11-1-2004

Optimal Control of Semilinear Evolution Inclusions via Discrete Approximations

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

Dong Wang
Wayne State University

OPTIMAL CONTROL OF SEMILINEAR EVOLUTION INCLUSIONS VIA DISCRETE APPROXIMATIONS

BORIS S. MORDUKHOVICH and DONG WANG

WAYNE STATE UNIVERSITY

Detroit, MI 48202

Department of Mathematics
Research Report

2004 Series
#11

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

OPTIMAL CONTROL OF SEMILINEAR EVOLUTION INCLUSIONS VIA DISCRETE APPROXIMATIONS[1]

by

Boris S. Mordukhovich and Dong Wang

Department of Mathematics, Wayne State University, Detroit, MI 48202
boris@math.wayne.edu, dwang@math.wayne.edu

Dedicated to Czeslaw Olech

Abstract: This paper studies a Mayer type optimal control problem with general endpoint constraints for semilinear unbounded evolution inclusions in reflexive and separable Banach spaces. First, we construct a sequence of discrete approximations to the original optimal control problem for evolution inclusions and prove that optimal solutions to discrete approximation problems uniformly converge to a given optimal solution for the original continuous-time problem. Then, based on advanced tools of generalized differentiation, we derive necessary optimality conditions for discrete-time problems under fairly general assumptions. Combining these results with recent achievements of variational analysis in infinite-dimensional spaces, we establish new necessary optimality conditions for constrained continuous-time evolution inclusions by passing to the limit from discrete approximations.

Keywords: optimal control, variational analysis, generalized differentiation, semilinear evolution inclusions, discrete approximations, necessary optimality conditions.

1 Introduction

Let X be a reflexive and separable Banach space, and let $F\colon X \times [a,b] \rightrightarrows X$ be a set-valued mapping. The primary object of this paper is the following *Mayer-type problem* (P) for *semilinear evolution inclusions* with general endpoint constraints:

$$\text{minimize} \quad J[x] := \varphi(x(b)) \tag{1.1}$$

over *mild* continuous trajectories $x\colon [a,b] \to X$ for the semilinear evolution inclusion

$$\dot{x}(t) \in Ax(t) + F(x(t),t), \quad x(a) = x_0 \in X \tag{1.2}$$

subject to the endpoint constraint

$$x(b) \in \Omega \subset X, \tag{1.3}$$

where $A\colon X \to X$ is an *unbounded* generator of the C_0-semigroup $\{e^{At} | \; t \geq 0\}$ and where $\Omega \subset X$ is a closed set. A special case of $F(x,t) = f(x,U,t)$ with a control set U relates

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

(1.2) to semilinear control evolution *equations* considered in PDE control theory for *smooth* data; see, e.g., the books by Fattorini [6] and Li and Yong [10] with their references and comprehensive discussions.

Optimal control problems governed by differential inclusions with *finite-dimensional* state spaces $X = I\!R^n$ (when there is no need to single out the linear term in (1.2)) have been intensively studied in many publications, mostly from the viewpoint of deriving necessary optimality conditions; see Clarke [3], Ioffe [9], Loewen and Rockafellar [11], Mordukhovich [12], Smirnov [18], Vinter [20], and the references therein.

Differential/evolution inclusions in *infinite dimensions* are essentially more involved and require new tools for their analysis, even in the case when A is a *bounded* operator (or $A = 0$) and F is a *compact-valued* mapping; see Tolstonogov [19] and Mordukhovich [13, 14] regarding various results for such inclusions. However, the above boundedness/compactness assumptions are quite restrictive for a number of important applications, especially to dynamic systems governed by partial differential equations and inclusions.

Although semilinear models of type (1.2) with control representations $F(x,t) = f(x,U,t)$ involving smooth functions $f(\cdot,u,t)$ have been studied in the literature in connection with optimal control problems for partial differential equations (see the references above), the *inclusion* models (1.2) have not drawn much attention. Let us mention the paper by Frankowska [7], where semilinear inclusions (1.2) were studied from the viewpoint of relaxation/convexification results and reachable set properties. We are not familiar with any work on necessary optimality conditions for semilinear evolution inclusions or even semilinear evolution equations with nonsmooth dynamics. Our previous results were announced in Mordukhovich and Wang [16] for problem (P) involving autonomous inclusions (1.2).

The primary goal of this paper is to extend the method of *discrete approximations* developed by Mordukhovich [12, 13, 14] for optimal control systems governed by bounded/compact differential inclusions. The *unboundedness* of the operator A in (1.2) is a principal issue for applications to control problems for partial differential equations and inclusions. First we establish *well-posedness/stability* of discrete approximations in the sense of the *uniform convergence* of their optimal solutions to the reference optimal solution for the original problem. Based on the advanced tools of variational analysis and generalized differentiation, we derive necessary optimality conditions for discrete-time problems and then, by passing to the limit from discrete approximations, obtain necessary conditions of the Euler-Lagrange type for the original problem (P).

The rest of the paper is organized as follows. Section 2 is devoted to the construction and justification of a *well-posed discrete approximation* for the original continuous-time problem. We first establish, under fairly general assumptions, that *any mild* continuous trajectory for (1.2) can be strongly approximated, in the $C([a,b];X)$-norm, by feasible trajectories of the corresponding discrete inclusions that are piecewise linearly extended to the continuous-time interval $[a,b]$. This allows us to justify the uniform convergence of *optimal solutions* for discrete approximation problems to *given* optimal solution for problem (P), thus making a bridge between discrete-time and continuous-time dynamic optimization problems.

In Section 3 we briefly review basic tools of generalized differentiation in variational analysis needed for deriving necessary optimality conditions in discrete approximation problems and then establishing, by passing to the limit, adequate necessary conditions for optimality of the given solution to (P).

Section 4 is devoted to necessary optimality conditions for the discrete-time problems appeared in our discrete approximation procedure. We pay the main attention to "fuzzy" optimality conditions that are more convenient and less restrictive for passing to the limit.

In Section 5 we develop the limiting procedure to establish necessary optimality conditions for the original continuous-time problem (P) by passing to the limit from discrete approximations. In this way we obtain new conditions in the *extended Euler-Lagrange form* involving mild solutions to a certain *adjoint* evolution inclusion.

2 Discrete approximations

The main goal of this section is to construct a well-posed sequence of discrete approximation problems to the continuous-time optimal control problem (P) under consideration. To achieve this goal, we obtain also some other results on discrete approximations that are certainly of independent interest.

We begin with clarifying the definition of mild solutions to the evolution inclusion (1.2), where A is an unbounded generator of the C_0-semigroup $\{e^{At}|\ t \geq 0\}$. A continuous trajectory/arc $x\colon [a,b] \to X$ is a *mild solution* to (1.2) if there is a Bochner integrable mapping $v \in L^1([a,b];X)$ such that

$$\begin{cases} x(t) = e^{A(t-a)}x_0 + \displaystyle\int_a^t e^{A(t-s)}v(s)\,ds \ \text{ for all } \ t \in [a,b] \\ \text{with } \ v(t) \in F(x(t),t) \ \text{ a.e. } \ t \in [a,b]. \end{cases} \tag{2.1}$$

In contrast to strong solutions for differential inclusions, we do *not* require the a.e. Fréchet differentiability of feasible arcs (which is not realistic for *unbounded* operators A) and actually replace (1.2) by the *integral inclusion* (2.1) considered in the space $C([a,b];X)$.

In what follows we always assume that the Banach space X is *reflexive and separable* and that A generates a *compact* C_0-semigroup e^{At} on X. We also suppose that A generates a semigroup of *contractions* on X, which does not restrict the generality. Indeed, given a Banach space X with the original norm $\|\cdot\|$ and an arbitrary C_0-semigroup $\{e^{At}|\ t \geq 0\}$ on X with $\|e^{At}\| \leq M$, let us renorm X by

$$\|x\|_1 = \sup_{t \geq 0} \|e^{At}x\|$$

and observe that $\|x\| \leq \|x\|_1 \leq M\|x\|$ for each $x \in X$. In addition one has

$$\|e^{At}x\|_1 = \sup_{\tau \geq 0} \|e^{A\tau}e^{At}x\| \leq \sup_{t \geq 0} \|e^{At}x\| = \|x\|_1,$$

which shows that $\{e^{At}|\ t \geq 0\}$ is a contraction semigroup on $(X, \|\cdot\|_1)$. It is easy to conclude that $\{e^{At}|\ t \geq 0\}$ is a C_0-semigroup on the renormed space $(X, \|\cdot\|_1)$; see, e.g., the book by Ahmed [1] for more details.

