1-1-2003

Optimal Control of Neutral Functional-Differential Inclusions

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

Lianwen Wang
Wayne State University

OPTIMAL CONTROL OF NEUTRAL FUNCTIONAL-DIFFERENTIAL INCLUSIONS

BORIS S. MORDUKHOVICH and LIANWEN WANG

WAYNE STATE UNIVERSITY

Detroit, MI 48202

Department of Mathematics

Research Report

2003 Series
#1

This research was partly supported by the National Science Foundation.

OPTIMAL CONTROL OF NEUTRAL FUNCTIONAL-DIFFERENTIAL INCLUSIONS[1]

BORIS S. MORDUKHOVICH and LIANWEN WANG

Department of Mathematics, Wayne State University, Detroit, MI 48202

boris@math.wayne.edu, lwang@math.wayne.edu

Dedicated to Jack Warga in honor of his 80th birthday

Abstract. This paper deals with optimal control problems for dynamical systems governed by constrained functional-differential inclusions of neutral type. Such control systems contain time-delays not only in state variables but also in velocity variables, which make them essentially more complicated than delay-differential (or differential-difference) inclusions. Our main goal is to derive necessary optimality conditions for general optimal control problems governed by neutral functional-differential inclusions with endpoint constraints. While some results are available for smooth control systems governed by neutral functional-differential equations, we are not familiar with any results for neutral functional-differential inclusions, even with smooth cost functionals in the absence of endpoint constraints. Developing the method of discrete approximations (which is certainly of independent interest) and employing advanced tools of generalized differentiation, we conduct a variational analysis of neutral functional-differential inclusions and obtain new necessary optimality conditions of both Euler-Lagrange and Hamiltonian types.

Key words. optimal control, functional-differential inclusions of neutral type, variational analysis, discrete approximations, generalized differentiation, necessary optimality conditions

AMS subject classification. 49K24, 49K25, 49J52, 49M25, 90C31.

1 Introduction

This paper concerns the study of optimal control problems for the so-called *neutral functional-differential inclusions*, which contain time-delays in both state and velocity variables. Such inclusions belong to the broad class of *hereditary* systems known also as systems with *memory* or *aftereffect*. They have been investigated in the form of controlled functional-differential *equations* being important for various practical applications, particularly to problems of automatic control, economic dynamics, modeling of ecological, biological, and chemical processes, etc.; see examples and discussions in [1, 4, 12, 10, 14, 15, 19] and their references. Note that some classes of *hyperbolic PDEs* can be reduced to neutral functional-differential equations, as shown in the above references.

To our knowledge, control problems for neutral functional-differential *inclusions* have not been

[1]Research was partly supported by the National Science Foundation under grant DMS-0072179.

sufficiently studied in the literature. We are only familiar with results concerning existence of optimal solutions, local controllability, and relaxation procedures mostly collected in [14].

In this paper we consider the following dynamic optimization (generalized optimal control) problem (P):

$$\text{(1.1)} \qquad \text{minimize} \quad J[x] := \varphi(x(a), x(b)) + \int_a^b f(x(t), x(t-\Delta), t)\, dt$$

over feasible arcs $x : [a-\Delta, b] \to I\!R^n$, which are continuous on $[a-\Delta, a)$ and $[a, b]$ (with a possible jump at $t = a$) and such that the combination $x(\cdot) - Ax(\cdot - \Delta)$ is absolutely continuous on $[a, b]$, satisfying the *neutral functional-differential inclusion*

$$\text{(1.2)} \qquad \frac{d}{dt}[x(t) - Ax(t-\Delta)] \in F(x(t), x(t-\Delta), t) \quad \text{a.e. } t \in [a, b],$$

$$\text{(1.3)} \qquad x(t) = c(t), \quad t \in [a-\Delta, a),$$

with the endpoint constraints

$$\text{(1.4)} \qquad (x(a), x(b)) \in \Omega \subset I\!R^{2n}.$$

We always assume that $F\colon I\!R^n \times I\!R^n \times [a, b] \rightrightarrows I\!R^n$ is a set-valued mapping of closed graph, Ω is a closed set, $\Delta \geq 0$ is a constant delay, and A is a constant $n \times n$ matrix.

Note that the neutral-type operator in the left-hand side of (1.2) is given in the *Hale form* [10] and that trajectories of the neutral inclusion may be assumed to be discontinuous not only at $t = a$ but also at the points $t = a + j\Delta \in [a, b]$, $j = 1, 2, \ldots$. Moreover, the results obtained in this paper can be easily extended to problems with the cost function φ depending on $x(a + j\Delta)$, the constraints (1.4) given at these intermediate points, and the integrand f in (1.1) depending on the velocity $\dot{z}(t)$ for $z(t) := x(t) - Ax(t-\Delta)$, $t \in [a, b]$. We can also consider the cases of *multiple delays* $\Delta_1 \geq \Delta_2 \geq \ldots \geq \Delta_m \geq 0$ as well as *variable delays* $\Delta(t) \geq 0$, where $\Delta(\cdot)$ is a Lipschitz continuous (hence a.e. differentiable) function satisfying the assumption

$$|\dot{\Delta}(t)| < \alpha \in (0, 1) \quad \text{a.e.} \quad t \in [a, b],$$

which ensures that the function $t - \Delta(t)$ is invertible on $[a, b]$. For simplicity we focus in what follows on problem (P) formulated in (1.1)–(1.4).

Our primary goal is to derive *necessary optimality conditions* for problem (P) under general assumptions on the initial data. For *nondelayed* systems governed by differential inclusions ($\Delta = 0$, $A = 0$) necessary optimality conditions have been studied intensively during recent years; see [5, 11, 16, 21, 27, 28, 29, 31] and the references therein. Some results are known for *delay-differential* (or differential-difference) inclusions corresponding to $A = 0$ in (1.2); see [6, 7, 17, 23, 24]. We are not familiar with any necessary optimality conditions obtained for problem (P) governed by neutral functional-differential inclusions with $A \neq 0$ in (1.2) besides the case of *smooth* control systems corresponding to

$$F(x,y,t) = \{v \in I\!R^n |\ v = g(x,y,u,t),\ u \in U\} \tag{1.5}$$

with continuously differentiable functions φ, f, g in (1.1) and (1.2) as well as those describing end-point constraints; see [3, 9, 13, 19] and their references.

Observe that neutral-type systems are essentially different from their counterparts with $A = 0$. In particular, it is well known that an analog of the Pontryagin maximum principle does *not* generally hold for neutral systems, even in the classical smooth framework of (1.5), with *no convexity* assumptions. In a sense, neutral-type systems combine properties of continuous-time and discrete-time control systems; indeed, they can be treated as discrete-time systems regarding velocity variables. On the other hand, neutral systems have some similarities with the so-called *hybrid* and *algebraic-differential* equations important in engineering control applications.

In this paper we derive necessary optimality conditions for the neutral-type control problem (P) under general assumptions on its initial data involving nonsmooth functions and nonconvex sets. These conditions are obtained in extended *Euler-Lagrange* and *Hamiltonian* forms involving advanced generalized differential constructions of variational analysis; see Section 4. Note that the results obtained seem to be new even in the case of nondelayed problems with $A \neq 0$ corresponding to *implicit* differential inclusions.

Our approach is based on the *method of discrete approximations*, in the line developed in [19, 21] for nondelayed differential inclusions and in [23, 24] for delay-differential systems with $A = 0$. This method, which is certainly of independent interest from both qualitative and numerical viewpoints, allows us to construct a *well-posed* parametric family of optimal control problems for approximating systems governed by *discrete-time* analogs of neutral functional-differential inclusions. A crucial issue is to establish *stability* of such approximations that ensures an appropriate *strong convergence* of optimal solutions. Convergence analysis of this method and its application to necessary optimality conditions for neutral systems are essentially more involved in comparison with the cases of differential and delay-differential inclusions.

The approximating discrete-time control problems can be reduced to special *finite-dimensional* problems of *nonsmooth programming* with an increasing number of *geometric constraints* that may have empty interiors. To handle such problems, we use suitable generalized differential tools of variational analysis satisfying a comprehensive calculus that allows us to derive general necessary optimality conditions for finite-difference analogs of neutral functional-differential inclusions. Then passing to the limit from well-posed discrete approximations with the strong convergence of optimal solutions and employing generalized differential calculus, we obtain necessary optimality conditions for the original problem (P).

The rest of the paper is organized as follows. In Section 2 we show that some combination built upon a given admissible trajectory of the neutral inclusion (1.2) can be *strongly approximated* by the corresponding combination built upon admissible trajectories of discrete-time systems. This result important for its own sake plays a crucial role in the construction of well-posed discrete approximations to the original problem (P) and in the subsequent justification of a strong convergence of their optimal solutions to the given optimal trajectory for (P).

Such a convergence analysis is conducted in Section 3 for a sequence of well-posed discrete approximations to (P) involving an appropriate perturbation of the endpoint constraints (1.4) that is consistent with the step of discretization. The required strong convergence of optimal solutions is established under an intrinsic property of the original problem (P) called *relaxation stability*. This property imposing the equality between the optimal values in (P) and its relaxation (convexification) goes far beyond the convexity assumption on the velocity sets $F(x,y,t)$.

Section 4 contains the basic constructions and required material on generalized differentiation needed for performing a variational analysis of discrete-time and continuous-time optimal control problems in the subsequent sections. These constructions and calculus rules are used in Section 5 for deriving general necessary optimality conditions for nonconvex discrete-time inclusions arising in discrete approximations of the original problem (P). The main results on the extended Euler-Lagrange and Hamiltonian conditions for neutral functional-differential inclusions are derived in Section 6 via passing to the limit from discrete approximations.

