for $(R)$ satisfying the relations

$$\widehat{J}[\widetilde{x}] = \varphi_0(\widetilde{x}(b)) + \int_a^b \widehat{f}_F(\widetilde{x}(t), \widetilde{x}(t-\Delta), \dot{\widetilde{x}}(t), t)\,dt + \rho \le J[\bar{x}]. \tag{3.15}$$

Since we suppose that $\rho > 0$ and we have $J[\bar{x}] = \widehat{J}[\bar{x}]$, the inequality in (3.15) is *strict*, and thus we get $\widehat{J}[\widetilde{x}] < \widehat{J}[\bar{x}]$ that *contradicts* the choice of $\bar{x}(\cdot)$ as a *relaxed intermediate local minimizer* for $(P)$. Thus (3.13) holds, which justifies (ii) and completes the proof of the theorem. $\triangle$

## 4 Generalized Differentiation

A characteristic feature of the original problem $(P)$ as well as of its discrete counterpart $(P_N)$ is *intrinsic nonsmoothness* primarily due to the presence of dynamic constraints (1.1) and (3.5). In what follows we deal with nonsmoothness by using appropriate generalized differential constructions studied in detail in the book [7]. For the reader's convenience, we briefly review these constructions and some of their important properties in this section. Since the corresponding constructions are used in the paper only in Asplund spaces, we adjust the definitions to this setting.

The *normal cone* to a set $\Omega \subset X$ at its point $\bar{x} \in \Omega$ (known as the basic, limiting, or Mordukhovich normal cone) is defined by

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

via the sequential Painlevé-Kuratowski outer/upper limit (1.10) of the *prenormal/Fréchet normal cone* to $\Omega$ 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 the symbol $x \xrightarrow{\Omega} \bar{x}$ indicates that $x \to \bar{x}$ with $x \in \Omega$. Note that for convex sets $\Omega$ we have

$$N(\bar{x};\Omega) = \widehat{N}(\bar{x};\Omega) = \big\{x^* \in X^* \big|\ \langle x^*, x - \bar{x}\rangle \le 0 \quad \text{for all } x \in \Omega\big\}. \tag{4.3}$$

Given a *set-valued mapping* $F\colon X \rightrightarrows Y$ and a point $(\bar{x},\bar{y}) \in \operatorname{gph} F$, define the *basic coderivative* of $F$ at $(\bar{x},\bar{y})$ and the *Fréchet coderivative* of $F$ at this point by, respectively,

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

$$\widehat{D}^*F(\bar{x},\bar{y})(y^*) := \big\{x^* \in X^* \big|\ (x^*,-y^*) \in \widehat{N}((\bar{x},\bar{y}); \operatorname{gph} F)\big\}. \tag{4.5}$$

Note that both coderivatives (4.4) and (4.5) are positively homogeneous set-valued mappings from $Y^*$ to $X^*$. They both are single-valued and linear

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

if $F\colon X \to Y$ is single-valued and $C^1$ around $\bar{x}$, or merely strictly differentiable at this point.

Given now an *extended-real-valued function* $\varphi\colon X \to \overline{I\!R} := (-\infty,\infty]$ finite at $\bar{x}$, the (basic, limiting, Mordukhovich) *subdifferential* of $\varphi$ at $\bar{x}$ is defined by

$$\partial\varphi(\bar{x}) := \operatorname*{Lim\,sup}_{x \xrightarrow{\varphi} \bar{x}} \widehat{\partial}\varphi(x), \tag{4.6}$$

where $x \xrightarrow{\varphi} \bar{x}$ means that $x \to \bar{x}$ with $\varphi(x) \to \varphi(\bar{x})$, and where $\widehat{\partial}\varphi(x)$ stands for the *Fréchet subdifferential* of $\varphi$ at $x$ defined by

$$\widehat{\partial}\varphi(x) := \Big\{x^* \in X^* \Big|\ \liminf_{u \to x} \frac{\varphi(u) - \varphi(x) - \langle x^*, u - x\rangle}{\|u - x\|} \ge 0\Big\}. \tag{4.7}$$

Besides the above generalized differential constructions, we employ their extended limiting versions for *moving* (parameter-dependent) objects needed in the case of *nonautonomous* systems. Given a moving set $\Omega\colon T \rightrightarrows X$, the *extended normal cone* to $\Omega(\bar{t})$ at $\bar{x} \in \Omega(\bar{t})$ is defined by

$$N_+(\bar{x};\Omega(\bar{t})) := \operatorname*{Lim\,sup}_{(x,t) \xrightarrow{\operatorname{gph}\Omega} (\bar{x},\bar{t})} \widehat{N}(x;\Omega(t)). \tag{4.8}$$

Given a parameter-dependent function $\varphi\colon X \times T \to \overline{I\!R}$ finite at $(\bar{x},\bar{t})$, the *extended subdifferential* of $\varphi(\cdot,\bar{t})$ at $\bar{x}$ is defined by

$$\partial_+\varphi(\bar{x},\bar{t}) = \operatorname*{Lim\,sup}_{(x,t) \xrightarrow{\varphi} (\bar{x},\bar{t})} \widehat{\partial}\varphi(x,t), \tag{4.9}$$

where $\widehat{\partial}\varphi(\cdot,t)$ is taken with respect to $x$ under fixed $t$. Obviously, the extended normal cone (4.8) and the extended subdifferential (4.9) reduce to the basic objects (4.1) and (4.6) if, respectively, $\Omega(\cdot)$ and $\varphi(\cdot,t)$ are independent of $t$. In the recent paper [10], the reader can find more details about the latter extended generalized differential constructions and calculus rules for them.