Fix an *arbitrary mild trajectory* $\bar{x}(\cdot)$ for the original inclusion (1.2) and impose the following *standing assumptions* on the set-valued mapping F:

(H1) There are an open set $U \subset X$ and positive numbers ℓ_F, m_F such that $\bar{x}(t) \in U$ as $t \in [a,b]$ and the sets $F(x,t)$ are compact and convex for all $x \in U$ and almost all $t \in [a,b]$. Moreover, one has

$$F(x,t) \subset m_F I\!B, \quad (x,t) \in U \times [a,b], \quad \text{and} \tag{2.2}$$

$$F(x_1,t) \subset F(x_2,t) + \ell_F \|x_1 - x_2\| I\!B, \quad x_1, x_2 \in U, \quad t \in [a,b], \tag{2.3}$$

where $I\!B$ stands for the closed unit ball of the space in question.

(H2) $F(x,\cdot)$ is Hausdorff continuous for a.e. $t \in [a,b]$ uniformly in $x \in U$.

Note that (2.3) signifies the local *Lipschitz continuity* of $F(\cdot,t)$ around $\bar{x}(t)$. To clarify the meaning of (H2), consider the so-called *averaged modulus of continuity* $\tau(F,h)$ for $F(x,t)$ in $t \in [a,b]$ when $x \in U$ defined by

$$\tau(F;h) := \int_a^b \sigma(F;t,h)\,dt, \tag{2.4}$$

where $\sigma(F;t,h) := \sup\{\omega(F;x,t,h) |\ x \in U\}$, where

$$\omega(F;x,t,h) := \sup\big\{\text{haus}(F(x,t_1),F(x,t_2)) \big|\ t_1,t_2 \in [t-h/2, t+h/2] \cap [a,b]\big\},$$

and where $\text{haus}(\cdot,\cdot)$ stands for the Hausdorff distance between compact sets. It is proved by Dontchev and Farkhi [5] that if $F(x,\cdot)$ is Hausdorff continuous for a.e. $t \in [a,b]$ uniformly in $x \in U$, then $\tau(F;h) \to 0$ as $h \to 0$.

Observe that the *convex-valuedness* assumption on $F(x,t)$ in (H1) is imposed for simplicity; it can be replaced by the so-called *relaxation stability* and can be actually dropped at all in some settings; cf. Mordukhovich [12, 13, 14] with $A = 0$ as well as Theorem 2.1 stated below in the general case. Note also that, everywhere except Theorem 2.1, we need the validity of assumptions (H1) and (H2) around a given *optimal solution* $\bar{x}(\cdot)$ to (P). In fact, the global optimality of $\bar{x}(\cdot)$ can be replaced by its *local strong optimality*, i.e., relative to a $C([a,b];X)$-neighborhood.

Our first step is to build *well-posed discrete approximations* of the *integral system* (2.1), i.e., for *mild* solutions of the initial inclusion (1.2) without taking into account the minimizing functional (1.1) and the endpoint constraint (1.3) in the original Mayer problem (P). For any natural number $N \in I\!N := \{1,2,\ldots\}$ consider the grid/partition

$$T_N := \big\{t_j = t_0 + jh_N \big|\ j = 0,\ldots,N\big\} \ \text{ with } \ t_0 = a, \ t_N = b, \ \text{ and stepsize } \ h_N := \frac{b-a}{N}.$$

The sequence of discrete inclusions approximating (2.1) is constructed as follows:

$$\left\{\begin{array}{l} x_N(t_{j+1}) \in e^{Ah_N} x_N(t_j) + h_N e^{Ah_N} F(x_N(t_j),t_j) \\ \text{with } \ j = 0,\ldots,N-1 \ \text{ and } \ x_N(t_0) = x_0. \end{array}\right. \tag{2.5}$$

Denote by $x_N(t)$, $a \le t \le b$, *piecewise linear extensions* of discrete trajectories $x_N(t_j)$ for (2.5) to the continuous-time interval $[a,b]$. The following result ensures, under the standing assumptions made, the uniform approximation of an arbitrary mild trajectory for (1.2) by a sequence of extended trajectories for the discrete inclusions (2.5).

Theorem 2.1 (uniform approximation of mild trajectories). *Let $\bar{x}(\cdot)$ be an arbitrary mild trajectory for* (1.2)*, i.e., satisfy the integral inclusion* (2.1) *under all the assumptions in* (H1) *and* (H2) *except the convex-valuedness of $F(x,t)$. Then there is a sequence of extended discrete trajectories $x_N(\cdot)$ for* (2.5) *that converges to $x(\cdot)$ in the norm of $C([a,b];X)$.*

Proof. Without loss of generality, assume in what follows that the operator A generates a C_0-semigroup $\{e^{At}|\ t \ge 0\}$ of contractions on X. Let $\{w_N(\cdot)\}$, $N = 1,2,\ldots$ be an arbitrary sequence of functions in $[a,b]$ such that $w_N(\cdot)$ are constant on $[t_j, t_{j+1})$ for every $j = 0,\ldots,N-1$ and $w_N(t)$ converge to some $v(t) \in F(\bar{x}(t),t)$ as $N \to \infty$ in the norm of $L^2([a,b];X)$. Such a sequence always exists because of the density of step-functions in $L^2([a,b];X)$. It is easy to see that

$$\int_a^t w_N(s)\,ds \to \int_a^t v(s)\,ds \quad \text{as} \quad N \to \infty \quad \text{uniformly on} \quad [a,b].$$

Then by the boundedness assumption in (2.2) and by the triangle inequality one gets

$$\left\| \int_a^t w_N(s)\,ds \right\| \le m_F(t-a) \quad \text{whenever} \quad t \in [a,b] \quad \text{and} \quad N \to \infty. \tag{2.6}$$

In the arguments and estimates below we use the numerical sequence

$$\xi_N := \left\| \int_a^b (v(t) - w_N(t))\,dt \right\| \to 0 \text{ as } N \to \infty. \tag{2.7}$$

Define the sequence of discrete functions $\{y_N(t_j)|\ j = 0,\ldots,N\}$ by

$$\begin{cases} y_N(t_{j+1}) = e^{Ah_N} y_N(t_j) + h_N e^{Ah_N} w_N(t_j) \\ \text{with } \ j = 0,\ldots,N-1 \ \text{ and } \ y_N(t_0) = x_0. \end{cases} \tag{2.8}$$

Note that the functions

$$y_N(t) := e^{A(t-a)} x_0 + \int_a^t e^{A(t-s)} w_N(s)\,ds, \quad t \in [a,b], \tag{2.9}$$

are piecewise linear extensions of (2.8) on the interval $[a,b]$ satisfying

$$\|y_N(t) - \bar{x}(t)\| \le \xi_N \quad \text{whenever} \quad t \in [a,b]. \tag{2.10}$$

The latter implies that $y_N(t) \in U$ for all $t \in [a,b]$ if N is sufficiently large. To proceed with the estimates below, observe that the Lipschitz condition (2.3) is clearly equivalent to

$$\operatorname{dist}(w, F(x_1,t)) \le \operatorname{dist}(w, F(x_2,t)) + \ell_F \|x_1 - x_2\| \quad \text{whenever} \quad w \in X,\ x_1, x_2 \in U, \quad \text{and} \quad t \in [a,b].$$

Furthermore, for any $w, x \in X$ and $t_1, t_2 \in [a, b]$ one has

$$\mathrm{dist}(w, F(x, t_1)) \le \mathrm{dist}(w, F(x, t_2)) + \mathrm{haus}(F(x, t_1), F(x, t_2)).$$

Now using the average modulus of continuity (2.4), we get

$$\begin{aligned}
\zeta_N : &= \sum_{j=0}^{N-1} h_N \mathrm{dist}(w_N(t_j), F(y_N(t_j), t_j)) = \sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} \mathrm{dist}(w_N(t_j), F(y_N(t_j), t_j))\, dt \\
&\le \sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} \mathrm{dist}(w_N(t_j), F(y_N(t_j), t))\, dt + \tau(F; h_N).
\end{aligned}$$

Since $w_N(t)$ are constants on $[t_j, t_{j+1})$, it follows from (2.3), (2.6), (2.10), and the contraction property of the C_0-semigroup $\{e^{At} |\, t \ge 0\}$ that

$$\begin{aligned}
\mathrm{dist}(w_N(t_j), F(y_N(t_j), t)) \quad &\le \mathrm{dist}(w_N(t), F(y_N(t), t)) + \ell_F \|y_N(t_j) - e^{A(t-t_j)} y_N(t_j)\| \\
&\quad + \ell_F m_F (t - t_j)
\end{aligned}$$

for all $t \in [t_j, t_{j+1})$, and that

$$\begin{aligned}
\mathrm{dist}(w_N(t), F(y_N(t), t)) \quad &\le \mathrm{dist}(w_N(t), F(x(t), t)) + \ell_F \|y_N(t) - x(t)\| \\
&\le \|w_N(t) - v(t)\| + \ell_F \xi_N
\end{aligned}$$

for a.e. $t \in [a, b]$. Thus one has

$$\left\{ \begin{aligned}
&\zeta_N \le \gamma_N := \big[1 + l_F(b-a)\big]\xi_N + l_F \theta_N + \tfrac{1}{2}(b-a)\ell_F m_F h_N + \tau(F; h_N) \\
&\text{with } \theta_N := \sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} \|y_N(t_j) - e^{A(t-t_j)} y_N(t_j)\|\, dt \to 0 \ \text{ as } \ N \to \infty.
\end{aligned} \right. \tag{2.11}$$

Observe that the discrete functions (2.8) are *not* trajectories for (2.5) because the inclusions $w_N(t_j) \in F(y_N(t_j), t_j)$ are not generally guaranteed for all $j = 0, \dots, N-1$. Now we use $w_N(t_j)$ to define *trajectories* for (2.5), which are close to $y_N(t_j)$ and have the convergence property stated in this theorem.