Our notation is basically standard; cf. [21] and [26]. Recall that, given a set-valued mapping (multifunction) $F\colon X \rightrightarrows Y$ between finite-dimensional spaces, the *Painlevé-Kuratowski upper/outer limit* of $F(x)$ as $x \to \bar{x}$ is defined by

$$\operatorname*{Lim\,sup}_{x\to\bar{x}} F(x) := \big\{y \in Y \mid \exists\, x_k \to \bar{x},\ \exists\, y_k \to y \ \text{ with } \ y_k \in F(x_k) \ \text{ for all } \ k \in I\!N\big\},$$

where $I\!N$ stands for the collection of all natural numbers.

2 Discrete approximations of neutral inclusions

This section concerns the study of discrete approximations of an *arbitrary* admissible trajectory to the neutral functional-differential inclusion (1.2) with the initial condition (1.3). We show that, under fairly general assumptions, any admissible trajectory to (1.2) and (1.3) can be *strongly approximated* in the sense indicated below by the corresponding trajectories to finite-difference inclusions obtained from (1.2) by the Euler scheme. This result is constructive providing efficient estimates for the approximation rate, and hence it is certainly of independent interest for numerical analysis and applications.

Let $\bar{x}(\cdot)$ be an *admissible trajectory* in (P), i.e., it is continuous on $[a-\Delta, a)$ and $[a,b]$ (with a possible jump at $t=a$), the combination $x(\cdot) - Ax(\cdot - \Delta)$ is absolutely continuous on $[a,b]$, and relations (1.2) and (1.3) are satisfied. Note that the endpoint constraints (1.4) may not hold for $\bar{x}(\cdot)$; if they do hold, $\bar{x}(\cdot)$ is *feasible* to (P). The following *standing assumptions* are imposed throughout the paper:

(H1) There are an open set $U \subset I\!R^n$ and two positive numbers ℓ_F and m_F such that $\bar{x}(t) \in U$ for all $t \in [a-\Delta, b]$, the sets $F(x,y,t)$ are closed, and one has

$$F(x,y,t) \subset m_F I\!B \ \text{ for all } \ (x,y,t) \in U \times U \times [a,b],$$
$$F(x_1,y_1,t) \subset F(x_2,y_2,t) + \ell_F(|x_1-x_2| + |y_1-y_2|)I\!B$$

if $(x_1,y_1),(x_2,y_2) \in U \times U$ and $t \in [a,b]$, where $I\!B$ stands for the closed unit ball in $I\!R^n$.

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

(H3) The function $c(\cdot)$ is continuous on $[a-\Delta, a]$.

Following [8], we consider the so-called *averaged modulus of continuity* for the multifunction $F(x,y,t)$ with $(x,y) \in U \times U$ and $t \in [a,b]$ that is defined by

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

where $\sigma(F;t,h) := \sup\left\{\vartheta(F;x,y,t,h)\,\big|\,(x,y) \in U \times U\right\}$ with

$$\vartheta(F;x,y,t,h) := \sup\left\{\text{haus}(F(x,y,t_1),F(x,y,t_2))\,\Big|\,(t_1,t_2) \in \left[t-\frac{h}{2}, t+\frac{h}{2}\right] \cap [a,b]\right\},$$

and where haus$(\cdot,\cdot)$ stands for the Hausdorff distance between two compact sets. It is proved in [8] that if $F(x,y,t)$ is Hausdorff continuous for a.e. $t \in [a,b]$ uniformly in $(x,y) \in U \times U$, then $\tau(F;h) \to 0$ as $h \to 0$. This fact is essentially used in the sequel.

Let us construct a sequence of discrete approximations of the given neutral-differential inclusion replacing the derivative in (1.2) by the *Euler finite difference*

$$\frac{d}{dt}[x(t) - Ax(t-\Delta)] \approx \frac{x(t+h) - Ax(t+h-\Delta) - x(t) + Ax(t-\Delta)}{h}.$$

In what follows we *assume that* $\Delta > 0$; the case of $\Delta = 0$ can be treated by the limiting procedure as $\Delta \downarrow 0$; see Remark 6.3. For any $N \in I\!N := \{1,2,\ldots\}$ we consider the *step of discretization* $h_N := \frac{\Delta}{N}$ and define the *discrete grid/partition* $t_j := a + jh_N$ as $j = -N,\ldots,k$ and $t_{k+1} := b$, where k is a natural number determined from $a + kh_N \le b < a + (k+1)h_N$. One clearly has $t_{-N} = a - \Delta$, $t_0 = a$, and $h_N \to 0$ as $N \to \infty$. Then the corresponding *neutral functional-difference inclusions* associated with (1.2) and (1.3) are given by

$$\text{(2.1)} \qquad \begin{cases} x_N(t_{j+1}) - Ax_N(t_{j+1}-\Delta) \in x_N(t_j) - Ax_N(t_j - \Delta) \\ +h_N F(x_N(t_j), x_N(t_j-\Delta), t_j) \quad \text{for} \quad j = 0,\ldots,k; \\ x_N(t_j) = c(t_j) \quad \text{for } j = -N,\ldots,-1. \end{cases}$$

A collection of vectors $\{x_N(t_j)|\, j=-N,\dots,k+1\}$ satisfying (2.1) is a *discrete trajectory* and the corresponding collection

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

is a *combined discrete velocity* for (2.1). We consider *extensions $x_N(t)$ of discrete trajectories* to the continuous-time interval $[a-\Delta,b]$ defined piecewise-linearly on $[a,b]$ and piecewise-constantly, continuously from the right on $[a-\Delta,a)$. We also define piecewise-constant *extensions of combined discrete velocities* on $[a,b]$ by

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

It is easy to verify that

$$x_N(t)-Ax_N(t-\Delta)=x_N(a)-Ax_N(a-\Delta)+\int_a^t v_N(s)\,ds\quad\text{for } t\in[a,b]\quad\text{and}$$
$$\frac{d}{dt}[x_N(t)-Ax_N(t-\Delta)]=v_N(t)\quad\text{a.e. } t\in[a,b].$$

Let $W^{1,2}[a,b]$ be a standard Sobolev space of absolutely continuous functions $x\colon[a,b]\to I\!R^n$ with the norm

$$\|x(\cdot)\|_{W^{1,2}}:=\max_{t\in[a,b]}|x(t)|+\Big(\int_a^b|\dot{x}(t)|^2dt\Big)^{1/2}.$$

The following theorem, which plays an essential role in the subsequent constructions and results of the paper being also important for its own sake, establishes a *strong approximation* of any admissible trajectory for the given neutral functional-differential inclusion by corresponding solutions to discrete approximations (2.1).

Theorem 2.1 *Let $\bar{x}(\cdot)$ be an admissible trajectory for* (1.2) *and* (1.3) *under hypotheses (H1)-(H3). Then there is a sequence $\{z_N(t_j)\mid j=-N,\dots,k+1\}$, $N\in I\!N$, of solutions to discrete inclusions* (2.1) *such that $z_N(t_0)=\bar{x}(a)$ for all $N\in I\!N$, the extended discrete trajectories $z_N(t)$, $a-\Delta\le t\le b$, converge uniformly to $\bar{x}(\cdot)$ on $[a-\Delta,b]$, and their extended combinations $z_N(t)-Az_N(t-\Delta)$ converge to $\bar{x}(t)-A\bar{x}(t-\Delta)$ in the $W^{1,2}$-norm on $[a,b]$ as $N\to\infty$. In particular, some subsequence of $\{\frac{d}{dt}[z_N(t)-Az_N(t-\Delta)]\}$ converges pointwisely to $\frac{d}{dt}[\bar{x}(t)-A\bar{x}(t-\Delta)]$ for a.e. $t\in[a,b]$.*

Proof. Using the density of step-functions in $L^1[a,b]$, we first select a sequence $\{\omega_N(\cdot)\}$, $N\in I\!N$, such that each $\omega_N(t)$ is constant on the interval $[t_j,t_{j+1})$ for $j=0,\dots,k$ and that $\omega_N(\cdot)$ converge to $\frac{d}{dt}[\bar{x}(\cdot)-A\bar{x}(\cdot-\Delta)]$ as $N\to\infty$ in the norm topology of $L^1[a,b]$. It follows from (H1) that

$$|\omega_N(t)|\le\Big|\omega_N(t)-\frac{d}{dt}[\bar{x}(t)-A\bar{x}(t-\Delta)]\Big|+\Big|\frac{d}{dt}[\bar{x}(t)-A\bar{x}(t-\Delta)]\Big|\le 1+m_F$$

for all $t \in [a,b]$ and $N \in I\!N$. In the estimates below we use the sequence

$$\xi_N := \int_a^b \left| \frac{d}{dt}[\bar{x}(t) - A\bar{x}(t-\Delta)] - \omega_N(t) \right| dt \to 0 \;\text{ as }\; N \to \infty.$$

Denote $\omega_{N_j} := \omega_N(t_j)$ and define discrete functions $\{u_N(t_j)\,|\,j = -N, \ldots, k+1\}$ recurrently by

$$\begin{cases} u_N(t_j) := \bar{x}(t_j) \;\text{ for }\; j = -N, \ldots, 0, \\ u_N(t_{j+1}) := Au_N(t_{j+1} - \Delta) + u_N(t_j) - Au_N(t_j - \Delta) + h_N \omega_{N_j} \quad \text{for } j = 0, \ldots, k. \end{cases}$$