## 5 Euler-Lagrange Conditions for Delay-Difference Inclusions

In this section we derive necessary conditions for optimal solutions to the discrete optimization problems $(P_N)$. We reduce these *discrete-time* dynamic optimization problems to problems of mathematical programming with functional, operator, and finitely many geometric constraints.

It is easy to observe that each discrete optimization problem $(P_N)$, for any fixed $N \in I\!N$ and the corresponding number $k \in I\!N$ defined in (3.1), can be equivalently written as the following problem of *mathematical programming* $(MP)$:

$$\begin{cases} \text{minimize } \phi_0(z) \quad \text{subject to} \\ \phi_j(z) \le 0,\ j = 1, \ldots, s, \\ g(z) = 0, \\ z \in \Theta_j \subset Z,\ j = 1, \ldots, l, \end{cases} \tag{5.1}$$

where $\phi_j$ are real-valued functions on the Banach space $Z := X^{N+2k+3}$, where $g\colon Z \to E$ is a mapping between Banach spaces, and where $\Theta_j \subset Z$. To see this, let

$$z^N = (x^N_{-N}, \ldots, x^N_{k+1}, y^N_0, \ldots, y^N_k) := (x^N(t_{-N}), \ldots, x^N(t_{k+1}), y^N(t_0), \ldots, y^N(t_k)) \in Z, \tag{5.2}$$

$E := X^N$, $s := k + 3 + m + 2r$, and $l := k + 2$, where $y^N_j := (x^N_{j+1} - x^N_j)/h_N$. Rewrite $(P_N)$ as an $(MP)$ problem (5.1) with the following data:

$$\begin{aligned} \phi_0(z^N) : &= \varphi_0(x^N_{k+1}) + \sum_{j=-N}^{-1} \int_{t_j}^{t_{j+1}} \|x^N_j - \bar{x}(t)\|^2 dt \\ &+ h_N \sum_{j=0}^{k} f(x^N_j, x^N_{j-N}, y^N_j, t_j) + \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} \|y^N_j - \dot{\bar{x}}(t)\|^2 dt, \end{aligned} \tag{5.3}$$

$$\phi_j(z^N) := \begin{cases} \|x^N_j - \bar{x}(t_j)\| - \dfrac{\varepsilon}{2}, \quad j = 1, \ldots, k+1, \\[2ex] \displaystyle\sum_{i=-N}^{-1} \int_{t_i}^{t_{i+1}} \|x^N_i - \bar{x}(t)\|^2 dt - \frac{\varepsilon}{2}, \quad j = k+2, \\[2ex] \displaystyle\sum_{i=0}^{k} \int_{t_i}^{t_{i+1}} \|y^N_i - \dot{\bar{x}}(t)\|^2 dt - \frac{\varepsilon}{2}, \quad j = k+3, \\[2ex] \varphi_i(x^N_{k+1}) - \ell\eta_N, \text{ for } j = k+3+i,\ i = 1, \ldots, m+r, \\[2ex] -\varphi_i(x^N_{k+1}) - \ell\eta_N, \text{ for } j = k+3+m+r+i,\ i = m+1, \ldots, m+r, \end{cases} \tag{5.4}$$

$$g(z^N) = (g_0(z^N), \ldots, g_k(z^N)) \ \text{ with } g_j(z^N) := x^N_{j+1} - x^N_j - h_N y^N_j, \quad j = 0, \ldots, k, \tag{5.5}$$

$$\Theta_j := \big\{(x^N_{-N}, \ldots, y^N_k) \mid \ x^N_j \in C(t_j)\big\}, \quad j = -N, \ldots, -1, \tag{5.6}$$

$$\Theta_j := \big\{(x^N_{-N}, \ldots, y^N_k) \mid \ y^N_j \in F(x^N_j, x^N_{j-N}, t_j)\big\}, \quad j = 0, \ldots, k. \tag{5.7}$$

The next theorem presents necessary conditions for optimal solutions to each problem $(P_N)$ in the *fuzzy/approximate* discrete-time forms of the *Euler-Lagrange* and *transversality* inclusions expressed in terms of the Fréchet-like generalized differential constructions reviewed in Section 4. The proof is based on applying the corresponding *fuzzy calculus* rules and *neighborhood criteria* for *metric regularity* and *Lipschitzian behavior* of mappings taken from [7]. 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 and convenience, we suppose in the formulation and proof of the next theorem that these arbitrary close points reduce to the reference ones in question. It makes no difference for the limiting procedure to derive the main necessary optimality conditions for constrained delay-differential inclusions given in Section 6.