Let us construct the desirable trajectories $\{\bar{x}_N(t_j) |\, j = 0, \dots, N\}$ by using the following *proximal algorithm:*

$$\left\{ \begin{aligned}
&x_N(t_0) = x_0, \ q_N(t_j) \in F(x_N(t_j), t_j), \\
&\|q_N(t_j) - w_N(t_j)\| = \mathrm{dist}(w_N(t_j), F(x_N(t_j), t_j)), \\
&x_N(t_{j+1}) = e^{Ah_N} x_N(t_j) + h_N e^{Ah_N} q_N(t_j), \quad j = 0, \dots, N-1.
\end{aligned} \right. \tag{2.12}$$

Then the piecewise linear extensions of $x_N(t_j)$, $j = 0, \dots, N$, are given by

$$x_N(t) = e^{A(t-a)} x_0 + \int_a^t e^{A(t-s)} q_N(s)\, ds, \quad t \in [a, b].$$

Now following the scheme of proving Theorem 2.1 in Mordukhovich [12] and adapting it to the case under consideration involving semilinear evolution inclusions in infinite-dimensional spaces, we show that the extensions $x_N(t),\ t \in [a,b]$, of the above discrete trajectories converge to $x(t)$ in the norm of $C([a,b];X)$. This completes the proof of the theorem. $\triangle$

Next suppose that $\bar{x}(\cdot)$ is an *optimal solution* to (P) and construct a sequence of optimization problems (P_N) for discrete inclusions (2.5) in such a way that optimal solutions to (P_N) *strongly* (in the norm of $C([a,b];X)$) converge to $\bar{x}(\cdot)$ as $N \to \infty$. Imposing assumptions (H1) and (H2) along $\bar{x}(\cdot)$ and using Theorem 2.1, we approximate $\bar{x}(\cdot)$ by discrete trajectories $\{\bar{x}_N(t_j)|\ j = 0,\dots,N\}$ and compute the numerical sequence

$$\eta_N := \gamma_N \exp\big[\ell_F(b-a)\big] + \xi_N \downarrow 0 \ \text{ as } \ N \to \infty, \tag{2.13}$$

where ξ_N and γ_N are defined in (2.7) and (2.11), respectively. Define a sequence of discrete approximation problems (P_N), $N \in I\!N$, as follows: minimize

$$J_N[x_N] := \varphi(x_N(b)) + h_N \sum_{j=0}^{N-1} \|x_N(t_{j+1}) - \bar{x}(t)\|^2 \tag{2.14}$$

subject to the discrete-time inclusions (2.5) and the *perturbed* endpoint constraints

$$x_N(b) \in \Omega_N := \Omega + \eta_N I\!B. \tag{2.15}$$

Note that nonzero perturbations η_N in (2.15) of the original endpoint constraint (1.3) are *crucial* for the validity of the next result, which makes a bridge between the continuous-time and discrete-time optimization problems under consideration.

Theorem 2.2 (uniform convergence of discrete optimal solutions). *Let $\bar{x}(\cdot)$ be an optimal solution to problem (P), and let the sequence $\{\eta_N\}$ be constructed in* (2.13). *In addition to the standing assumptions on A and F, suppose that the cost function φ is lower semicontinuous on U and continuous at $\bar{x}(b)$ and that the constraint set Ω is locally closed around this point. Then for each $N \in I\!N$ the discrete-time optimization problem (P_N) admits an optimal solution. Furthermore, any sequence $\{\bar{x}_N(t)\}$, $t \in [a,b]$, of extended optimal solutions for (P_N) converges to $\bar{x}(\cdot)$ strongly in $C([a,b];X)$ as $N \to \infty$.*

Proof. It follows from the proof of Theorem 2.1 and the choice of η_N in (2.13) and (2.15) that the discrete trajectories $\{x_N(t_j)|\ j = 0,\dots,N\}$ constructed in Theorem 2.1 for the given optimal solution $\bar{x}(\cdot)$ to (P) are *feasible* solutions to (P_N) for all $N \in I\!N$ sufficiently large. Then the classical Weierstrass theorem ensures the existence of *optimal* solutions $\bar{x}_N(\cdot) = (\bar{x}_N(t_0), \bar{x}_N(t_1), \dots, \bar{x}_N(t_N))$ to (P_N) with $\bar{x}_N(t_0) = x_0$ for such N under the assumptions made. Let us prove that for *any* sequence of optimal solutions $\bar{x}_N(\cdot)$ to (P_N) we have the inequality

$$\limsup_{N\to\infty}\ J_N[\bar{x}_N] \ \le \ J[\bar{x}] \tag{2.16}$$

To accomplish this, it suffices to show that

$$J_N[x_N] = \varphi(x_N(b)) + h_N \sum_{j=0}^{N-1} \|x_N(t_{j+1}) - \bar{x}(t)\|^2 \to J[\bar{x}] = \varphi(\bar{x}(b)) \quad \text{as} \ \ N \to \infty \tag{2.17}$$

for the sequence of discrete trajectories $x_N(\cdot)$ approximating $\bar{x}(\cdot)$ due to Theorem 2.1.

Since $x_N(b) \to \bar{x}(b)$, the convergence $\varphi(x_N(b)) \to \varphi(\bar{x}(b))$ as $N \to \infty$ follows directly from the continuity of φ at $\bar{x}(b)$. To justify (2.17), it remains showing that

$$h_N \sum_{j=0}^{N-1} \|x_N(t_{j+1}) - \bar{x}(t)\|^2 \to 0 \quad \text{as} \quad N \to \infty.$$

The latter follows from the estimate

$$\|x_N(t_{j+1}) - \bar{x}(t_{j+1})\| \le \eta_N \downarrow 0 \quad \text{as} \quad N \to \infty,$$

which can be distilled from the proof of Theorem 2.1; cf. Mordukhovich [12].

To proceed further, observe that the piecewise linear extensions of the optimal solutions $\bar{x}_N(\cdot)$ to (P_N) admit the integral representation

$$\bar{x}_N(t) = e^{A(t-a)}x_0 + \int_a^t e^{A(t-s)}v_N(s)\,ds, \quad a \le t \le b,$$

with some $v_N(t) \in F(\bar{x}_N(t), t)$ a.e. $t \in [a, b]$ for all $N \in \mathbb{N}$. Let us prove that $\bar{x}_N(t) \to \bar{x}(t)$ *uniformly* on $[a, b]$. Assuming the contrary, we have without loss of generality that

$$c := \lim_{N\to\infty} \max_{t\in[a,b]} \|\bar{x}_N(t) - \bar{x}(t)\| > 0.$$

Now following the scheme in the proof of Theorem 2.7 from Frankowska [7], we find a mild solution $\widetilde{x}(\cdot)$ to (1.2) such that $\bar{x}_N(t) \to \widetilde{x}(t)$ uniformly on $[a, b]$. Since

$$h_N \sum_{j=0}^{N-1} \|\bar{x}_N(t_{j+1}) - \bar{x}(t)\|^2 \le \int_a^b c^2\,dt = c^2(b-a)$$

for all N sufficiently large, we get from (2.14) and (2.16) that

$$J[\widetilde{x}] < \varphi(\widetilde{x}(b)) + c^2(b-a) \le \liminf_{N\to\infty} J_N[\bar{x}_N] \le \limsup_{N\to\infty} J_N[\bar{x}_N] \le J[\bar{x}].$$

The latter contradicts the optimality of $\bar{x}(\cdot)$ in the original problem (P). Hence $c = 0$, which completes the proof of the theorem. $\triangle$

3 Tools of generalized differentiation

This section contains some preliminary material on generalized differentiation widely used in the variational analysis of evolution inclusions conducted in what follows. We refer the reader to the book by Mordukhovich [14] for more details, discussions, and the extensive bibliography; a finite-dimensional counterpart of the generalized differential theory is available in the book by Rockafellar and Wets [17]. Since the standing framework of the paper confines ourselves to reflexive and separable Banach spaces, we present formulations of the main constructions and results holding in this setting. Note however that all the results presented in this section hold true in the (essentially more general) framework of Asplund spaces, while some of them are valid in other (even arbitrary) Banach space settings under appropriate modifications of definitions.