Then the extended (in the above way) discrete functions satisfy

$$\begin{cases} u_N(t) = \bar{x}(t_j) \quad \text{for } t \in [t_j, t_{j+1}),\; j = -N, \ldots, -1, \\ u_N(t) - Au_N(t-\Delta) = \bar{x}(a) - A\bar{x}(a-\Delta) + \displaystyle\int_a^t \omega_N(s)\,ds \quad \text{for } t \in [a,b]. \end{cases}$$

Next we denote $r_N(t) := u_N(t) - \bar{x}(t)$, $y_N(t) := |r_N(t) - Ar_N(t-\Delta)|$ and prove that $|r_N(t)| \to 0$ uniformly in $[a,b]$ as $N \to \infty$. Indeed, for any $t \in [a,b]$ one has

$$\begin{aligned} y_N(t) &:= \left|u_N(t) - Au_N(t-\Delta) - [\bar{x}(t) - A\bar{x}(t-\Delta)]\right| \\ &\leq \int_a^t \left|\omega_N(s) - \frac{d}{dt}[\bar{x}(s) - A\bar{x}(s-\Delta)]\right| ds \leq \xi_N, \end{aligned}$$

which implies the estimates

$$\begin{aligned} |r_N(t)| &\leq y_N(t) + |A| \cdot |r_N(t-\Delta)| \leq y_N(t) + |A|y_N(t-\Delta) + |A|^2|r_N(t-2\Delta)| \leq \ldots \\ &\leq y_N(t) + |A|y_N(t-\Delta) + \ldots + |A|^m y_N(t-m\Delta) + |A|^{m+1}|r_N(t-(m+1)\Delta)|. \end{aligned}$$

Observe that $c(\cdot)$ is *uniformly continuous* on $[a-\Delta, a]$ due to assumption (H3). Picking an arbitrary sequence $\beta_N \downarrow 0$ as $N \to \infty$, we therefore have

$$|c(t') - c(t'')| \leq \beta_N \;\text{ whenever }\; t', t'' \in [t_j, t_{j+1}], \quad j = -N, \ldots, -1.$$

Choose an integer number m such that $a - \Delta \leq b - (m+1)\Delta < a$. Then $t - (m+1)\Delta \in [t_j, t_{j+1})$ for some $j \in \{-N, \ldots, -1\}$, which implies that

$$|r_N(t - (m+1)\Delta)| \leq |c(t_j) - c(t-(m+1)\Delta)| \leq \beta_N.$$

Since $m \in I\!N$ does not depend on N, this gives

$$|r_N(t)| \leq \xi_N(1 + |A| + \ldots + |A|^m) + |A|^{m+1}\beta_N \to 0 \quad \text{as } N \to \infty. \tag{2.2}$$

Now consider a sequence $\{\zeta_N\}$ defined by

$$\zeta_N := h_N \sum_{j=0}^{k} \operatorname{dist}(\omega_{N_j}; F(u_N(t_j), u_N(t_j - \Delta), t_j))$$

and show that $\zeta_N \downarrow 0$ as $N \to \infty$. By construction of ζ_N and the averaged modulus of continuity $\tau(F;h)$ we get the following estimates:

$$\begin{aligned}
\zeta_N &= \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} \operatorname{dist}(\omega_{N_j}; F(u_N(t_j), u_N(t_j-\Delta), t_j))\,dt \\
&= \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} \operatorname{dist}(\omega_{N_j}; F(u_N(t_j), u_N(t_j-\Delta), t))\,dt \\
&+ \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} \left[\operatorname{dist}(\omega_{N_j}; F(u_N(t_j), u_N(t_j-\Delta), t_j))\,dt - \operatorname{dist}(\omega_{N_j}; F(u_N(t_j), u_N(t_j-\Delta), t))\right] dt \\
&\le \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} \operatorname{dist}(\omega_{N_j}; F(u_N(t_j), u_N(t_j-\Delta), t))\,dt + \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} \sigma(F; t, h_N)\,dt \\
&\le \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} \operatorname{dist}(\omega_{N_j}; F(u_N(t_j), u_N(t_j-\Delta), t))\,dt + \tau(F; h_N).
\end{aligned}$$

Further, assumption (H1) implies that for any $t \in [t_j, t_{j+1})$ with $j = 0, \ldots, k$ one has

$$\begin{aligned}
&\operatorname{dist}(\omega_{N_j}; F(u_N(t_j), u_N(t_j-\Delta), t)) - \operatorname{dist}(\omega_{N_j}; F(u_N(t), u_N(t-\Delta), t)) \\
&\le \operatorname{dist}(F(u_N(t_j), u_N(t_j-\Delta), t), F(u_N(t), u_N(t-\Delta), t)) \\
&\le \ell_F(|u_N(t_j) - u_N(t)| + |u_N(t_j-\Delta) - u_N(t-\Delta)|).
\end{aligned}$$

Taking into account that

$$\begin{aligned}
&\left|u_N(t_j) - Au_N(t_j-\Delta) - [u_N(t) - Au_N(t-\Delta)]\right| \\
&= \left|\int_{t_j}^{t} \omega_N(s)\,ds\right| \le (1+m_F)(t_{j+1}-t_j) = (1+m_F)h_N := a_N \downarrow 0,
\end{aligned}$$

we arrive at

$$\begin{aligned}
|u_N(t) - u_N(t_j)| &\le a_N + |A| \cdot |u_N(t-\Delta) - u_N(t_j-\Delta)| \\
&\le a_N(1 + |A| + \ldots + |A|^m) + |A|^{m+1}|u_N(t-(m+1)\Delta) - u_N(t_j-(m+1)\Delta)| \\
&\le a_N(1 + |A| + \ldots + |A|^m) + |A|^{m+1}\beta_N := b_N \downarrow 0 \quad \text{as} \quad N \to \infty
\end{aligned}$$

and hence ensure that

$$\operatorname{dist}(\omega_{N_j}; F(u_N(t_j), u_N(t_j-\Delta), t)) - \operatorname{dist}(\omega_{N_j}; F(u_N(t), u_N(t-\Delta), t)) \le 2\ell_F b_N.$$

It follows from (H1) and (2.2) that for any $t \in [t_j, t_{j+1})$ and $j = 0, \ldots, k$ one has

$$\begin{aligned}
&\operatorname{dist}(\omega_{N_j}; F(u_N(t), u_N(t-\Delta), t)) - \operatorname{dist}(\omega_N(t); F(\bar{x}(t), \bar{x}(t-\Delta), t)) \\
&\le \operatorname{dist}(F(u_N(t), u_N(t-\Delta), t),\ F(\bar{x}(t), \bar{x}(t-\Delta), t)) \\
&\le \ell_F(|u_N(t) - \bar{x}(t)| + |u_N(t-\Delta) - \bar{x}(t-\Delta)|) \\
&\le 2\ell_F\xi_N(1 + |A| + \ldots + |A|^m) + 2\ell_F|A|^{m+1}\beta_N.
\end{aligned}$$

Combining the above estimates and denoting

$$\mu_N := 2b_N + 2\xi_N(1 + |A| + \ldots + |A|^m) + 2|A|^{m+1}\beta_N,$$

we arrive at

$$\begin{aligned}
&\mathrm{dist}(\omega_{N_j}; F(u_N(t_j), u_N(t_j - \Delta), t)) \le\ 2\ell_F b_N + \mathrm{dist}(\omega_{N_j}; F(u_N(t), u_N(t - \Delta), t)) \\
&\le 2\ell_F b_N + 2\ell_F \xi_N(1 + |A| + \ldots + |A|^m) + 2\ell_F |A|^{m+1}\beta_N + \mathrm{dist}(\omega_{N_j}; F(\bar{x}(t), \bar{x}(t - \Delta), t)) \\
&\le \ell_F \mu_N + \left|\omega_{N_j} - \frac{d}{dt}[\bar{x}(t) - A\bar{x}(t - \Delta)]\right|
\end{aligned}$$

and finally conclude that

$$\begin{aligned}
\zeta_N\ &\le \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} \left(\left|\omega_{N_j} - \frac{d}{dt}[\bar{x}(t) - A\bar{x}(t - \Delta)]\right| + \ell_F \mu_N\right) dt + \tau(F; h_N) \\
&= \xi_N + \ell_F \mu_N(b - a) + \tau(F; h_N) := \gamma_N \downarrow 0 \ \text{ as } \ N \to \infty.
\end{aligned} \tag{2.3}$$

Note that the discrete functions $\{u_N(t_j) \mid j = -N, \ldots, k+1\}$ may *not* be trajectories for (2.1), since one does not generally have $\omega_{N_i} \in F(u_N(t_j), u_N(t_j - \Delta), t_j)$ for $j = 0, \ldots, k$. Let us construct the desired trajectories $\{z_N(t_j) \mid j = -N, \ldots, k+1\}$ by the following *proximal algorithm*:

$$\begin{cases}
z_N(t_j) = c(t_j) \ \text{ for } \ j = -N, \ldots, -1, \quad z_N(t_0) = \bar{x}(a), \\
z_N(t_{j+1}) - Az_N(t_{j+1} - \Delta) = z_N(t_j) - Az_N(t_j - \Delta) + h_N v_{N_j} \ \text{ for } \ j = 0, \ldots, k, \\
v_{N_j} \in F(z_N(t_j), z_N(t_j - \Delta), t_j) \ \text{ with} \\
|v_{N_j} - \omega_{N_j}| = \mathrm{dist}(\omega_{N_j}; F(z_N(t_j), z_N(t_j - \Delta), t_j)) \ \text{ for } \ j = 0, \ldots, k.
\end{cases} \tag{2.4}$$

It follows from the construction (2.4) that $z_N(t_j)$ is a feasible trajectory to the discrete inclusion (2.1) for each $N \in I\!N$. Note that

$$|z_N(t) - \bar{x}(t)| = |z_N(t_j) - \bar{x}(t)| = |c(t_j) - c(t)| < \beta_N \quad \text{for } t \in [t_j, t_{j+1}),\ j = -N, \ldots, -1,$$

which implies that the extensions of $z_N(\cdot)$ converge to $\bar{x}(t)$ uniformly on $[a - \Delta, a)$. Let us analyze the situation on $[a, b]$.