**Theorem 5.1 (approximate Euler-Lagrange conditions for delay-difference inclusions).** *Let $\bar{z}^N(\cdot)$ be an optimal solution to problem $(P_N)$ with any fixed $N \in \mathbb{N}$ sufficiently large under the standing hypotheses* (H1)–(H5). *Denote $F_j := F(\cdot,\cdot,t_j)$ and $f_j := f(\cdot,\cdot,\cdot,t_j)$ and assume in addition that $X$ is Asplund and that the functions $\varphi_i$ and $f_j$ are Lipschitz continuous around $\bar{x}^N_{k+1}$ and $(\bar{x}^N_j, \bar{x}^N_{j-N}, \bar{y}^N_j)$, respectively, for $i = 0,\ldots,m+r$ and $j = 0,\ldots,k$. Consider the quantities*

$$\begin{cases} \theta^N_j := 2\displaystyle\int_{t_j}^{t_{j+1}} \left\| \frac{\bar{x}^N_{j+1} - \bar{x}^N_j}{h_N} - \dot{\bar{x}}(t) \right\| dt, \quad j = 0,\ldots,k, \\ \sigma^N_j := 2\displaystyle\int_{t_j}^{t_{j+1}} \|\bar{x}^N_j - \bar{x}(t)\|\, dt, \quad j = -N,\ldots,-1. \end{cases} \tag{5.8}$$

*Then for any sequence of positive numbers $\varepsilon_N \to 0$ as $N \to \infty$ there are sequences of Lagrange multipliers $\lambda^N_i$, $i = 0,\ldots,m+r$, and sequences of the discrete adjoint arcs $p^N_j \in X^*$, $j = 0,\ldots,k+1$, and $q^N_j \in X^*$, $j = -N,\ldots,k+1$, satisfying the following relationships:*

- *the sign and nontriviality conditions*

$$\lambda^N_i \geq 0 \quad \text{for all} \quad i = 0,\ldots,m+r, \quad \sum_{i=0}^{m+r} \lambda^N_i = 1; \tag{5.9}$$

- *the complementary slackness conditions*

$$\lambda^N_i \left[\varphi_i(\bar{x}^N_{k+1}) - \ell\eta_N\right] = 0 \text{ for } i = 1,\ldots,m; \tag{5.10}$$

- *the approximate Euler-Lagrange inclusion*

$$\begin{cases} \Big(\dfrac{p^N_{j+1} - p^N_j}{h_N}, \dfrac{q^N_{j-N+1} - q^N_{j-N}}{h_N}, -\dfrac{\lambda^N_0 \theta^N_j}{h_N} a^N_j + p^N_{j+1} + q^N_{j+1}\Big) \in \lambda^N_0 \widehat{\partial} f(\bar{x}^N_j, \bar{x}^N_{j-N}, \bar{y}^N_j, t_j) \\ + \widehat{N}((\bar{x}^N_j, \bar{x}^N_{j-N}, \bar{y}^N_j); \text{gph}\, F_j) + \varepsilon_N \mathbb{B}^* \quad \text{with some } a^N_j \in \mathbb{B}^*,\ j = 0,\ldots,k; \end{cases} \tag{5.11}$$

- *the approximate tail conditions*

$$\begin{cases} -\dfrac{q^N_{j+1} - q^N_j}{h_N} - \lambda^N_0 \dfrac{\sigma^N_j}{h_N} b^N_j \in \widehat{N}(\bar{x}^N_j; C(t_j)) + \varepsilon_N \mathbb{B}^* \quad \text{with some } b^N_j \in \mathbb{B}^*,\ j = -N,\ldots,-1, \\ q^N_j = 0, \quad j = k-N+1,\ldots,k+1; \end{cases} \tag{5.12}$$

- *and the approximate transversality inclusion*

$$-p^N_{k+1} \in \sum_{i=0}^{m} \lambda^N_i \widehat{\partial}\varphi_i(\bar{x}^N_{k+1}) + \sum_{i=m+1}^{m+r} \lambda^N_i \Big[\widehat{\partial}\varphi_i(\bar{x}^N_{k+1}) \bigcup \widehat{\partial}(-\varphi_i)(\bar{x}^N_{k+1})\Big] + \varepsilon_N h_N \mathbb{B}^*. \tag{5.13}$$

**Proof.** We basically follow the procedure developed in the proof of [8, Theorem 6.19] given for the case of discrete approximations of nondelayed differential inclusions with geometric endpoint constraints, while here we take into account new elements in the structure of the initial data in the constrained delay problem under consideration. We present a detailed proof of the theorem in the major case of metric regularity of operator constraints while referring the reader to our similar previous consideration in the remaining case, which does not actually incorporate the new specific features of the problem under consideration; see below.

Consider problem $(P_N)$ in the equivalent *mathematical programming* form (5.1) for the decision variable $z^N \in Z$ in (5.2) with the initial data defined in (5.3)–(5.7). Given $\varepsilon > 0$ in (3.9)–(3.11), take $N \in I\!\!N$ so large that constraints (3.9)–(3.11) hold as *strict* inequalities; this is ensured by Theorem 3.2. Then *all* the inequality constraints in (5.4) are *inactive* at the point

$$\bar{z}^N := (\bar{x}^N_{-N}, \ldots, \bar{x}^N_{k+1}, \bar{y}^N_0, \ldots, \bar{y}^N_k) := (\bar{x}^N(t_{-N}), \ldots, \bar{x}^N(t_{k+1}), \bar{y}^N(t_0), \ldots, \bar{y}^N(t_k)),$$

and thus the functions $\phi_j$, $j = 1, \ldots, k+3$, can be *ignored* in the arguments below.