Given $\Omega \subset X$, define the (limiting, basic) *normal cone* to Ω at $\bar{x} \in \Omega$ by

$$N(\bar{x};\Omega) := \operatorname*{Lim\,sup}_{x \xrightarrow{\Omega} \bar{x}} \widehat{N}(x;\Omega), \tag{3.1}$$

where "Lim sup" signifies the *sequential* Painlevé-Kuratowski upper/outer limit of a set-valued mapping from X to X^* in the norm topology of X and the weak*(=weak) topology of X^*, where $x \xrightarrow{\Omega} \bar{x}$ means that $x \to \bar{x}$ with $x \in \Omega$, and where $\widehat{N}(x;\Omega)$ stands for the *prenormal* (or *Fréchet normal*) cone to Ω at $x \in \Omega$ given by

$$\widehat{N}(x;\Omega) := \Big\{ x^* \in X^* \Big|\ \limsup_{u \xrightarrow{\Omega} x} \frac{\langle x^*, u - x\rangle}{\|u - x\|} \le 0 \Big\} \tag{3.2}$$

via the classical upper limit "lim sup" for scalar functions.

Given a set-valued mapping $F\colon X \rightrightarrows Y$ and a point $(\bar{x},\bar{y}) \in \operatorname{gph} F$, define the *coderivative* of F at $(\bar{x},\bar{y})$ as a positive homogeneous mapping $D^*F(\bar{x},\bar{y})\colon Y^* \rightrightarrows X^*$ with

$$D^*F(\bar{x},\bar{y})(y^*) := \big\{ x^* \in X^* \big|\ (x^*, -y^*) \in N((\bar{x},\bar{y}); \operatorname{gph} F) \big\}. \tag{3.3}$$

If F is single-valued and C^1 around $\bar{x}$ (or merely strictly differentiable at this point), then

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

i.e., the coderivative (3.3) is an appropriate extension of the classical *adjoint* derivative operator to nonsmooth and set-valued mappings. Note that (3.3) can be equivalently represented in the limiting form

$$D^*F(\bar{x},\bar{y})(y^*) = \Big\{ x^* \in X^* \Big|\ \begin{array}{l} \exists \text{ sequences } (x_k, y_k) \xrightarrow{\operatorname{gph} F} (\bar{x},\bar{y}),\ \ (x_k^*, y_k^*) \xrightarrow{w} (x^*, y^*) \\ \text{with } \ x_k^* \in \widehat{D}^*F(x_k,y_k)(y_k^*),\ \ k \in I\!N \Big\}, \end{array}$$

where w signifies the weak convergence on X^*, $I\!N := \{1, 2, \ldots\}$, and where

$$\widehat{D}^*F(\bar{x},\bar{y})(y^*) := \big\{ x^* \in X^* \big|\ (x^*, -y^*) \in \widehat{N}((\bar{x},\bar{y}); \operatorname{gph} F) \big\} \tag{3.4}$$

stands for the *Fréchet coderivative* of F at $(\bar{x},\bar{y}) \in \operatorname{gph} F$. Using (3.4), we have the following characterization of the classical *local Lipschitzian* property of compact-valued multifunctions: F is locally Lipschitzian around $\bar{x} \in \operatorname{dom} F$ with modulus ℓ_F *if and only if* there is $\eta > 0$ such that

$$\sup\big\{ \|x^*\| \ \big|\ x^* \in \widehat{D}^*F(x,y)(y^*) \big\} \le \ell_F \|y^*\| \tag{3.5}$$

whenever $x \in \bar{x} + \eta I\!B$, $y \in F(x)$, and $y^* \in Y^*$.

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

$$\partial\varphi(\bar{x}) := \operatorname*{Lim\,sup}_{x \xrightarrow{\varphi} \bar{x}} \widehat{\partial}\varphi(x), \tag{3.6}$$

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 φ at x defined by

$$\widehat{\partial}\varphi(x) := \Big\{x^* \in X^* \Big| \liminf_{x \to \bar{x}} \frac{\varphi(u) - \varphi(x) - \langle x^*, u - x\rangle}{\|u - x\|} \ge 0\Big\}. \tag{3.7}$$

The subgradient set (3.7) is widely used in the theory of *viscosity solutions* to nonlinear partial differential equations under the name of "viscosity subdifferential." Observe that

$$N(\bar{x}; \Omega) = \partial\delta(\bar{x}; \Omega) \text{ and } \widehat{N}(\bar{x}; \Omega) = \widehat{\partial}\delta(\bar{x}; \Omega),$$

where $\delta(\cdot; \Omega)$ stands for the *indicator function* of Ω defined by $\delta(x; \Omega) = 0$ for $x \in \Omega$ and $\delta(x; \Omega) = \infty$ otherwise.

The above normal cones, subdifferentials, and coderivatives enjoy comprehensive *calculus rules*: in *fuzzy/approximate* forms for Fréchet-like constructions (3.2), (3.4), (3.7), and in *exact/pointwise* forms for their limiting counterparts. The driving force for these calculi is the usage of certain *variational principles*, or *extremal principles* in the geometric framework, which are at the very heart of variational analysis. We formulate the fuzzy rule for Fréchet subgradients of *semi-Lipschitzian sums* used in what follows, where $I\!B^*$ stands for the closed unit ball of X^*: given any $\varepsilon > 0$, one has the inclusion

$$\begin{aligned}\widehat{\partial}(\varphi_1 + \varphi_2)(\bar{x}) \subset \bigcup\Big\{ & \widehat{\partial}\varphi_1(x_1) + \widehat{\partial}\varphi_2(x_2)\Big| x_i \in \bar{x} + \varepsilon I\!B, \\ & |\varphi_i(x_i) - \varphi_i(\bar{x})| \le \varepsilon, \ i = 1, 2\Big\} + \varepsilon I\!B^*\end{aligned} \tag{3.8}$$

provided that φ_1 is Lipschitz continuous around $\bar{x}$ while φ_2 is finite at $\bar{x}$ and lower semicontinuous around this point.

Besides calculus rules for generalized differentiation that are equally important in finite and infinite dimensions, major ingredients of infinite-dimensional variational analysis are "normal compactness" properties of sets, set-valued mappings, and extended-real-valued functions that are automatic in finite dimensions while playing a crucial role in many aspects in infinite-dimensional analysis, especially those related to passing to the limit. In this paper we employ only one of such properties needed for closed sets. This property called *sequential normal compactness* (SNC) is defined as follows: $\Omega \subset X$ is SNC at $\bar{x} \in \Omega$ if for any sequences of $(x_k, x_k^*) \in X \times X^*$ satisfying

$$x_k \to \bar{x} \text{ with } x_k \in \Omega \text{ and } x_k^* \in \widehat{N}(x_k; \Omega) \text{ as } k \in I\!N$$

one has the implication $x_k^* \xrightarrow{w} 0 \Longrightarrow \|x_k^*\| \to 0$ as $k \to \infty$. When Ω is *convex* with $\operatorname{ri}\Omega \ne \emptyset$, its SNC property is *equivalent* to Ω being of *finite codimension*; the latter is widely used in optimal control of partial differential equations; cf. Fattorini [6] and Li and Yong [10].

Finally, let us mention an extension of the limiting normal cone (3.1) to the case of *moving* (parameter-dependent) sets useful in the study of nonautonomous objects. Given a moving set $\Omega\colon T \rightrightarrows X$ on a topological space of parameters, define the *extended normal cone* to $\Omega(\bar{t})$ at $\bar{x} \in \Omega(\bar{t})$ by

$$\widetilde{N}(\bar{x}; \Omega(\bar{t})) := \mathop{\mathrm{Lim\,sup}}_{(x,t) \xrightarrow{\mathrm{gph}\, F} (\bar{t}, \bar{x})} \widehat{N}(x; \Omega(t)).$$

Furthermore, $\Omega(\cdot)$ is said to be *normally semicontinuous* at $(\bar{x},\bar{t})$ if $\widetilde{N}(\bar{x};\Omega(\bar{t})) = N(\bar{x};\Omega(\bar{t}))$. The latter property holds not only for parameter-independent sets $\Omega(t) \equiv \Omega$ but in much more general settings; see, e.g., Bounkhel and Thibault [2] and Mordukhovich [12, 14].