First we claim that $z_N(t_j) \in U$ for $j = 0, \ldots, k+1$, where $U \subset I\!R^n$ is a neighborhood of $\bar{x}(\cdot)$ given in (H1). Arguing by induction, we obviously have $z_N(t_0) \in U$ and assume that $z_N(t_j) \in U$ for all $j = 1, \ldots, m$ with some fixed $m \in \{1, \ldots, k\}$. Then

$$\begin{aligned}
&|z_N(t_{m+1}) - u_N(t_{m+1})| = |Az_N(t_{m+1} - \Delta) + z_N(t_m) - Az_N(t_m - \Delta) + h_N v_{N_m} \\
&- (Au_N(t_{m+1} - \Delta) + u_N(t_m) - Au_N(t_m - \Delta) + h_N \omega_{N_m})| \\
&\le |A| \cdot |z_N(t_{m+1} - \Delta) - u_N(t_{m+1} - \Delta)| + |A| \cdot |z_N(t_m - \Delta) - u_N(t_m - \Delta)| \\
&+ |z_N(t_m) - u_N(t_m)| + h_N \mathrm{dist}(\omega_{N_m}; F(z_N(t_m), z_N(t_m - \Delta), t_m)).
\end{aligned}$$

Taking into account that

$$\begin{aligned}|z_N(t_m) - u_N(t_m)| &\le |z_N(t_{m-1}) - u_N(t_{m-1})| + |A| \cdot |z_N(t_{m-1-N}) - u_N(t_{m-1-N})| \\ &+ |A| \cdot |z_N(t_{m-N}) - u_N(t_{m-N})| + h_N \operatorname{dist}(\omega_{N_{m-1}}; F(z_N(t_{m-1}), z_N(t_{m-1} - \Delta))t_{m-1}))\end{aligned}$$

and that $|z_N(t_j) - u_N(t_j)| = 0$ for $j \le 0$, one has

$$|z_N(t_{m+1}) - u_N(t_{m+1})| \le Mh_N \sum_{j=0}^{m} \operatorname{dist}(\omega_{N_j}; F(u_N(t_j), u_N(t_j - \Delta), t_j)) \le M\gamma_N \tag{2.5}$$

with some constant $M > 0$. Now invoking (2.2) and increasing M if necessary, we arrive at

$$|z_N(t_{m+1}) - \bar{x}(t_{m+1})| \le \xi_N + M\gamma_N \to 0 \text{ as } N \to \infty.$$

It remains to prove that the extended combinations $z_N(t) - Az_N(t - \Delta)$ converge to $\bar{x}(t) - Ax(t - \Delta)$ in the $W^{1,2}$-norm on $[a, b]$, which means that

$$\begin{aligned}&\max_{t\in[a,b]} \left|z_N(t) - Az_N(t - \Delta) - [\bar{x}(t) - A\bar{x}(t - \Delta)]\right| \\ &+ \int_a^b \left|\frac{d}{dt}[z_N(t) - Az_N(t - \Delta)] - \frac{d}{dt}[\bar{x}(t) - A\bar{x}(t - \Delta)]\right|^2 dt \to 0 \quad \text{as } N \to \infty.\end{aligned} \tag{2.6}$$

To furnish this, we use (2.5) and get the estimate

$$\begin{aligned}\sum_{j=0}^{k+1} \left|z_N(t_j) - u_N(t_j)\right| &\le \sum_{j=0}^{k+1} M \sum_{m=0}^{j-1} h_N \operatorname{dist}(\omega_{N_m}; F(u_N(t_m), u_N(t_m - \Delta), t_m)) \\ &\le M(b - a) \sum_{j=0}^{k} \operatorname{dist}(\omega_{N_j}; F(u_N(t_j), u_N(t_j - \Delta), t_j)),\end{aligned}$$

which implies by (H1) that

$$\begin{aligned}&\int_a^b \left|\frac{d}{dt}[z_N(t) - Az_N(t - \Delta)] - \omega_N(t)\right| dt = \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} \left|\frac{d}{dt}[z_N(t) - Az_N(t - \Delta)] - \omega_N(t)\right| dt \\ &= \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} |v_{N_j} - \omega_{N_j}|\, dt = \sum_{j=0}^{k} h_N \operatorname{dist}(\omega_{N_j}; F(z_N(t_j), z_N(t_j - \Delta), t_j)) \\ &= \sum_{j=0}^{k} h_N \operatorname{dist}(\omega_{N_j}; F(u_N(t_j), u_N(t_j - \Delta), t_j)) \\ &+ \sum_{j=0}^{k} h_N [\operatorname{dist}(\omega_{N_j}; F(z_N(t_j), z_N(t_j - \Delta), t_j)) - \operatorname{dist}(\omega_{N_j}; F(u_N(t_j), u_N(t_j - \Delta), t_j))] \\ &\le \textstyle\sum_{j=0}^{k} h_N \operatorname{dist}(\omega_{N_j}; F(u_N(t_j), u_N(t_j - \Delta), t_j)) \\ &+ \sum_{j=0}^{k} \ell_F h_N \left[|z_N(t_j) - u_N(t_j)| + |z_N(t_j - \Delta) - u_N(t_j - \Delta)|\right] \\ &\le \gamma_N + \sum_{j=0}^{k} \ell_F h_N \left[|z_N(t_j) - u_N(t_j)| + |z_N(t_j - \Delta) - u_N(t_j - \Delta)|\right] \\ &\le \gamma_N + 2M(b - a)\ell_F \sum_{j=0}^{k} h_N \operatorname{dist}(\omega_{N_j}; F(u_N(t_j), u_N(t_j - \Delta), t_j)) \le \gamma_N + 2M\ell_F(b - a)\gamma_N.\end{aligned}$$

The latter ensures the estimate

$$\begin{aligned}
&\int_a^b \left|\frac{d}{dt}[z_N(t) - Az_N(t-\Delta)] - \frac{d}{dt}[\bar{x}(t) - A\bar{x}(t-\Delta)]\right| dt \\
&\le \int_a^b \left|\frac{d}{dt}[z_N(t) - Az_N(t-\Delta)] - \omega_N(t)\right| dt + \int_a^b \left|\omega_N(t) - \frac{d}{dt}[\bar{x}(t) - A\bar{x}(t-\Delta)]\right| dt \\
&\le \gamma_N(1 + 2M(b-a)\ell_F) + \xi_N.
\end{aligned}$$

Since $z_N(t) \in U$, it follows from (H1) by (1.2) and (2.4) that $\left|\frac{d}{dt}[z_N(t) - Az_N(t-\Delta)]\right| \le m_F$, $\left|\frac{d}{dt}[\bar{x}(t) - A\bar{x}(t-\Delta)]\right| \le m_F$, and hence

$$\begin{aligned}
&\int_a^b \left|\frac{d}{dt}[z_N(t) - Az_N(t-\Delta)] - \frac{d}{dt}[\bar{x}(t) - A\bar{x}(t-\Delta)]\right|^2 dt \\
&= \int_a^b \left|\frac{d}{dt}[z_N(t) - Az_N(t-\Delta)] - \frac{d}{dt}[\bar{x}(t) - A\bar{x}(t-\Delta)]\right| \\
&\qquad \times \left|\frac{d}{dt}[z_N(t) - Az_N(t-\Delta)] + \frac{d}{dt}[\bar{x}(t) - A\bar{x}(t-\Delta)]\right| dt \\
&\le \; 2m_F[\gamma_N(1 + 2M(b-a)\ell_F) + \xi_N] \downarrow 0 \;\text{ as }\; N \to \infty.
\end{aligned}$$

Observing that

$$\begin{aligned}
&\max_{t\in[a,b]} \left|z_N(t) - Az_N(t-\Delta) - [\bar{x}(t) - A\bar{x}(t-\Delta)]\right| \\
&\le \int_a^b \left|\frac{d}{dt}[z_N(t) - Az_N(t-\Delta)] - \frac{d}{dt}[\bar{x}(t) - A\bar{x}(t-\Delta)]\right|^2 dt,
\end{aligned}$$

we arrive at (2.6) and complete the proof of the theorem. $\triangle$

3 Strong convergence of discrete optimal solutions

Our next goal is to construct a sequence of *well-posed* discrete approximations of the whole dynamic optimization problem (P) given in (1.1)–(1.4) such that optimal solutions to discrete approximation problems *strongly converge*, in the sense described below, to a *given* optimal solution $\bar{x}(\cdot)$ to the original optimization problem governed by neutral functional-differential inclusions. The following construction explicitly involves the optimal solution $\bar{x}(\cdot)$ to the problem (P) under consideration for which we aim to derive necessary optimality conditions in the subsequent sections.