Let us examine the following two mutually exclusive cases in the proof of the theorem, which are complemented to each other.

**Case 1.** Assume that the operator constraint mapping $g\colon X^{N+2k+3} \to X^{k+1}$ in (5.5) is *metrically regular* at $\bar{z}^N$ relative to the set

$$\Theta := \bigcap_{j=-N}^{k} \Theta_j, \tag{5.14}$$

with $\Theta_j$ taken from (5.6) and (5.7), in the sense that there is a constant $\mu > 0$ and a neighborhood $V$ of $\bar{z}^N$ such that the distance estimate

$$\text{dist}(z; S) \le \mu \|g(z) - g(\bar{z}^N)\| \quad \text{for all} \quad z \in \Theta \cap V \quad \text{with} \quad S := \{z \in \Theta \mid g(z) = g(\bar{z}^N)\}$$

is satisfied. Then, by Ioffe's exact penalization theorem from [4] (see also [8, Theorem 5.16]), we conclude that $\bar{z}^N$ is a local optimal solution to the unconstrained minimization problem:

$$\text{minimize} \;\; \max\big\{\phi_0(z) - \phi_0(\bar{z}^N), \max_{i \in I(\bar{x}_N)} \varphi_i(x^N_{k+1})\big\} + \mu\big(\|g(z)\| + \text{dist}(z; \Theta)\big) \tag{5.15}$$

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

$$I(\bar{x}_N) := \Big\{ i \in \{1, \ldots, m\} \Big| \; \varphi_i(\bar{x}^N_{k+1}) = \ell\eta_N \Big\} \bigcup \Big\{ \;\; i \in \{m+1, \ldots, m+r\} \Big| \; \text{either} \;\; \varphi_i(\bar{x}^N_{k+1}) = \ell\eta_N$$

$$\text{or} \;\; -\varphi_i(\bar{x}^N_{k+1}) = \ell\eta_N \Big\}.$$

Applying the generalized Fermat rule from [7, Proposition 1.114] to the local optimal solution $\bar{z}^N$ for (5.15), we arrive at the subdifferential inclusion

$$0 \in \widehat{\partial}\big[\max\{\phi_0(\cdot) - \phi_0(\bar{z}^N), \max_{i \in I(\bar{x}_N)} \varphi_i(\cdot)\} + \mu\|g(\cdot)\| + \mu\text{dist}(\cdot, \Theta)\big](\bar{z}^N). \tag{5.16}$$

in terms of the Fréchet subdifferential (4.7). Picking then any sequence $\varepsilon_N \downarrow 0$ as $N \to \infty$ and employing in (5.16) the fuzzy sum rule for Fréchet subgradients from [7, Theorem 2.33(b)], we have by taking into account our notational convention that

$$0 \in \widehat{\partial}\big[\max\{\phi_0(\cdot) - \phi_0(\bar{z}^N), \max_{i \in I(\bar{x}_N)} \varphi_i(\cdot)\}\big](\bar{z}^N) + \mu\widehat{\partial}\|g(\cdot)\|(\bar{z}^N) + \mu\widehat{\partial}\text{dist}(\cdot; \Theta)(\bar{z}^N) + \frac{\varepsilon_N h_N}{4} I\!\!B^*.$$

Computing now by [7, Proposition 1.85] the Fréchet subdifferential of the distance function $\text{dist}(\bar{z};\Theta)$ and using the simple chain rule for the composition $\|g(z)\| = (\phi \circ g)(z)$ with $\phi(y) := \|y\|$ and the smooth mapping $g$ from (5.5), we get

$$0 \in \widehat{\partial}\big[\max\{\phi_0(\cdot) - \phi_0(\bar{z}^N), \max_{i\in I(\bar{x}_N)} \varphi_i(\cdot)\}\big](\bar{z}^N) + \sum_{j=0}^{N-1} \nabla g_j(\bar{z}^N)^* e_j^* + \widehat{N}(\bar{z}^N;\Theta) + \frac{\varepsilon_N h_N}{4} I\!B^* \quad (5.17)$$

for some $e_j^* \in X^*$ satisfying

$$\sum_{j=0}^{k} \nabla g_j(\bar{z}^N)^* e_j^* = \big(0,\ldots,0,-e_0^*, e_0^* - e_1^*,\ldots, e_{k-1}^* - e_k^*, e_k^*, -h_N e_0^*,\ldots,-h_N e_k^*\big) \quad (5.18)$$

due to the specific structure of the operator constraints in (5.1) and (5.5). By the fuzzy rule for Fréchet subgradients of the maximum function from [7, Theorem 3.46] we have the inclusion

$$\widehat{\partial}\Big[\max\{\phi_0(\cdot) - \phi_0(\bar{z}^N), \max_{i\in I(\bar{x}_N)} \varphi_i(\cdot)\}\Big](\bar{z}^N) \subset \lambda_0^N \widehat{\partial}\phi_0(\bar{z}^N) + \sum_{i=1}^{m} \lambda_i^N \widehat{\partial}\varphi_i(\bar{x}_{k+1}^N)$$
$$+ \sum_{i=m+1}^{m+r} \lambda_i^N \Big[\widehat{\partial}\varphi_i(\bar{x}_{k+1}^N) \bigcup \widehat{\partial}(-\varphi_i)(\bar{x}_{k+1}^N)\Big] + \frac{\varepsilon_N h_N}{4} I\!B^*, \quad (5.19)$$