4 Optimality conditions for discrete approximations

The primary objective of this section is to obtain necessary conditions for optimal solutions to the discrete approximation problems (P_N) governed by difference evolution inclusions in infinite-dimensional spaces. We reduce these dynamic optimization problems to "non-dynamic" problems (MP_N) of mathematical programming with *operator and many geometric constraints.* To conduct a variational analysis of problems (MP_N) and then of (P_N), we employ the tools and calculus rules of generalized differentiation discussed in Section 3. The main attention is paid to "fuzzy" results derived under minimal assumptions. They happen to be more convenient for furnishing limiting procedures to establish necessary optimality conditions in the original problem (P) developed in Section 5.

Fix $N \in I\!N$ and consider a "long" vector $z \in X^{2N+1}$ defined by

$$z = (x_0^N, x_1^N, \ldots, x_N^N, y_0^N, \ldots, y_{N-1}^N) := (x^N(t_0), x^N(t_1), \ldots, x^N(t_N), y^N(t_0) \ldots, y^N(t_{N-1})),$$

with $x_0^N = x^N(t_0) = x_0$, where the discrete "mild derivative" vectors y_j^N are given by

$$y_j^N = \frac{x_{j+1}^N - e^{Ah_N} x_j^N}{h_N} \quad \text{as} \quad j = 0, \ldots, N-1.$$

For each $N \in I\!N$ consider the following problem of mathematical programming (MP_N): minimize the cost function

$$\phi_0(z) := \varphi(x_N^N) + h_N \sum_{j=0}^{N-1} \|e^{Ah_N} x_j^N + h_N y_j^N - \bar{x}(t)\|^2 \tag{4.1}$$

subject to the constraints

$$g_j(z) := x_{j+1}^N - e^{Ah_N} x_j^N - h_N y_j^N = 0 \quad \text{for} \quad j = 0, \ldots, N-1, \tag{4.2}$$

$$\Lambda_j := \big\{(x_0^N, x_1^N, \ldots, y_{N-1}^N) \big|\, y_j^N \in e^{Ah_N} F(x_j^N, t_j)\big\} \quad \text{for} \quad j = 0, \ldots, N-1, \tag{4.3}$$

$$\Lambda_N := \{(x_0^N, x_1^N, \ldots, y_{N-1}^N) \big|\, x_N^N \in \Omega_N\}. \tag{4.4}$$

Note that constraints (4.2) are of the *operator type*, while constraints (4.3) and (4.4) are *geometric* the number of which is *increasing* as $N \to \infty$. It is easy to see that each problem (MP_N) defined in (4.1)–(4.4) is equivalent to the discrete approximation problem (P_N) given in (2.5), (2.14), and (2.15) as $N \in I\!N$. Denote

$$g(z) := (g_0(z), g_1(z), \ldots, g_{N-1}(z)) \quad \text{and} \quad \Lambda := \bigcap_{j=0}^{N} \Lambda_j.$$

The next *fuzzy intersection rule* for the sets Λ_j is implied by the general result in Mordukhovich [14, Lemma 3.1] due to the automatic fulfillment of the "fuzzy qualification condition" therein that follows from the specific structure of (4.3) and (4.4).

Lemma 4.1 (fuzzy intersection rule). *Let $\bar{z}^N = (x_0, \bar{x}_1^N, \ldots, \bar{x}_N^N, \bar{y}_0^N, \ldots, \bar{y}_{N-1}^N)$ be an optimal solution to problem (MP_N). Assume that the sets Ω and $\operatorname{gph} F(\cdot, t_j)$ are locally closed. Then for any $\varepsilon > 0$ and $\widetilde{z} \in \bar{z}^N + \varepsilon I\!B$ we have*

$$\widehat{N}(\widetilde{z}; \Lambda) \subset \widehat{N}(z_0; \Lambda_0) + \widehat{N}(z_1; \Lambda_1) + \cdots + \widehat{N}(z_N; \Lambda_N) + \varepsilon I\!B^*$$

with some $z_j \in \Lambda_j \cap (\widetilde{z} + \varepsilon I\!B)$ as $j = 0, \ldots, N$.

The following theorem gives necessary optimality conditions of a *fuzzy Lagrange multiplier* type for the infinite-dimensional mathematical programming problems (MP_N) with operator and many geometric constraints.

Theorem 4.2 (fuzzy Lagrange multiplier rule). *Let $\bar{z}^N = (x_0, \bar{x}_1^N, \ldots, \bar{x}_N^N, \bar{y}_0^N, \cdots, \bar{y}_{N-1}^N)$ be an optimal solution to problem (MP_N) as $N \in I\!N$. Assume that the cost function ϕ_0 is locally Lipschitzian and that the sets Ω and $\operatorname{gph} F(\cdot, t_j)$ are locally closed around $\bar{z}^N$. Then for any $\varepsilon > 0$ there are a number $\mu_0 \geq 0$ and adjoint vectors $\psi_j^* \in X^*$ as $j = 0, \ldots, N-1$ and $z_j^* \in (X^*)^{2N+1}$ as $j = 0, \ldots, N$ satisfying the relationships*

$$z_j^* \in \widehat{N}(z_j; \Lambda_j) \quad \text{with some} \quad z_j \in \Lambda_j \cap (\bar{z}^N + \varepsilon I\!B) \quad \text{as} \quad j = 0, \ldots, N, \tag{4.5}$$

$$-\sum_{j=0}^{N} z_j^* \in \mu_0 \widehat{\partial} \phi_0(\widetilde{z}_0) + \sum_{j=0}^{N-1} \nabla g_j(\widetilde{z})^* \psi_j^* + \varepsilon I\!B^* \quad \text{with some} \quad \widetilde{z}_0, \ \widetilde{z} \in \bar{z}^N + \varepsilon I\!B \tag{4.6}$$

and the nontriviality condition

$$\mu_0 + \sum_{j=0}^{N-1} \|\psi_j^*\| \geq 1. \tag{4.7}$$

Proof. By the above construction and notation made, each $\bar{z}^N$ is an optimal solution to the optimization problem

$$\text{minimize} \ \ \phi_0(z) \ \ \text{subject to} \ g(z) = 0 \ \ \text{and} \ \ z \in \Lambda,$$

where the index "N" is omitted for simplicity.

Assume first that $\bar{z}^N$ is a *regular point* for ϕ_0 relative to Λ, i.e., there are $\alpha > 0$ and a neighborhood U of $\bar{z}^N$ such that

$$\operatorname{dist}(x; Q) \leq \alpha \|\phi_0(z) - \phi_0(\bar{z}^N)\| \ \ \text{for all} \ \ z \in \Lambda \cap U,$$

where $Q := \{z \in \Lambda |\ \phi_0(z) = \phi_0(\bar{z}^N)\}$ and where $\operatorname{dist}(\cdot; Q)$ stands for the distance function. Then by the *reduction theorem* from Ioffe [8], $\bar{z}^N$ is a local solution to the following problem:

$$\text{minimize} \ \ \phi_0(z) + \mu \|g(z)\| \ \ \text{subject to} \ \ z \in \Lambda$$

for all $\mu > 0$ sufficiently large. This easily implies that

$$0 \in \widehat{\partial}\big[\phi_0(\cdot) + \mu \|g(\cdot)\| + \delta(\cdot; \Lambda)\big](\bar{z}^N). \tag{4.8}$$

Picking any $\varepsilon > 0$ and applying the fuzzy sum rule (3.8) to (4.8) and then the intersection rule of Lemma 4.1 to the set Λ, we find (z_j, z_j^*) satisfying (4.5) as well as $\widetilde{z}_0$, $\widetilde{z} \in \bar{z}^N + \varepsilon I\!B$, $\widetilde{z}_0^* \in \widehat{\partial}\phi_0(\widetilde{z}_0)$, and $\psi_j^* \in X^*$ satisfying

$$0 \in \widetilde{z}_0^* + \sum_{j=0}^{k-1} \nabla g_j(\widetilde{z})^* \psi_j^* + \sum_{j=0}^{k} z_j^* + \varepsilon I\!B^*.$$

Thus we arrive at (4.5)–(4.7) with $\mu_0 = 1$ in the regular case.