Given $\bar{x}(t)$, $a - \Delta \le t \le b$, take its approximation $z_N(t)$ from Theorem 2.1 and denote $\eta_N := |z_N(t_{k+1}) - \bar{x}(b)|$. For any natural number N we consider the following *discrete-time* dynamic optimization problem (P_N):

$$\begin{aligned}
&\text{minimize} \quad J_N[x_N] := \varphi(x_N(t_0), x_N(t_{k+1})) + |x_N(t_0) - \bar{x}(a)|^2 \\
&+h_N \sum_{j=0}^{k} f(x_N(t_j), x_N(t_j - \Delta), t_j) + \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} \left|\frac{d}{dt}[\bar{x}(t) - A\bar{x}(t-\Delta)]\right. \\
&\left. - \frac{x_N(t_{j+1}) - Ax_N(t_{j+1} - \Delta) - x_N(t_j) + Ax_N(t_j - \Delta)}{h_N}\right|^2 dt
\end{aligned} \tag{3.1}$$

subject to the *dynamic constraints* governed by neutral functional-difference inclusions (2.1), the *endpoint constraints*

$$(3.2) \qquad (x_N(t_0), x_N(t_{k+1})) \in \Omega_N := \Omega + \eta_N I\!B,$$

which are η_N*-perturbations* of the original endpoint constraints (1.4), and the auxiliary constraints

$$(3.3) \qquad |x_N(t_j) - \bar{x}(t_j)| \le \varepsilon, \quad j = 1, \ldots, k+1,$$

with some $\varepsilon > 0$. The latter auxiliary constraints are needed to guarantee the existence of optimal solutions in (P_N) and can be ignored in the derivation of necessary optimality conditions; see below.

In what follows we select $\varepsilon > 0$ in (3.3) such that $\bar{x}(t) + \varepsilon I\!B \subset U$ for all $t \in [a - \Delta, b]$ and take sufficiently large N ensuring that $\eta_N < \varepsilon$. Note that problems (P_N) have *feasible* solutions, since the trajectories z_N from Theorem 2.1 satisfy all the constraints (2.1), (3.2), and (3.3). Therefore, by the classical Weierstrass theorem in finite dimensions, each (P_N) admits an *optimal* solution $\bar{x}_N(\cdot)$ under the following assumption imposed in addition to (H1)–(H3), where U stands for a neighborhood of the optimal trajectory $\bar{x}(\cdot)$ to (P):

(H4) φ is continuous on $U \times U$, $f(x, y, \cdot)$ is continuous for a.e. $t \in [a, b]$ uniformly in $(x, y) \in U \times U$, $f(\cdot, \cdot, t)$ is continuous on $U \times U$ uniformly in $t \in [a, b]$, and Ω is locally closed around $(\bar{x}(a), \bar{x}(b))$.

We are going to justify the strong convergence of $\bar{x}_N(\cdot) \to \bar{x}(\cdot)$ in the sense of Theorem 2.1. To proceed, we need to involve an important intrinsic property of the original problem (P) called *relaxation stability.* Following the line originated by Jack Warga in optimal control theory (see the book [30] and its references), we consider the *relaxed* problem (R) of minimizing the cost functional (1.1) on admissible trajectories of the *convexified* functional-differential inclusion of the neutral type

$$(3.4) \qquad \frac{d}{dt}[x(t) - Ax(t - \Delta)] \in \operatorname{co} F(x(t), x(t - \Delta), t) \quad \text{a.e. } t \in [a, b]$$

with the initial "tail" condition (1.3) and the endpoint constraints (1.4). Any admissible trajectory for (3.4) satisfying (1.3) is called a *relaxed trajectory* for (1.2).

One clearly has $\inf(R) \le \inf(P)$ for the optimal values of the cost functionals in the relaxed and original problems. We say that the original problem (P) is *stable with respect to relaxation* if

$$\inf(P) = \inf(R).$$

This property, which obviously holds under the convexity assumption on the sets $F(x, y, t)$, goes far beyond the convexity. General sufficient conditions for the relaxation stability of the neutral-type problem (P) follows from [14]. We also refer the reader to [2, 21, 23, 30] for more detailed discussions on the validity of the relaxation stability property for various classes of differential, functional-differential, and functional-integral control systems.

Now we are ready to establish the following strong convergence theorem for optimal solutions to discrete approximations, which makes a bridge between optimal control problems governed by neutral functional-differential and functional-difference inclusions.

Theorem 3.1 *Let $\bar{x}(\cdot)$ be an optimal solution to problem (P), which is assumed to be stable with respect to relaxation. Suppose also that hypotheses (H1)–(H4) hold. Then any sequence $\{\bar{x}_N(\cdot)\}$, $N \in I\!N$, of optimal solutions to (P_N) extended to to the continuous interval $[a-\Delta, b]$ converges uniformly to $\bar{x}(\cdot)$ on $[a-\Delta, b]$, and the sequence of their combinations $\bar{x}_N(\cdot) - A\bar{x}_N(\cdot - \Delta)$ converges to $\bar{x}(\cdot) - A\bar{x}(\cdot - \Delta)$ in the $W^{1,2}$-norm on $[a,b]$ as $N \to \infty$.*

Proof. We know from the above discussion that (P_N) has an optimal solution $\bar{x}_N(\cdot)$ for all N sufficiently large; suppose that it happens for all $N \in I\!N$ without loss of generality. Given $\bar{x}(\cdot)$, we consider the sequence $\{z_N(\cdot)\}$ strongly approximating $\bar{x}(\cdot)$ by Theorem 2.1. Since each z_N is feasible to (P_N), one has

$$J_N[\bar{x}_N] \le J_N[z_N] \quad \text{for all} \quad N \in I\!N.$$

For convenience we represent $J_N[z_N]$ as the sum of three terms:

$$J_N[z_N] = \varphi(z_N(t_0), z_N(t_{k+1})) + h_N \sum_{j=0}^{k} f(z_N(t_j), z_N(t_j - \Delta), t_j)$$
$$+ \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} \left| \frac{z_N(t_{j+1}) - Az_N(t_{j+1} - \Delta) - z_N(t_j) + Az_N(t_j - \Delta)}{h_N} - \frac{d}{dt}[\bar{x}(t) - A\bar{x}(t-\Delta)] \right|^2 dt$$
$$:= I_{1N} + I_{2N} + I_{3N}.$$

It follows from Theorem 2.1 and the assumption on φ in (H4) that

$$I_{1N} \to \varphi(\bar{x}(a), \bar{x}(b)) \quad \text{as} \quad N \to \infty$$

and that, using the sign "$\approx$" for expressions that are equivalent as $N \to \infty$,

$$I_{2N} = h_N \sum_{j=0}^{k} f(z_N(t_j), z_N(t_j - \Delta), t_j) = \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} f(z_N(t_j), z_N(t_j - \Delta), t)\, dt$$
$$+ \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} [f(z_N(t_j), z_N(t_j - \Delta), t_j) - f(z_N(t_j), z_N(t_j - \Delta), t)]\, dt$$
$$= \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} f(z_N(t_j), z_N(t_j - \Delta), t)\, dt + \tau(f; h_N)$$
$$\approx \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} f(z_N(t_j), z_N(t_j - \Delta), t)\, dt \approx \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} f(\bar{x}(t_j), \bar{x}(t_j - \Delta), t)\, dt$$
$$\to \int_a^b f(\bar{x}(t), \bar{x}(t-\Delta), t)\, dt \quad \text{as} \quad N \to \infty,$$

$$I_{3N} = \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} \left| v_N(t) - \frac{d}{dt}[\bar{x}(t) - A\bar{x}(t-\Delta)] \right|^2 dt = \int_a^b \left| v_N(t) - \frac{d}{dt}[\bar{x}(t) - A\bar{x}(t-\Delta)] \right|^2 dt$$
$$= \int_a^b \left| \frac{d}{dt}[z_N(t) - Az_N(t-\Delta)] - \frac{d}{dt}[\bar{x}(t) - A\bar{x}(t-\Delta)] \right|^2 dt \to 0 \text{ as } N \to \infty,$$

where $v_N(t)$ stand for the extensions of combined discrete velocities for $z_N(\cdot)$; cf. Section 2. This implies that $J_N[z_N] \to J[\bar{x}]$ as $N \to \infty$, and therefore

$$\limsup_{N\to\infty} J_N[\bar{x}_N] \leq J[\bar{x}]. \tag{3.5}$$

It is easy to observe that the strong convergence claimed in the theorem follows from

$$\rho_N := \left|\bar{x}_N(a) - \bar{x}(a)\right|^2 + \int_a^b \left| \frac{d}{dt}[\bar{x}_N(t) - A\bar{x}_N(t-\Delta)] - \frac{d}{dt}[\bar{x}(t) - A\bar{x}(t-\Delta)] \right|^2 dt \to 0 \text{ as } N \to \infty.$$