where the multipliers $\lambda_i^N, i = 0,\ldots,m+r$, satisfy the sign, nontriviality, and complementary slackness conditions in (5.9) and (5.10). Taking into account the structure of cost functional $\phi_0$ in (5.3) and the specific forms of its terms, we get from the aforementioned fuzzy sum rule that

$$\begin{aligned}\widehat{\partial}\phi_0(\bar{z}^N) &\subset \widehat{\partial}\varphi_0(\bar{x}_{k+1}^N) + \sum_{j=-N}^{-1} \Big[\int_{t_j}^{t_{j+1}} 2\|\bar{x}_j^N - \bar{x}(t)\|\,dt\Big] I\!B^* \\ &+ h_N \sum_{j=0}^{k} \widehat{\partial} f(\bar{x}_j^N, \bar{x}_{j-N}^N, \bar{y}_j^N, t_j) + \sum_{j=0}^{k} \Big[\int_{t_j}^{t_{j+1}} 2\|\bar{y}_j^N - \dot{\bar{x}}(t)\|\,dt\Big] I\!B^* + \frac{\varepsilon_N h_N}{4} I\!B^*,\end{aligned} \quad (5.20)$$

where the Fréchet subdifferential of the function $f$ is considered with respect to its all but $t$ variables, and where the classical relationship $\partial\|\cdot\|^2(x) \subset 2\|x\| I\!B^*$ is used together with the subdifferentiation formula under the integral sign in (5.3) well known from convex analysis.

Apply further the fuzzy intersection rule from [7, Lemma 3.1] to the set $\Theta$ in (5.14) and get

$$\widehat{N}(\bar{z}^N;\Theta) \subset \widehat{N}(\bar{z}^N;\Theta_{-N}) + \ldots + \widehat{N}(\bar{z}^N;\Theta_k) + \frac{\varepsilon_N h_N}{4} I\!B^*. \quad (5.21)$$

Let $z_j^* = (x_{-N,j}^*,\ldots,x_{k+1,j}^*, y_{0,j}^*,\ldots,y_{k,j}^*)$ and observe from the set structures in (5.6) that for any $z_j^* \in \widehat{N}(\bar{z}^N;\Theta_j)$, $j = -N,\ldots,-1$, all but one components of $z_j^*$ are zero with the remaining one satisfying $x_{j,j}^* \in \widehat{N}(\bar{x}_j^N; C(t_j))$, $j = -N,\ldots,-1$. Similarly the relationship $z_j^* \in \widehat{N}(\bar{z}^N;\Theta_j)$ for $j = 0,\ldots,k$ implies that

$$(x_{j,j}^*, x_{j-N,j}^*, y_{j,j}^*) \in \widehat{N}((\bar{x}_j^N, \bar{x}_{j-N}^N, \bar{y}_j^N); \text{gph}\, F_j), \quad j = 0,\ldots,k, \quad (5.22)$$

with all the other components of $z_j^*$, $j = 0,\ldots,k$, equal to zero. Combining these relationships with (5.17)–(5.21) and using the notation

$$u_{k+1}^N \in \widehat{\partial}\varphi_0(\bar{x}_{k+1}^N), \quad (v_j^N, \kappa_{j-N}^N, w_j^N) \in \widehat{\partial} f(\bar{x}_j^N, \bar{x}_{j-N}^N, \bar{y}_j^N, t_j),\ j = 0,\ldots,k,$$

with $\bar{y}_j^N = (\bar{x}_{j+1}^N - \bar{x}_j^N)/h_N$ due to $g(\bar{z}^N) = 0$ in (5.5), we arrive at

$$\begin{cases} -x_{j,j}^* - x_{j,j+N}^* \in \lambda_0^N h_N \kappa_j^N + \lambda_0^N \sigma_j^N \mathbb{B}^* + \varepsilon_N h_N \mathbb{B}^*, \quad j = -N, \ldots, -1, \\ -x_{j,j}^* - x_{j,j+N}^* \in \lambda_0^N h_N \kappa_j^N + \lambda_0^N h_N v_j^N + e_{j-1}^* - e_j^* + \varepsilon_N h_N \mathbb{B}^*, \quad j = 1, \ldots, k-N, \\ -x_{j,j}^* \in \lambda_0^N h_N v_j^N + e_{j-1}^* - e_j^* + \varepsilon_N h_N \mathbb{B}^*, \quad j = k-N+1, \ldots, k, \\ -y_{j,j}^* \in \lambda_0^N h_N w_j^N + \lambda_0^N \theta_j^N \mathbb{B}^* - h_N e_j^* + \varepsilon_N h_N \mathbb{B}^*, \quad j = 0, \ldots, k, \\ 0 \in x_{k+1,k+1}^* + \lambda_0^N u_{k+1}^N + e_k^* + \varepsilon_N h_N \mathbb{B}^*, \\ -x_{0,0}^* \in \lambda_0^N h_N \kappa_0^N + \lambda_0^N h_N v_0^N - e_0^* + h_N \varepsilon_N \mathbb{B}^*, \end{cases} \tag{5.23}$$

where the triples $(x_{j,j}^*, x_{j-N,j}^*, y_{j,j}^*)$ satisfy (5.22) for all $j = 0, \ldots, k$ and where

$$x_{k+1,k+1}^* \in \sum_{i=1}^{m} \lambda_i^N \widehat{\partial}\varphi_i(\bar{x}_{k+1}^N) + \sum_{i=m+1}^{m+r} \lambda_i^N \Big[\widehat{\partial}\varphi_i(\bar{x}_{k+1}^N) \bigcup \widehat{\partial}(-\varphi_i)(\bar{x}_{k+1}^N)\Big]. \tag{5.24}$$