Consider next the remaining case, which implies that the mapping $g_\Lambda := g + \Delta(\cdot;\Lambda)$ is *not metrically regular* around $\bar{z}^N$ in the conventional sense; here $\Delta(z;\Lambda)$ stands for the *indicator mapping* of Λ defined by $\Delta(z;\Lambda) = 0$ if $z \in \Lambda$ and $\Delta(z;\Lambda) = \emptyset$ otherwise. Applying the coderivative criterion of metric regularity from Mordukhovich and Shao [15, Theorem 5.6], for any $\varepsilon > 0$ we find $\widetilde{z} \in \bar{z}^N + \varepsilon I\!B$ and $\psi^* = (\psi_0^*, \ldots, \psi_{N-1}^*) \in (X^*)^N$ such that

$$0 \in \widehat{D}^* g_\Lambda(\widetilde{z})(\psi^*) \quad \text{with} \quad \|\psi^*\| > 1.$$

By Lemma 4.1 and elementary coderivative calculus involving a smooth mapping g, we have

$$\begin{aligned} 0 \ &\in \widehat{D}^* g_\Lambda(\widetilde{z})(\psi^*) = \widehat{D}^*\big[g(\cdot) + \Delta(\cdot;\Lambda)\big](\widetilde{z})(\psi^*) = \nabla g(\widetilde{z})^*\psi^* + \widehat{N}(\widetilde{z};\Lambda) \\ &\subset \sum_{j=0}^{N-1} \nabla g_j(\widetilde{z})^* \psi_j^* + \widehat{N}(z_0;\Lambda_0) + \cdots + \widehat{N}(z_N;\Lambda_N) + \varepsilon I\!B^* \end{aligned}$$

with some $z_j \in \Lambda_j \cap (\widetilde{z} + \varepsilon I\!B)$ as $j = 0, \ldots, N$. The latter implies (4.5)–(4.7) with $\mu_0 = 0$ and thus completes the proof of the theorem. $\triangle$

Based on the above necessary optimality conditions for problems of mathematical programming, we now derive the following "fuzzy" necessary optimality conditions in the extended *Euler-Lagrange* form for discrete approximations of the original problem.

Theorem 4.3 (fuzzy Euler-Lagrange conditions for discrete approximations). *Let $\bar{x}^N(\cdot) = (x_0, \bar{x}_1^N, \cdots, \bar{x}_N^N)$ be an optimal solution to problem (P_N) with any fixed $N \in I\!N$. Assume that the cost function φ is locally Lipschitzian and that the sets Ω and $\operatorname{gph} F(\cdot, t_j)$, $j = 0, \ldots, N-1$, are locally closed around $\bar{x}_N(\cdot)$. Then, given an arbitrary $\varepsilon > 0$, there exist a number $\lambda^N \geq 0$ and a discrete adjoint trajectory $p^N(\cdot) = (p_1^N, \ldots, p_N^N) \in (X^*)^N$ satisfying the following relations:*

— *the fuzzy Euler-Lagrange inclusion: there are (x_j^N, y_j^N), $(\widetilde{x}_j^N, \widetilde{y}_j^N) \in (\bar{x}_j^N, \bar{y}_j^N) + \varepsilon I\!B$, and $a_j^* \in I\!B^*$ for $j = 1, \ldots, N-1$ such that*

$$\Big(\frac{e^{A^* h_N} p_{j+1}^N - p_j^N}{h_N} - \lambda^N \theta_j^N a_j^*,\ p_{j+1}^N\Big) \in \widehat{N}((x_j^N, y_j^N); \operatorname{gph}(e^{A h_N} F(\cdot, t_j))) + \varepsilon I\!B^* \tag{4.9}$$

as $j = 1, \ldots, N-1$, where the numbers θ_j^N are defined by

$$\theta_j^N := 2\Big\| e^{A h_N} \widetilde{x}_j^N + h_N \widetilde{y}_j^N - \bar{x}(t) \Big\|, \quad j = 1, \ldots, N-1; \tag{4.10}$$

—*the fuzzy transversality inclusion: there are* $x_N^N,\ \widetilde{x}_N^N \in \bar{x}_N^N + \varepsilon I\!B$ *such that*

$$-p_N^N \in \lambda^N \widehat{\partial}\varphi(\widetilde{x}_N^N) + \widehat{N}(x_N^N;\Omega_N) + h_N\varepsilon I\!B^*; \tag{4.11}$$

—*the nontriviality condition:*

$$\lambda^N + \sum_{j=1}^{N} \|p_j^N\| \geq 1 \quad \textit{for all} \quad N \in I\!N. \tag{4.12}$$

Proof. Apply Theorem 4.2 to the optimal solution

$$\bar{z} := \bar{z}^N = (x_0, \bar{x}_1^N, \ldots, \bar{x}_N^N, \bar{y}_0^N, \ldots, \bar{y}_{N-1}^N)$$

for problem (4.1)–(4.4) as $N \in I\!N$, where $\bar{x}^N = (x_0, \bar{x}_1^N, \ldots, \bar{x}_N^N)$ is a given optimal solution to (P_N). According to this result, there are a real number $\mu_0 \geq 0$ and adjoint vectors $(\psi_0^*, \ldots, \psi_{N-1}^*) \in (X^*)^N$ and $z_j^* \in (X^*)^{2N+1}$, $j = 0, \ldots, N$, satisfying the extended Lagrange-type relations (4.5)–(4.7). Taking into account the structure of Λ_j in (4.3) and (4.4), present z_j^* and the corresponding vectors z_j from (4.5) as

$$z_j = (x_0, x_{1j}^N, \ldots, x_{Nj}^N, y_{0j}^N, \ldots, y_{N-1j}^N) \in X^{2N+1} \quad \text{and} \quad z_j^* = (x_{0j}^*, x_{1j}^*, \ldots, x_{Nj}^*, y_{0j}^*, \ldots, y_{N-1j}^*).$$

It is easy to derive from (4.3)–(4.5) the following relationships:

$$\begin{cases} (x_{jj}^*, y_{jj}^*) \in \widehat{N}\big((x_{jj}^N, y_{jj}^N); \operatorname{gph}(e^{Ah_N}F(\cdot, t_j))\big), \\ x_{ij}^* = y_{ij}^* = 0 \ \text{ otherwise}, \quad j = 0, \ldots, N-1; \end{cases} \tag{4.13}$$

$$x_{NN}^* \in \widehat{N}(x_{NN}^N; \Omega_N) \ \text{ and } \ x_{iN}^* = y_{iN}^* = 0 \ \text{ otherwise} \tag{4.14}$$

with some $z_j^N \in \bar{z}^N + \varepsilon I\!B$. Further, by the structure of g_j in (4.2) we observe that

$$\begin{aligned} \sum_{j=0}^{N-1} \nabla g_j(z)^* \psi_j^* &= \big(- e^{A^* h_N}\psi_0^*, \psi_0^* - e^{A^* h_N}\psi_1^*, \ldots, \psi_{N-2}^* - e^{A^* h_N}\psi_{N-1}^*, \psi_{N-1}^*, \\ &\quad -h_N\psi_0^*, -h_N\psi_1^*, \ldots, -h_N\psi_{N-1}^* \big) \end{aligned} \tag{4.15}$$

for any $z \in X^{2N+1}$. Then applying the extended fuzzy Lagrange multiplier rule (4.6) with the notation $\lambda^N := \mu_0 \geq 0$ and then the fuzzy sum rule (3.8) for the cost function ϕ_0 in (4.1) with taking into account its specific structure as well as the above relationships (4.13)–(4.15), we arrive at the inclusions

$$-x_{00}^* \in -e^{A^* h_N}\psi_0^* + h_N\varepsilon I\!B^*, \tag{4.16}$$

$$-x_{jj}^* \in h_N\lambda^N\theta_j^N I\!B^* + \psi_{j-1}^* - e^{A^* h_N}\psi_j^* + h_N\varepsilon I\!B^* \ \text{ for } \ j = 1, \ldots, N-1, \tag{4.17}$$

$$-x_{NN}^* \in \lambda^N \widetilde{x}_N^* + \psi_{N-1}^* + h_N\varepsilon I\!B^*, \tag{4.18}$$

$$-y_{jj}^* \in -h_N\psi_j^* + h_N\varepsilon I\!B^* \ \text{ for } \ j = 0, \ldots, N-1, \tag{4.19}$$

where the numbers θ_j^N are defined in (4.10) with some $(\widetilde{x}_j^N, \widetilde{y}_j^N) \in (\bar{x}_j^N, \bar{y}_j^N) + \varepsilon I\!B$ for $j = 1, \ldots, N-1$, and where

$$\widetilde{x}_N^* \in \widehat{\partial}\varphi(\widetilde{x}_N^N) \text{ with some } \widetilde{x}_N^N \in \bar{x}_N^N + \varepsilon I\!B.$$

Finally, by changing the notation in (4.13), (4.14), and (4.16)–(4.19) to

$$(x_j^N, y_j^N) = (x_{jj}^N, y_{jj}^N),\ (x_j^*, y_j^*) := (x_{jj}^*, y_{jj}^*),\ j = 0, \ldots, N, \text{ and } p_j^N := \psi_{j-1}^*,\ j = 1, \ldots, N,$$

we arrive at (4.9), (4.11), and (4.12), which ends the proof of the theorem. △

The nontriviality condition (4.12) in Theorem 4.3 can be essentially improved under the additional assumptions on F, which are parts of our standing hypotheses.