On the contrary, suppose that the latter does not hold. Then there are a constant $\alpha > 0$ and a subsequence $\{N_m\} \subset I\!N$ for which $\rho_{N_m} \to \alpha$ as $m \to \infty$. Employing the standard compactness arguments based on (2.1) and the boundedness assumption in (H1), we find an absolutely continuous function $\widetilde{z}\colon [a,b] \to I\!R^n$ and a function $\widetilde{x}\colon [a-\Delta,b]$ continuous on $[a-\Delta,a)$ and $[a,b]$ such that

$$\frac{d}{dt}[\bar{x}_N(t) - A\bar{x}_N(t-\Delta)] \to \dot{\widetilde{z}}(t) \quad \text{weakly in } L^2[a,b],$$

that $\bar{x}_N(t) \to \widetilde{x}(t)$ uniformly on $[a-\Delta,b]$ as $N \to \infty$ (without loss of generality), and that $\widetilde{z}(t) = \widetilde{x}(t) - A\widetilde{x}(t-\Delta)$ for $t \in [a,b]$. By the classical Mazur theorem there is a sequence of *convex combinations* of $\frac{d}{dt}[\bar{x}_N(t) - A\bar{x}_N(t-\Delta)]$ that converges to $\frac{d}{dt}[\widetilde{x}(t) - A\widetilde{x}(t-\Delta)]$ in the norm topology of $L^2[a,b]$ and hence *pointwisely* for a.e. $t \in [a,b]$ along some subsequence. Therefore

$$\frac{d}{dt}[\widetilde{x}(t) - A\widetilde{x}(t-\Delta)] \in \text{co}\, F(\widetilde{x}(t), \widetilde{x}(t-\Delta), t) \;\text{ a.e. }\; t \in [a,b].$$

Since $\widetilde{x}(\cdot)$ obviously satisfies the tail condition (1.3) and the endpoint constraints (1.4), it is a feasible solution to the relaxed problem (R). Note that

$$h_N \sum_{j=0}^{k} f(\bar{x}_N(t_j), \bar{x}_N(t_j-\Delta), t_j) = \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} f(\bar{x}_N(t_j), \bar{x}_N(t_j-\Delta), t_j)dt \to \int_a^b f(\widetilde{x}(t), \widetilde{x}(t-\Delta), t)dt$$

as $N \to \infty$ due to the assumptions made. Observe also that the integral functional

$$I[v] := \int_a^b \left| v(t) - \frac{d}{dt}[\bar{x}(t) - A\bar{x}(t-\Delta)] \right|^2 dt$$

is lower semicontinuous in the weak topology of $L^2[a,b]$ by the *convexity* of the integrand in v. Since one has

$$\sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} \left| \frac{\bar{x}_N(t_{j+1}) - A\bar{x}_N(t_{j+1}-\Delta) - [\bar{x}_N(t_j) - A\bar{x}_N(t_j-\Delta)]}{h_N} - \frac{d}{dt}[\bar{x}(t) - A\bar{x}(t-\Delta)] \right|^2 dt$$
$$= \int_a^b \left| \frac{d}{dt}[\bar{x}_N(t) - A\bar{x}_N(t-\Delta)] - \frac{d}{dt}[\bar{x}(t) - A\bar{x}(t-\Delta)] \right|^2 dt,$$

the latter implies that

$$\int_a^b \left|\frac{d}{dt}[\tilde{x}(t) - A\tilde{x}(t-\Delta)] - \frac{d}{dt}[\bar{x}(t) - A\bar{x}(t-\Delta)]\right|^2 dt \le$$

$$\liminf_{N\to\infty} \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} \left|\frac{\bar{x}_N(t_{j+1}) - A\bar{x}_N(t_{j+1}-\Delta) - [\bar{x}_N(t_j) - A\bar{x}_N(t_j-\Delta)]}{h_N} - \frac{d}{dt}[\bar{x}(t) - A\bar{x}(t-\Delta)]\right|^2 dt.$$

Using the above relationships and passing to the limit in the expression (3.1) for $J_N[\bar{x}_N]$, we get

$$J[\tilde{x}] + \alpha \le \lim_{N\to\infty} J_N[\bar{x}_N].$$

By (3.5) one therefore has

$$J[\tilde{x}] \le J[\bar{x}] - \alpha < J[\bar{x}] \ \text{ if } \ \alpha > 0.$$

This clearly contradicts the optimality of $\bar{x}(\cdot)$ in the relaxed problem (R) due to the assumption on relaxation stability. Thus $\alpha = 0$, which completes the proof of the theorem. $\triangle$

4 Tools of variational analysis

The convegence/stability results of the previous section allow us to make a bridge between the original infinite-dimensional optimization problem (P) for neutral functional-differential inclusions and the sequence of finite-dimensional dynamic optimization problems (P_N) for neutral functional-difference inclusions. Our strategy is first to obtain necessary optimality conditions for the latter finite-dimensional problems and then derive necessary optimality conditions for the original problem (P) by passing to the limit from the ones for (P_N) as $N \to \infty$.

Observe that problems (P_N) are essentially *nonsmooth*, even in the case of smooth functions φ and f in the cost functional and the absence of endpoint constraints. The main source of nonsmoothness comes from the (increasing number of) geometric constraints in (2.1), which reflect the discrete dynamics and may have empty interiors. To conduct a variational analysis of such problems, we use appropriate tools of generalized differentiation introduced in [18] and then developed and applied in many publications; see, in particular, the books [19, 26] for detailed treatments and further references.

Recall the the *basic/limiting normal cone* to the set $\Omega \subset I\!R^n$ at the point $\bar{x} \in \Omega$ is

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

where $x \xrightarrow{\Omega} \bar{x}$ means that $x \to \bar{x}$ with $x \in \Omega$, and where

$$\widehat{N}(\bar{x};\Omega) := \left\{x^* \in I\!R^n \,\middle|\, \limsup_{x \xrightarrow{\Omega} \bar{x}} \frac{\langle x^*, x - \bar{x}\rangle}{|x - \bar{x}|} \le 0\right\}. \tag{4.2}$$

is the cone of Fréchet (or regular) normals to Ω at $\bar{x}$. For convex sets Ω both cones $N(\bar{x};\Omega)$ and $\widehat{N}(\bar{x};\Omega)$ reduce to the normal cone of convex analysis. Note that the basic normal cone (4.1) is often *nonconvex* while satisfying a comprehensive calculus, in contrast to (4.2).

Given an extended-real-valued function $\varphi\colon I\!R^n \to \overline{I\!R} := [-\infty,\infty]$ finite at $\bar{x}$, the *subdifferential* of φ at $\bar{x}$ is defined geometrically

$$\partial\varphi(\bar{x}) := \big\{x^* \in I\!R^n \big|\ (x^*,-1) \in N((\bar{x},\varphi(\bar{x}));\operatorname{epi}\varphi)\big\} \tag{4.3}$$

via basic normals to the epigraph $\operatorname{epi}\varphi := \{(x,\mu) \in I\!R^{n+1} |\ \mu \ge \varphi(x)\}$; equivalent analytic representations of (4.3) can be found in [19, 26].

Given a set-valued mapping $F\colon I\!R^n \rightrightarrows I\!R^m$ with the graph $\operatorname{gph} F := \{(x,y) \in I\!R^n \times I\!R^m |\ y \in F(x)\}$, the *coderivative* $D^*F(\bar{x},\bar{y})\colon I\!R^m \rightrightarrows I\!R^n$ of F at $(\bar{x},\bar{y}) \in \operatorname{gph} F$ is defined by

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

Note the useful relationships

$$\partial\varphi(\bar{x}) = D^*E_\varphi(\bar{x},\varphi(\bar{x}))(1) \quad \text{and} \quad D^*g(\bar{x})(y^*) = \partial\langle y^*,g\rangle(\bar{x}), \quad y^* \in I\!R^m,$$

between the subdifferential and coderivative introduced, where $E_\varphi(x) := \{\mu \in I\!R |\ \mu \ge \varphi(x)\}$ is the epigraphical multifunctions associated with $\varphi\colon I\!R^n \to \overline{I\!R}$ and where $\langle y^*,g\rangle(x) := \langle y^*,g(x)\rangle$ is the scalarized function associated with a locally Lipschitzian mapping $g\colon I\!R^n \to I\!R^m$.

The subdifferential/coderivative constructions (4.3) and (4.4) enjoy a variety of useful calculus rules that can be found in the books mentioned above and their references. Let us formulate two results crucial in what follows. The first one gives a complete coderivative characterization of the classical local Lipschitzian property of multifunctions imposed in our standing assumption (H1); cf. [20, Theorem 5.11] and [26, Theorem 9.40].

Theorem 4.1 *Let $F\colon I\!R^n \rightrightarrows I\!R^m$ be a closed-graph multifunction locally bounded around $\bar{x}$. Then the following conditions are equivalent:*

(i) *F is locally Lipschitzian around $\bar{x}$.*

(ii) *There exist a neighborhood U of $\bar{x}$ and a number $\ell > 0$ such that*

$$\sup\Big\{|x^*| \ \Big|\ x^* \in D^*F(x,y)(y^*)\Big\} \le \ell|y^*| \quad \textit{for all } x \in U,\ y \in F(x),\ y^* \in I\!R^m.$$

The next result taken from [19, Corollary 7.5] provides necessary optimality conditions for a general problem (MP) of *nonsmooth mathematical programming* with many geometric constraints:

$$\begin{cases} \text{minimize } \phi_0(z) \quad \text{subject to} \\ \varphi_j(z) \le 0, \quad j = 1,\ldots,r, \\ g_j(z) = 0, \quad j = 0,\ldots,m, \\ z \in \Lambda_j, \quad j = 0,\ldots,l, \end{cases}$$

where $\phi_j\colon I\!R^d \to I\!R$, $g_j\colon I\!R^d \to I\!R^n$, and $\Lambda_j \subset I\!R^d$.