Further, we denote

$$\begin{cases} \widetilde{p}_j^N := e_{j-1}^* \ \text{ for } \ j = 1, \ldots, k+1, \quad \widetilde{q}_j^N := \lambda_0^N \kappa_j^N + \dfrac{x_{j,j+N}^*}{h_N} \ \text{ for } \ j = -N, \ldots, k-N, \\ \widetilde{q}_j^N := 0 \ \text{ for } \ j = k-N+1, \ldots, k+1 \end{cases} \tag{5.25}$$

and define the the *adjoint discrete trajectories* $(p_j^N, q_j^N)$ by

$$\begin{cases} q_{k+1}^N := 0, \quad q_j^N := q_{j+1}^N - \widetilde{q}_j^N h_N \ \text{ for } \ j = -N, \ldots, k+1, \\ p_0^N := -q_0^N, \ \text{and } p_j^N := \widetilde{p}_j^N - q_j^N \ \text{ for } \ j = 1, \ldots, k+1. \end{cases} \tag{5.26}$$

It is easy to check that $q_j^N = 0$ for $j = k-N+1, \ldots, k+1$. Combining finally the relationships and notation (5.22)–(5.26), we get the optimality conditions (5.11)–(5.13) of the theorem along an arbitrarily chosen sequence $\varepsilon_N \downarrow 0$ as $N \to \infty$. This completes the proof of the theorem in Case 1.

**Case 2.** It remains to consider the situation when the mapping $g$ from (5.5) is *not metrically regular* at $\bar{z}^N$ relative to the set $\Theta$. In this case the restriction

$$g_\Theta(z) := \begin{cases} g(z) & \text{if } \ z \in \Theta, \\ \emptyset & \text{otherwise} \end{cases} \tag{5.27}$$

of the mapping $g$ on the set $\Theta$ from (5.14) is *not* metrically regular *around* $\bar{z}_N$ in the standard sense; see, e.g. [7, Definition 1.47]. Observe that neither $g$ nor $\Theta$ involves the functional constraints of the problems $(P_N)$ and (5.1) under consideration. Thus we can proceed as in the proofs of [8, Theorem 6.19] and [11, Theorem 5.1] for the geometric constraint cases therein and, employing the *neighborhood characterizations* of the metric regularity and Lipschitz-like properties from [7, Theorem 4.5 and Theorem 4.7] as well as calculations similar to the above Case 1, arrive at the conclusions of the theorem in the remaining case. This completes the proof of the theorem. $\triangle$

## 6 Euler-Lagrange Conditions for Delay-Differential Inclusions

The concluding section of the paper presents the main result on new necessary optimality conditions for *relaxed intermediate local minimizers* in the *delay-differential systems* under consideration given in the *extended Euler-Lagrange form.* These conditions and their proof are based on passing to the limit from the optimality conditions for discrete approximations obtained in Section 5 with the use of the well-posedness/strong convergence results for discrete approximations established in Section 3 and special properties of the generalized differential constructions reviewed in Section 4 that allow us to justify the appropriate convergence of *adjoint trajectories.*

In this section we keep the standing assumptions (H1)–(H4) and (H6), but instead of (H5) impose its following strengthened modification:

**(H5')** The integrand $f(x,y,v,\cdot)$ is continuous for a.e. $t \in [a,b]$ and bounded uniformly with respect to $(x,y,v) \in U \times (M_C I\!B) \times (M_F I\!B)$; furthermore, there are numbers $\mu > 0$ and $L_f \geq 0$ such that $f(\cdot,\cdot,\cdot,t)$ is Lipschitz continuous on the set $A_\mu(t)$ from (H5) with constraint $L_f$ uniformly in $t \in [a,b]$.

Now we are ready to formulate and prove aforementioned necessary optimality conditions for relaxed intermediate local minimizers in $(P)$ *without* imposing any sequential normal compactness assumptions on endpoint constraints given by finitely many Lipschitzian equalities and inequalities.