Corollary 4.4 (fuzzy Euler-Lagrange conditions with enhanced nontriviality). *In addition to the assumptions of Theorem 4.3, suppose that for each $j = 0, \ldots, N-1$ the multifunction $F(\cdot, t_j)$ is compact-valued and Lipschitz continuous around $\bar{x}_j^N$. Then there is a number $\gamma > 0$ independent of N and such that for some sequences of natural numbers $N \to \infty$ and positive numbers $\varepsilon_N \downarrow 0$ there are multipliers λ^N and adjoint trajectories $p^N(\cdot)$ satisfying (4.9)–(4.11) with $\varepsilon = \varepsilon_N$ and the enhanced nontriviality condition*

$$\lambda^N + \|p_N^N\| \ge \gamma \ \text{ as } \ N \to \infty. \tag{4.20}$$

Proof. It follows from the proof of Theorems 4.2 and 4.3 that either $\lambda^N = 1$ or $\lambda^N = 0$ for all $N \in I\!N$. It remains to show that if $\lambda^N = 0$, then $\|p_N^N\| \ge \gamma$ for some number $\gamma > 0$ and all N sufficiently large. To proceed, we first estimate $\|p_j^N\|$ via $\|p_N^N\|$ when $\lambda^N = 0$. Indeed, in the latter case the Euler-Lagrange inclusion (4.9) can be written in terms of the coderivative (3.4) as

$$\frac{e^{A^* h_N} p_{j+1}^N - p_j^N}{h_N} - \varepsilon b_j^* \in \widehat{D}^*(e^{Ah_N} F(\cdot, t_j))(x_j^N, y_j^N)\big(-p_{j+1}^N + \varepsilon c_j^*\big) \tag{4.21}$$

for all $j = 0, \ldots, N-1$, where $b_j^*, c_j^* \in I\!B^*$. Since $\{e^{At} |\ t \ge 0\}$ is a C_0-semigroup of contractions, we have $\|e^{A^* h_N}\| \le 1$. Involving the *coderivative characterization* (3.5) of the local Lipschitzian property for compact-valued multifunctions, we derive from (4.21) that

$$\begin{aligned}
\|p_j^N\| &\le \|e^{A^* h_N} p_{j+1}^N - p_j^N - h_N \varepsilon b_j^*\| + \|e^{A^* h_N} p_{j+1}^N - h_N \varepsilon b_j^*\| \\
&\le (1 + h_N \ell_F)\|p_{j+1}^N\| + h_N(1 + \ell_F)\varepsilon \\
&\le \cdots\cdots\cdots \\
&< \exp\big[\ell_F(b-a)\big]\big[\|p_N^N\| + (b-a)(1+\ell_F)\varepsilon\big].
\end{aligned} \tag{4.22}$$

Suppose that the nontriviality condition (4.20) does not hold along with (4.9) and (4.11) in the case of $\lambda^N = 0$ under consideration. Take a sequence $\gamma_k \downarrow 0$ as $k \to \infty$ and choose numbers N_k and ε_k such that

$$N_k := \big[1/\gamma_k\big],\ \ \varepsilon_k \le \gamma_k^2, \text{ and } \|p^N(t_N)\| \le \gamma_k^2$$

for $k \in I\!N$, where $[\cdot]$ stands for the greatest integer less than or equal to the given real number. By the adjoint trajectory estimate (4.22) we have

$$\sum_{j=1}^{N_k} \|p^{N_k}(t_j)\| \le N_k \gamma_k \exp\big[\ell_F(b-a)\big] + \varepsilon_k N_k (b-a)(1+\ell_F)$$

$$\le \gamma_k \exp\big[\ell_F(b-a)\big] + \gamma_k (b-a)(1+\ell_F) \downarrow 0 \quad \text{as} \quad k \in I\!N,$$

which contradicts (4.12) with $\lambda^N = 0$. This completes the proof of the corollary. $\triangle$

5 Optimality conditions for semilinear evolution inclusions

In this section we derive necessary optimality conditions for the original problem (P) governed by infinite-dimensional evolution inclusions by passing to the limit from those for the discrete-time problems (P_N) established in Section 4. Our limiting procedure is based on the stability/convergence results for discrete approximations obtained in Section 2 and the robust tools of generalized differentiation reviewed in Section 3. A crucial component of the variational analysis developed in this section is justifying an appropriate convergence of *adjoint arcs* in the limiting procedure from discrete approximations. This is mainly based on the above *coderivative criterion* for the local Lipschitzian property of set-valued mappings. To furnish the limiting process, we keep in this section all the standing assumptions imposed in Section 2 with adding the following requirements on the cost function and target/constraint set in (1.1) and (1.3) around the optimal endpoint under consideration:

(H3) The cost function φ is Lipschitz continuous around $\bar{x}(b)$ and the target set Ω is SNC at this point.

Note that the Lipschitzian requirement on φ can be weakened to the lower semicontinuity (with some change in the transversality condition; cf. Mordukhovich [12, 14]), while the SNC requirement is very essential in infinite dimensions. It has been well recognized that necessary optimality conditions of the Pontryagin maximum principle type do not hold even in simple control problems for the heat equation with a singleton target set, which is never SNC in infinite-dimensional spaces; cf. Fattorini [6] and Li and Yong [10].

To formulate the main result, consider the *Hamiltonian* function

$$H(x,p,t) := \sup\big\{\langle p,v\rangle \big|\ v \in F(x,t)\big\}, \quad p \in X^*,$$

and form the *argmaximum* sets defined by

$$M(x,p,t) := \big\{v \in F(x,t) \big|\ \langle p,v\rangle = H(x,p,t)\big\}.$$

In what follows we use the (limiting) normal/coderivative/subdifferential constructions of Section 3 with respect to all the variables but t.

Theorem 5.1 (extended Euler-Lagrange conditions for semilinear evolution inclusions). *Let $\bar{x}(\cdot)$ be an optimal solution to the continuous-time problem (P) under the standing assumptions made. Then there exist a number $\lambda \geq 0$ and a weakly continuous arc $p\colon [a,b] \to X^*$, not both zero, satisfying:*

—the extended Euler-Lagrange inclusion

$$\begin{cases} p(t) \in e^{A^*(b-t)}p(b) + \displaystyle\int_b^t e^{A^*(s-t)}\Big\{u \in X^* \Big| \ (u, p(s)) \in \widetilde{N}((\bar{x}(s), v); \operatorname{gph} F(\cdot, s)), \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad v \in M(\bar{x}, p, s)\Big\}\, ds \quad \text{for all} \ \ t \in [a,b], \end{cases} \tag{5.1}$$

which is equivalent to

$$p(t) \in e^{A^*(b-t)}p(b) + \int_b^t e^{A^*(s-t)}D^*F(\bar{x}(s), v)(-p(s))\Big| \ v \in M(\bar{x}, p, s)\Big\}\, ds \tag{5.2}$$

if $\operatorname{gph} F(\cdot, t)$ *is normally semicontinuous at* $(\bar{x}(t), v)$ *for all* $v \in M(\bar{x}(t), p(t), t)$ *and a.e.* $t \in [a,b]$, *in particular, if* F *is autonomous;*

—the Weierstrass-Pontryagin maximum condition

$$\langle p(t), \dot{\bar{x}}(t)\rangle = H(\bar{x}(t), p(t), t) \quad a.e. \quad t \in [a,b]; \tag{5.3}$$

—the transversality condition

$$-p(b) \in \lambda\partial\varphi(\bar{x}(b)) + N(\bar{x}(b); \Omega). \tag{5.4}$$

Proof. Assume without loss of generality that the operator A generates a C_0-semigroup $\{e^{At} |\ t \geq 0\}$ of contractions on X and build a sequence of discrete approximations (P_N) for (P), which approximates $\bar{x}(\cdot)$ in the sense of Theorem 2.2. Employing necessary conditions of Corollary 4.4 for optimal solutions $\bar{x}^N(\cdot) = (x_0, \bar{x}_1^N, \ldots, \bar{x}_N^N)$ to (P_N), we find sequences of numbers $\lambda^N \geq 0$ and adjoint discrete trajectories $p^N(\cdot) = (p_1^N, \ldots, p_N^N)$ satisfying conditions (4.9)–(4.11) and (4.20) with some $\varepsilon_N \downarrow 0$ as $N \to \infty$. Observe that the nontriviality condition (4.20) can be equivalently written as

$$\lambda^N + \|p^N(b)\| = 1 \ \text{ for all } \ N \in I\!N, \tag{5.5}$$

since the number $\gamma > 0$ therein is independent of N. We may always assume that λ^N converge to some $\lambda \geq 0$ as $N \to \infty$.