Theorem 4.2 *Let $\bar z$ be an optimal solution to (MP). Assume that all ϕ_i are Lipschitz continuous, that g_j are continuously differentiable, and that Λ_j are locally closed near $\bar z$. Then there exist real numbers $\{\mu_j|\ j=0,\ldots,r\}$ as well as vectors $\{\psi_j \in I\!R^n|\ j=0,\ldots,m\}$ and $\{z_j^* \in I\!R^d|\ j=0,\ldots,l\}$, not all zero, such that*

$$\mu_j \ge 0 \quad \text{for} \quad j=0,\ldots,r, \tag{4.5}$$

$$\mu_j\phi_j(\bar z) = 0 \quad \text{for} \quad j=1,\ldots,r, \tag{4.6}$$

$$z_j^* \in N(\bar z;\Lambda_j) \quad \text{for} \quad j=0,\ldots,l, \tag{4.7}$$

$$-\sum_{j=0}^{l} z_j^* \in \partial\Big(\sum_{j=0}^{r}\mu_j\phi_j\Big)(\bar z) + \sum_{j=0}^{m}\nabla g_j(\bar z)^*\psi_j. \tag{4.8}$$

For applications in this paper in the case of nonautonomous continuous-time systems we need the following modifications of the basic constructions (4.1), (4.3), and (4.4) for sets, functions, and set-valued mappings *depending on a parameter t* from a topological space T (in our case $T=[a,b]$).

Given $\Omega\colon T \rightrightarrows I\!R^n$ and $\bar x \in \Omega(\bar t)$, we define the *extended normal cone* to $\Omega(\bar t)$ at $\bar x$ by

$$\widetilde N(\bar x;\Omega(\bar t)) := \mathop{\mathrm{Limsup}}_{(t,x)\overset{\mathrm{gph}\,\Omega}{\to}(\bar t,\bar x)} \widehat N(x;\Omega(t)). \tag{4.9}$$

For $\varphi\colon I\!R^n \times T \to \overline{I\!R}$ finite at $(\bar x,\bar t)$ and for $F\colon I\!R^n \times T \rightrightarrows I\!R^m$ with $\bar y \in F(\bar x,\bar t)$, the *extended subdifferential* of φ at $(\bar x,\bar t)$ and the *extended coderivative* of F at $(\bar x,\bar y,\bar t)$ with respect to x are given, respectively, by

$$\widetilde\partial_x\varphi(\bar x,\bar t) := \big\{x^* \in I\!R^n \big|\ (x^*,-1) \in \widetilde N((\bar x,\varphi(\bar x,\bar t));\mathrm{epi}\,\varphi(\cdot,\bar t))\big\} \tag{4.10}$$

and, whenever $y^* \in I\!R^m$, by

$$\widetilde D_x^*F(\bar x,\bar y,\bar t)(y^*) := \big\{x^* \in I\!R^n \big|\ (x^*,-y^*) \in \widetilde N((\bar x,\bar y);\mathrm{gph}\,F(\cdot,\bar t))\big\} \tag{4.11}$$

Note that the sets (4.9)–(4.11) may be bigger in some situations than the corresponding sets $N(\bar x;\Omega(\bar t))$, $\partial_x\varphi(\bar x,\bar t)$, and $D_x^*F(\bar x,\bar y,\bar t)(y^*)$, where the latter two sets stand for the subdifferential (4.3) of $\varphi(\cdot,\bar t)$ at $\bar x$ and the coderivative (4.4) of $F(\cdot,\bar t)$ at $(\bar x,\bar y,\bar t)$, respectively. Efficient conditions ensuring equalities for these sets are discussed in [21, 22, 24].

It is not hard to check that the extended constructions (4.9)–(4.11) are *robust* with respect to their variables, which is important for performing limiting procedures in what follows. In particular,

$$\widetilde N(\bar x;\Omega(\bar t)) = \mathop{\mathrm{Limsup}}_{(t,x)\overset{\mathrm{gph}\,\Omega}{\to}(\bar t,\bar x)} \widetilde N(x;\Omega(t)). \tag{4.12}$$

Note also that the constructions (4.9)–(4.11) enjoy a full generalized differential calculus similar to one for (4.1), (4.3), and (4.4). We are not going to use this calculus in the present paper.

5 Necessary optimality conditions for discrete approximations

This section concerns necessary optimality conditions for discrete approximation problems (P_N) governed by neutral functional-difference inclusions. We derive such conditions in the extended Euler-Lagrange form by reducing (P_N) to nonsmooth mathematical programs with many geometric constraints and employing generalized differential calculus for the basic constructions (4.1), (4.3), and (4.4).

Let us reduce the dynamic optimization problem (P_N) for each $N \in I\!N$ to the mathematical programming problem (MP) considered in Section 4 with the decision vector

$$z := (x_0^N, x_1^N, \ldots, x_{k+1}^N, v_0^N, v_1^N, \ldots, v_k^N) \in I\!R^{n(2k+3)}$$

and the following data:

$$\begin{aligned}\phi_0(z): \quad &= \varphi(x_0^N, x_{k+1}^N) + |x_0^N - \bar{x}(a)|^2 + h_N \sum_{j=0}^{k} f(x_j^N, x_{j-N}^N, t_j)\\ &+ \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} \left| v_j^N - \frac{d}{dt}[\bar{x}(t) - A\bar{x}(t-\Delta)] \right|^2 dt,\end{aligned} \tag{5.1}$$

$$\begin{aligned}\phi_j(z): \quad &= |x_j^N - \bar{x}(t_j)| - \varepsilon, \quad j = 1, \ldots, k+1,\\ \Lambda_j: \quad &= \left\{(x_0^N, \ldots, v_k^N) \mid v_j^N \in F(x_j^N, x_{j-N}^N, t_j)\right\}, \quad j = 0, \ldots, k,\\ \Lambda_{k+1}: \quad &= \left\{(x_0^N, \ldots, v_k^N) \mid (x_0^N, x_{k+1}^N) \in \Omega_N\right\},\\ g_j(z): \quad &= x_{j+1}^N - Ax_{j+1-N}^N - x_j^N + Ax_{j-N}^N - h_N v_j^N, \quad j = 0, \ldots, k,\end{aligned}$$

where $x_j^N := c(t_j)$ for $j < 0$. Let $\bar{z}^N = (\bar{x}_0^N, \ldots, \bar{x}_{k+1}^N, \bar{v}_0^N, \ldots, \bar{v}_k^N)$ be an optimal solution to (MP). Applying Theorem 4.2, we find real numbers μ_j^N and vectors $z_j^* \in I\!R^{n(2k+3)}$ for $j = 0, \ldots, k+1$ as well as vectors $\psi_j^N \in I\!R^n$ for $j = 0, \ldots, k$, not all zero, such that conditions (4.5)–(4.8) are satisfied.

Taking $z_j^* = (x_{0,j}^*, \ldots, x_{k+1,j}^*, v_{0,j}^*, \ldots, v_{k,j}^*) \in N(\bar{z}^N; \Lambda_j)$ for $j = 0, \ldots, k$, we observe that all but one components of z_j^* are zero and the remaining one satisfies

$$(x_{j,j}^*, x_{j-N,j}^*, v_{j,j}^*) \in N((\bar{x}_j^N, \bar{x}_{j-N}^N, \bar{v}_j^N); \operatorname{gph} F(\cdot, \cdot, t_j)), \quad j = 0, \ldots, k.$$

Similarly, the condition $z_{k+1}^* \in N(\bar{z}^N; \Lambda_{k+1})$ is equivalent to

$$(x_{0,k+1}^*, x_{k+1,k+1}^*) \in N((\bar{x}_0^N, \bar{x}_{k+1}^N); \Omega_N)$$

with all the other components of z_{k+1}^* equal to zero. Employing Theorem 3.1 on the convergence of discrete approximations, we have $\phi_j(\bar{z}^N) < 0$ for $j = 1, \ldots, k+1$ whenever N is sufficiently large. Thus $\mu_j^N = 0$ for these indexes due to the complementary slackness conditions (4.6). Let

$\lambda^N := \mu_0^N \geq 0$. Observe further that

$$\sum_{j=0}^{k}(\nabla g_j(\bar z^N))^*\psi_j^N = \Big(-\psi_0 + A^*(\psi_N^N - \psi_{N-1}^N), \psi_0 - \psi_1 + A^*(\psi_{N+1}^N - \psi_N^N), \ldots,$$
$$\psi_{k-N-1} - \psi_{k-N} + A^*(\psi_k^N - \psi_{k-1}^N), \psi_{k-N} - \psi_{k-N+1} + A^*\psi_k^N, \ldots,$$
$$\psi_{k-1}^N - \psi_k^N, \psi_k^N, -h_N\psi_0^N, \ldots, -h_N\psi_k^N\Big).$$