**Theorem 6.1 (extended Euler-Lagrange conditions for relaxed intermediated local minimizers in delay-differential inclusions with functional endpoint constraints).** *Let $\bar{x}(\cdot)$ be a relaxed intermediate local minimizer for problem $(P)$ under hypotheses* (H1)–(H4), (H5'), *and* (H6). *Assume in addition that both spaces $X$ and $X^*$ are Asplund. Then there are multipliers $(\lambda_0,\ldots,\lambda_{m+r}) \in I\!R^{m+r+1}$ and absolutely continuous dual arcs $p\colon [a,b] \to X^*$ and $q\colon [a-\Delta,b] \to X^*$ satisfying the following relationships:*

- *the sign and nontriviality conditions*

$$\lambda_i \geq 0 \;\; \text{for all} \;\; i = 0,\ldots,m+r, \quad \text{and} \quad \sum_{i=0}^{m+r} \lambda_i = 1; \tag{6.1}$$

- *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*

$$\begin{aligned}(\dot{p}(t),\dot{q}(t-\Delta)) \in \operatorname{clco}\Big\{(u,w)\Big|\; &(u,w,p(t)+q(t)) \in \lambda_0 \partial_+ f\big(\bar{x}(t),\bar{x}(t-\Delta),\dot{\bar{x}}(t),t\big) \\ &+ N_+\big((\bar{x}(t),\bar{x}(t-\Delta),\dot{\bar{x}}(t)); \operatorname{gph} F(\cdot,\cdot,t)\big)\Big\} \quad \text{a.e. } t \in [a,b],\end{aligned} \tag{6.3}$$

*where the norm-closure operation can be omitted when the state space $X$ is reflexive;*

- *the optimal tail conditions*

$$\begin{cases} \langle \dot{q}(t),\bar{x}(t)\rangle = \min\limits_{c \in C(t)} \langle \dot{q}(t), c\rangle & \text{a.e. } t \in [a-\Delta,a), \\ q(t) = 0, \quad t \in [b-\Delta,b]; \end{cases} \tag{6.4}$$

• *the transversality inclusion*

$$-p(b) \in \sum_{i=0}^{m} \lambda_i \partial\varphi_i(\bar{x}(b)) + \sum_{i=m+1}^{m+r} \lambda_i \big[\partial\varphi_i(\bar{x}(b)) \cup \partial(-\varphi_i)(\bar{x}(b))\big]. \tag{6.5}$$

**Proof.** Given the relaxed intermediate local minimizer $\bar{x}(\cdot)$ for the original problem $(P)$, we first employ Theorem 3.2 that ensures the strong $L^1$-approximation of $\bar{x}(\cdot)$ on the initial interval $[a-\Delta, a]$ and the strong $W^{1,1}$-approximation of $\bar{x}(\cdot)$ on the main interval $[a,b]$ by a sequence of optimal solutions $\bar{x}_N(\cdot)$ to the discrete approximation problems $(P_N)$. As mentioned, we actually have the $L^2/W^{1,2}$-approximation under the assumptions made. Picking a sequence $\varepsilon_N \downarrow 0$ as $N \to \infty$ and using the necessary optimality conditions for $\bar{x}_N(\cdot)$ obtained in Theorem 5.1, find the corresponding sequences of multipliers $\lambda_i^N$, $i = 0,\ldots,m+r$, and of the discrete adjoint arcs $p_j^N \in X^*$, $j = 0,\ldots,k+1$, and $q_j^N \in X^*$, $j = -N,\ldots,k+1$, satisfying all the relationships in (5.9)–(5.13). By (5.9), we suppose without loss of generality that

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

where the limiting multipliers $\lambda_i$, $i = 0,\ldots,m+r$, satisfy the sign and nontriviality conditions in (6.1). We easily get the complementarity slackness conditions (6.2) by passing to the limit in (5.10) with taking into account that $\eta_N \to 0$ as $N \to \infty$.

Consider the piecewise linear extensions $p^N(t)$ and $q^N(t)$ of the discrete adjoint trajectories to the continuous-time intervals $[a,b]$ and $[a-\Delta, b]$, respectively, and define by

$$\theta^N(t) := \frac{\theta_j^N}{h_N} a_j^N \ \text{ for } \ t \in [t_j, t_{j+1}), \quad j = 0,\ldots,k,$$

$$\sigma^N(t) := \frac{\sigma_j^N}{h_N} b_j^N \ \text{ for } \ t \in [t_j, t_{j+1}), \quad j = -N,\ldots,-1,$$

the piecewise constraint extensions of the discrete quantities (5.8) to the corresponding intervals, where $a_j^N$ and $b_j^N$ are taken from (5.11) and (5.12), respectively.

By the constructions of $\theta^N(t)$ and $\sigma^N(t)$ we have the estimates

$$\int_a^b \|\theta^N(t)\|\,dt = \sum_{j=0}^{k} \|\theta_j^N\| \le 2\sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} \left\|\frac{\bar{x}_{j+1}^N - \bar{x}_j^N}{h_N} - \dot{\bar{x}}(t)\right\| dt = 2\int_a^b \|\dot{\bar{x}}^N(t) - \dot{\bar{x}}(t)\|\,dt,$$

$$\int_{a-\Delta}^{a} \|\sigma^N(t)\|\,dt = \sum_{j=-N}^{-1} \|\sigma_j^N\| \le 2\sum_{j=-N}^{-1} \int_{t_j}^{t_{j+1}} \|\bar{x}_j^N - \bar{x}(t)\|\,dt = 2\int_{a-\Delta}^{a} \|\bar{x}^N(t) - \bar{x}(t)\|\,dt,$$

which imply by Theorem 3.2 and classical real analysis that

$$\theta^N(t) \to 0 \ \text{ a.e. } \ t \in [a,b], \ \sigma^N(t) \to 0 \ \text{ a.e. } \ t \in [a-\Delta, a] \ \text{ as } \ N \to \infty \tag{6.6}$$