As usual, the notation $x^N(t)$ and $p^N(t)$ stand for the piecewise linear extensions of the corresponding discrete functions on $[a,b]$. From the proof of Theorem 4.3 and Corollary 4.4 we observe that the adjoint trajectories $p^N(t)$ are uniformly bounded on $[a,b]$. Furthermore, it follows from (4.9) that the functions

$$p^N(t) = e^{A^*(b-t)}p^N(b) - \int_t^b e^{A^*(s-t)}g^N(s)\, ds \tag{5.6}$$

are the piecewise linear extensions of $p^N = (p_1^N, \ldots, p_N^N)$ on $[a,b]$, where

$$\begin{aligned} g^N(t) \in \ & \Big\{u \in X^* \Big| \ (u, p^N(t_{j+1})) \in \widehat{N}\big((x^N(t_j), y_j^N); \operatorname{gph}(e^{Ah_N}F(\cdot, t_j))\big) + \varepsilon_N I\!B^* \Big\} \\ & + \lambda^N\theta_j^N I\!B^* \ \text{ for } \ t \in [t_j, t_{j+1}), \quad j = 0, \ldots, N-1. \end{aligned} \tag{5.7}$$

It follows from (5.7) and the coderivative criterion (3.5) that the functions $g^N(\cdot)$ are uniformly bounded in $L^2([a,b];X)$, and hence they weakly converge to some $g(\cdot)\in L^2([a,b];X)$. By Theorem 2.2 we have that $x^N(t)\to\bar{x}(t)$ uniformly on $[a,b]$, and hence $\theta_j^N\to 0$ for all $j=1,\ldots,N-1$ as $N\to\infty$. Observe also that

$$\left\{\begin{array}{l}\displaystyle\lim_{N\to\infty} e^{A^*(b-t)}p^N(b)=e^{A^*(b-t)}p(b)\ \text{ and}\\ \displaystyle\lim_{N\to\infty}\int_t^b e^{A^*(s-t)}g^N(s)\,ds=\int_t^b e^{A^*(s-t)}g(s)\,ds,\end{array}\right. \tag{5.8}$$

since the C_0-semigroup $\{e^{At}|\ t\ge 0\}$ is compact, which implies the compactness of the one $\{e^{A^*t}|\ t\ge 0\}$. By (5.8) and the weak continuity of the Bochner integral as a linear operator from $L^2([a,b];X^*)$ to X^*, we get from (5.6) by passing to the limit as $N\to\infty$ that the adjoint arcs $p^N(t)$ weakly converge for each $t\in[a,b]$ to some function $p(t)\in X^*$, which is weakly continuous on $[a,b]$. Furthermore, (5.7) and the convexity of the sets $F(x_N(t_j),t_j)$ whenever $j=0,\ldots,N-1$ imply that

$$\begin{array}{l}g^N(t)\in\Big\{w\in X^*\Big|\ \ (w,e^{A^*h_N}(p^N(t_{j+1})-\varepsilon_N b_N^*))\in\widehat{N}((x^N(t_j),v);\operatorname{gph}F(\cdot,t_j)),\\ \qquad\qquad v\in M\big(x^N(t_j),e^{A^*h_N}(p^N(t_{j+1})-\varepsilon_N b_N^*),t_j\big)\Big\}+\lambda^N\theta_j^N I\!B^*,\end{array} \tag{5.9}$$

where $w:=u-\varepsilon_N a_N^*$ with some $a_N^*,b_N^*\in I\!B^*$. Passing to the limit in (5.9) as $N\to\infty$ and using the classical Mazur weak closure theorem, we arrive at

$$\begin{array}{l}g(t)\in\operatorname{co}\Big\{\ \ u\in X^*\Big|\ (u,p(t))\in\widetilde{N}((\bar{x}(t),v);\operatorname{gph}F(\cdot,t)),\\ \qquad\qquad\qquad v\in M(\bar{x}(t),p(t),t)\Big\}\ \text{ a.e. }\ t\in[a,b],\end{array} \tag{5.10}$$

where the closure operation for the convex hull in (5.10) can be omitted due to the reflexivity of the space X. Then passing to the limit in (5.6) as $N\to\infty$ with taking into account (5.8) and (5.10), we obtain the inclusion

$$\left\{\begin{array}{l}\displaystyle p(t)\in e^{A^*(b-t)}p(b)+\int_b^t e^{A^*(s-t)}\operatorname{co}\Big\{u\in X^*\Big|\ \ (u,p(s))\in\widetilde{N}((\bar{x}(s),v);\operatorname{gph}F(\cdot,s)),\\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad v\in M(\bar{x}(s),p(s),s)\Big\}\,ds\ \text{ for all }\ t\in[a,b],\end{array}\right.$$

where the convexity operation under the integral can be omitted due to the fundamental Lyapunov-Aumann integration theorem in reflexive and separable Banach spaces (see, e.g., Diestel and Uhl [4] and Tolstonogov [19]) by the above compactness arguments involving the compact semigroup $\{e^{A^*t}|\ t\ge 0\}$. Thus we have the extended Euler-Lagrange inclusion (5.1), which automatically implies the maximum condition (5.3) as well as the coderivative form (5.2) under the normal semicontinuity. The transversality condition (5.4) follows from (4.11) by passing to the limit as $N\to\infty$ and taking into account the structure (2.15) of the set Ω_N with $\eta_N\downarrow 0$.

It remains to justify the nontriviality condition $(p(\cdot),\lambda)\ne 0$ under the SNC assumption imposed on the set Ω at $\bar{x}(b)$. Supposing the contrary, we have that $p^N(b)\xrightarrow{w}0$ as $N\to\infty$. Due to the fuzzy transversality condition (4.11) with $p_N^N=p^N(b)$ and the convergence $\lambda^N\to 0$ and $x_N^N\to\bar{x}(b)$, the latter implies that $\|p^N(b)\|\to 0$ as $N\to\infty$ by the SNC property of Ω and the structure of Ω_N. This contradicts the discrete nontriviality condition (5.5) for large $N\in I\!N$ and completes the proof of the theorem. △

References

[1] AHMED, N.U. (1991) *Semigroup Theory with Applications to Systems and Control.* Longman, Harlow, UK.

[2] BOUNKHEL, M. and THIBAULT, L. (2005), Further characterizations of regular sets in Hilbert spaces and their applications to nonconvex sweeping processes. To appear in *SIAM J. Optim.*

[3] CLARKE, F.H. (2004) Necessary conditions in dynamic optimization. To appear in *Mem. Amer. Math. Soc.*

[4] DIESTEL, J. and UHL, J.J. (1977) *Vector Measures.* American Mathematical Society, Providence, Rhode Island.

[5] DONTCHEV, A.L. and FARKHI, E.M. (1989) Error estimates for discretized differential inclusions. *Computing*, **41**, 349–358.

[6] FATTORINI, H.O. (1999) *Infinite-Dimensional Optimization and Control Theory.* Cambridge University Press, Cambridge, UK.

[7] FRANKOWSKA, H. (1990) A priori estimates for operational differential inclusions. *J. Diff. Eq.*, **84**, 100–128.

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

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

[10] LI, X.J. and YONG, J. (1995) *Optimal Control Theory for Infinite-Dimensional Systems.* Birkhäuser, Boston.

[11] LOEWEN, P.D. and ROCKAFELLAR, R.T. (1996) New necessary conditions for the generalized problem of Bolza. *SIAM J. Control Optim.*, **34**, 1496–1511.

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

[13] MORDUKHOVICH, B.S. (2004) Optimal control of evolution inclusions. To appear in *Nonlinear Anal.*

[14] MORDUKHOVICH, B.S. (2005) *Variational Analysis and Generalized Differentiation, Vol. I: Basic Theory, Vol. II: Applications*, To appear in Springer, Berlin.

[15] MORDUKHOVICH, B.S. and SHAO, Y. (1995) Differential characterizations of covering, metric regularity, and Lipschitzian properties of multifunctions between Banach spaces. *Nonlinear Anal.*, **25**, 1401–1424.

[16] MORDUKHOVICH, B.S. and WANG, D. (2005) Optimal control of semilinear unbounded differential inclusions. To appear in *Proc. Fourth World Congress Nonlinear Anal.*

[17] ROCKAFELLAR, R.T. and WETS, R. J-B. (1998) *Variational Analysis.* Springer, Berlin.

[18] SMIRNOV, G.V. (2002) *Introduction to the Theory of Differential Inclusions.* American Mathematical Society, Providence, Rhode Island.

[19] TOLSTONOGOV, A.A. (2000) *Differential Inclusions in a Banach Spaces.* Kluwer, Dordrecht, The Netherlands.

[20] VINTER, R.B. (2000) *Optimal Control.* Birkhäuser, Boston.