along a subsequence of $I\!N$ that is assumed to be the whole natural series. Proceeding further similarly to the proofs of [8, Theorem 6.21] and [11, Theorem 6.1], we derive from the approximate Euler-Lagrange and transversality inclusions of Theorem 5.1 with the use of the *coderivative condition* for the Lipschitzian property of set-valued mappings from [7, Theorem 1.43] and the uniform boundedness of Fréchet subgradients for Lipschitzian functions [7, Proposition 1.85] that the sequences $\{\dot p^N(t)\}$ and $\{\dot q^N(t-\Delta)\}$ are uniformly bounded in $L^1([a,b];X^*)$. Since both spaces $X$ and $X^*$ have the Radon-Nikoým property, we conclude without loss of generality, by using the Dunford weak compactness theorem and arguing similarly to the proof of Theorem 3.2, that

$$\begin{cases} p^N(t) \to p(t) \ \text{ weak}^* \text{ in } X^* \ \text{ for all } \ t\in[a,b],\\ \dot p^N(\cdot) \to \dot p(\cdot) \ \text{ weakly in } \ L^1([a,b];X^*) \end{cases} \tag{6.7}$$

as $I\!N \to \infty$ with the absolutely continuous limiting function $p\colon [a,b]\to X^*$ and

$$\begin{cases} q^N(t-\Delta) \to q(t-\Delta) \ \text{ weak}^* \text{ in } X^* \ \text{ for all } \ t\in[a,b],\\ \dot q^N(\cdot-\Delta) \to \dot q(\cdot-\Delta) \ \text{ weakly in } \ L^1([a,b];X^*) \end{cases} \tag{6.8}$$

as $I\!N \to \infty$ with the absolutely continuous limiting function $q\colon [a-\Delta,b]\to X^*$. The latter implies the optimal tail conditions in (6.4) by passing to the limit in (5.12) as $N\to\infty$ and taking into account the specific structure of the normal cone to the convex sets $C(t)$ therein as well as the pointwise convergence of $\sigma^N(\cdot)$ in (6.6).

Further, it is easy to observe that the approximate Euler-Lagrange inclusion (5.11) can be equivalently written as

$$\begin{aligned} &(\dot p^N(t),\dot q^N(t-\Delta)) \in \Big\{(u,v)\in X^*\times X^*\Big|\ \Big(u,v,p^N(t_{j+1})+q^N(t_{j+1})\frac{\lambda_0^N\theta_j^N a_j^N}{h_N}\Big)\\ &\in \lambda_0^N\widehat\partial f(\bar x(t_j),\bar x(t_j-\Delta),\dot{\bar x}^N(t),t_j)\\ &+\widehat N\big((\bar x^N(t_j),\bar x^N(t_j-\Delta),\dot{\bar x}^N(t));\operatorname{gph}F(\cdot,\cdot,t_j)\big)\Big\}+\varepsilon_N I\!B^* \end{aligned} \tag{6.9}$$

for all $t\in[t_j,t_{j+1})$, $j=0,\ldots,k$, and $N\in I\!N$. Using (6.6)–(6.8) and applying the classical Mazur convexification theorem to (6.9), we get the extended Euler-Lagrange inclusion (6.3) by passing to the limit in (6.9) as $N\to\infty$ and taking into account the extended normal cone and subdifferential constructions in (4.8) and (4.9). Note that the closure operation in (6.3) can be omitted of $X$ is reflexive. Indeed, in this case the weak and weak* topology agree and, furthermore, every bounded and convex set is weakly compact in $X^*$ being therefore automatically norm-closed closed in the latter space due to the aforementioned Mazur theorem. Hence the arguments above allow us to drop the closure operation in the limiting convexification procedure. Finally, passing to the limit in (5.13) as $N\to\infty$ and taking into account the basic subdifferential construction (4.6), we arrive at the transversality inclusion (6.5) and thus complete the proof of the theorem. $\triangle$

# References


[1] J. M. Borwein and Q. J. Zhu, *Techniques of Variational Analysis*, CMS Books in Mathematics **20**, Springer, New York, 2005.

[2] J. Diestel and J. J. Uhl, *Vector Measures*, American Mathematical Society, Providence, R.I., 1977.

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

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

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

[6] B. S. Mordukhovich, Discrete approximations and refined Euler-Lagrange conditions for non-convex differential inclusions, *SIAM J. Control and Optim.* **33** (1995), 882–815.

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

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

[9] B. S. Mordukhovich, Variational analysis of evolution inclusions, *SIAM J. Optim.* **18** (2007), 752–777.

[10] B. S. Mordukhovich and B. Wang, Generalized differentiation of parameter-dependent sets and mappings, *Optimization* **57** (2008), 17–40.

[11] B. S. Mordukhovich, D. Wang and L. Wang, Optimal control of delay-differential inclusions with multivalued initial conditions in infinite dimensions, *Control Cybern.* **37** (2008), No. 2.

[12] B. S. Mordukhovich and L. Wang, Optimal control of constrained delay-differential inclusions with multivalued initial conditions, *Control Cybern.* **28** (2003) 585–609.

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

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

[15] J. Warga, *Optimal Control of Differential and Functional Equations*, Academic Press, New York, 1972.