From the subdifferential sum rule for ϕ_0 in (5.1) one has

$$\partial\phi(\bar z^N) \subset \partial\varphi(\bar x_0^N, \bar x_{k+1}^N) + 2(\bar x_0^N - \bar x(a)) + h_N \sum_{j=0}^{k} \partial f(\bar x_j^N, \bar x_{j-N}^N, t_j)$$
$$+ \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} 2\Big(\bar v_j^N - \frac{d}{dt}[\bar x(t) - A\bar x(t-\Delta)]\Big)\, dt$$

with ∂f standing here and in what follows for the basic subdifferential of f with respect to the first two variables. Thus the inclusion (4.8) in Theorem 4.2 is equivalent to the relationships:

$$-x_{0,0}^* - x_{0,N}^* - x_{0,k+1}^* = \lambda^N u_0^N + \lambda^N h_N \vartheta_0^N + \lambda^N h_N \kappa_0^N + 2\lambda^N(\bar x_0^N - \bar x(a)) - \psi_0^N - A^*(\psi_{N-1}^N - \psi_N^N),$$
$$-x_{j,j}^* - x_{j,j+N}^* = \lambda^N h_N \kappa_j^N + \lambda^N h_N \vartheta_j^N + \psi_{j-1}^N - \psi_j^N - A^*(\psi_{j+N-1}^N - \psi_{j+N}^N), \quad j = 1, \ldots, k-N,$$
$$-x_{k-N+1,k-N+1}^* = \lambda^N h_N v_{k-N+1}^N + \psi_{k-N}^N - \psi_{k-N+1}^N + A^*\psi_k^N,$$
$$-x_{j,j}^* = \lambda^N h_N \vartheta_j^N + \psi_{j-1}^N - \psi_j^N, \quad j = k-N+2, \ldots, k,$$
$$-x_{k+1,k+1}^* = \lambda^N u_{k+1}^N + \psi_k^N,$$
$$-v_{j,j}^* = \lambda^N \theta_j^N - h_N \psi_j^N, \quad j = 0, \ldots, k$$

with the notation

$$(u_0^N, u_{k+1}^N) \in \partial\varphi(\bar x_0^N, \bar x_{k+1}^N), \quad (\vartheta_j^N, \kappa_{j-N}^N) \in \partial f(\bar x_j^N, \bar x_{j-N}^N, t_j),$$
$$\theta_j^N := -2\int_{t_j}^{t_{j+1}} \Big(\frac{d}{dt}[\bar x(t) - A\bar x(t-\Delta)] - \bar v_j^N\Big)\, dt.$$

Based on the above relationships, we arrive at the following necessary optimality conditions for discrete-time problems (P_N), where $f_j(\cdot,\cdot) := f(\cdot,\cdot,t_j)$ and $F_j(\cdot,\cdot) := F(\cdot,\cdot,t_j)$.

Theorem 5.1 *Let $\bar z^N$ be an optimal solution to problem (P_N). Assume that the sets Ω and* gph F_j *are closed and that the functions φ and f_j are Lipschitz continuous around the points $(\bar x_0^N, \bar x_{k+1}^N)$ and $(\bar x_j^N, \bar x_{j-N}^N)$, respectively, for all $j = 0, \ldots, k$. Then there exist $\lambda^N \geq 0$, p_j^N $(j = 0, \ldots, k+N+1)$, and q_j^N $(j = -N, \ldots, k+1)$, not all zero, such that*

$$p_j^N = 0, \quad j = k+2, \ldots, k+N+1, \tag{5.2}$$

$$q_j^N = 0, \quad j = k-N+1, \ldots, k+1, \tag{5.3}$$

(5.4) $$(p_0^N+q_0^N,-p_{k+1}^N)\in\lambda^N\partial\varphi(\bar{x}_0^N,\bar{x}_{k+1}^N)+N((\bar{x}_0^N,\bar{x}_{k+1}^N);\Omega_N),$$

(5.5) $$\Big(\frac{P_{j+1}^N-P_j^N}{h_N},\frac{Q_{j-N+1}^N-Q_{j-N}^N}{h_N},-\frac{\lambda^N\theta_j^N}{h_N}+p_{j+1}^N+q_{j+1}^N\Big)$$
$$\in\lambda^N\big(\partial f_j(\bar{x}_j^N,\bar{x}_{j-N}^N),0\big)+N((\bar{x}_j^N,\bar{x}_{j-N}^N,\bar{v}_j^N);\operatorname{gph}F_j),\quad j=0,\ldots,k,$$

with the notation

$$P_j^N:=p_j^N-A^*p_{j+N}^N,\quad Q_j^N:=q_j^N-A^*q_{j+N}^N,$$

$$\bar{v}_j^N:=\frac{(\bar{x}_{j+1}^N-\bar{x}_j^N)+A(x_{j-N}^N-x_{j-N+1}^N)}{h_N}.$$

Proof. Most of the proof has been actually done above, and we just need to change notation in the relationships formulated right before the theorem. Let first

$$\widetilde{p}_j^N:=\begin{cases}\psi_{j-1}^N & \text{for } j=1,\ldots,k+1,\\ 0 & \text{for } j=k+2,\ldots,k+N+1.\end{cases}$$

$$\widetilde{q}_j^N:=\begin{cases}x_{j,j+N}^*/h_N & \text{for } j=-N,\ldots,-1,\\ \lambda^N\kappa_j^N+x_{j,j+N}^*/h_N & \text{for } j=0,\ldots,k-N,\\ 0 & \text{for } j=k-N+1,\ldots,k+1,\end{cases}$$

and then define q_j^N for $j=-N,\ldots,k+1$ by the recurrent formula

$$q_j^N:=q_{j+1}^N-A^*(q_{j+N+1}^N+q_{j+N}^N)-h_N\widetilde{q}_j^N,$$

where we put $q_j^N:=0$ for $j>k+1$. Observe that

$$\Big(\frac{(\widetilde{p}_{j+1}^N-q_{j+1}^N)-A^*(\widetilde{p}_{j+N+1}^N-q_{j+N+1}^N)-(\widetilde{p}_j^N-q_j^N)+A^*(\widetilde{p}_{j+N}^N-q_{j+N}^N)}{h_N},$$
$$\frac{(q_{j-N+1}^N-A^*q_{j+1}^N)-(q_{j-N}^N-A^*q_j^N)}{h_N},-\frac{\lambda^N\theta_j^N}{h_N}+\widetilde{p}_{j+1}^N\Big)$$
$$\in\lambda^N\big(\partial f_j(\bar{x}_j^N,\bar{x}_{j-N}^N),0\big)+N((\bar{x}_j^N,\bar{x}_{j-N}^N,\bar{v}_j^N);\operatorname{gph}F_j),\quad j=0,\ldots,k.$$

Letting finally

$$p_0^N:=\lambda^Nu_0^N+x_{0,k+1}^*-q_0^N,\quad p_j^N:=\widetilde{p}_j^N-q_j^N\quad\text{for } j=1,\ldots,k+N+1,$$

we can easily check that all the relationships (5.2)–(5.5) hold. △

Corollary 5.2 *In addition to the assumptions of Theorem 5.1, suppose that the mapping F_j is bounded and Lipschitz continuous around $(\bar{x}_j^N,\bar{x}_{j-N}^N)$ for each $j=0,\ldots,k$. Then conditions $\lambda^N\geq 0$ and (5.2)–(5.5) hold with $(\lambda^N,p_{k+1}^N)\neq 0$, i.e., one can let*

(5.6) $$(\lambda^N)^2+|p_{k+1}^N|^2=1.$$

Proof. If $\lambda^N = 0$, then (5.5) together with (5.2) and (5.3) imply that

$$\Big(\frac{p_{k+1}^N - p_k^N}{h_N}, \frac{-q_{k-N}^N}{h_N}\Big) \in D^*F_k(\bar{x}_k^N, \bar{x}_{k-N}^N, \bar{v}_k^N)(-p_{k+1}^N).$$

Assuming now that $p_{k+1}^N = 0$, we get

$$\Big(\frac{-p_k^N}{h_N}, \frac{-q_{k-N}^N}{h_N}\Big) \in D^*F_k(\bar{x}_k^N, \bar{x}_{k-N}^N, \bar{v}_k^N)(0),$$

which yield $p_k^N = q_{k-N}^N = 0$ by Theorem 4.1. Repeating the above procedure, we arrive at contradiction with the nontriviality assertion in Theorem 5.1. $\triangle$

6 Optimality conditions for functional-differential inclusions

In this section we obtain the main results of the paper providing necessary optimality conditions for the original dynamic optimization problem (P) in both extended Euler-Lagrange and Hamiltonian forms involving generalized differential constructions of Section 4. Our major theorem establishes the following conditions of the Euler-Lagrange type derived by the limiting procedure from discrete approximations. Note that here $\Delta > 0$ as was assumed in Section 2.

Theorem 6.1 *Let $\bar{x}(\cdot)$ be an optimal solution to problem (P) under hypotheses (H1)-(H4), where φ and $f(\cdot,\cdot,t)$ are assumed to be Lipschitz continuous instead of the plain continuity. Suppose also that (P) is stable with respect to relaxation. Then there exist a number $\lambda \geq 0$ and piecewise continuous functions $p\colon [a, b+\Delta] \to I\!R^n$ and $q\colon [a-\Delta, b] \to I\!R^n$ such that $p(t) - A^*p(t+\Delta)$ and $q(t-\Delta) - A^*q(t)$ are absolutely continuous on $[a,b]$ and the following conditions hold:*

$$\lambda + |p(b)| = 1, \tag{6.1}$$

$$p(t) = 0 \ \ for \ \ t \in (b, b+\Delta], \quad q(t) = 0 \ \ for \ \ t \in (b-\Delta, b], \tag{6.2}$$

$$(p(a) + q(a), -p(b)) \in \lambda\partial\varphi(\bar{x}(a), \bar{x}(b)) + N((\bar{x}(a), \bar{x}(b)); \Omega), \tag{6.3}$$

$$\begin{aligned} &\Big(\frac{d}{dt}[p(t) - A^*p(t+\Delta)], \frac{d}{dt}[q(t-\Delta) - A^*q(t)]\Big) \\ &\in \text{co}\Big\{(u, w, p(t) + q(t)) \in \lambda(\widetilde{\partial} f(\bar{x}(t), \bar{x}(t-\Delta), t), 0) \\ &+\widetilde{N}\big((\bar{x}(t), \bar{x}(t-\Delta), \frac{d}{dt}[\bar{x}(t) - A\bar{x}(t-\Delta)]); \text{gph}\, F(\cdot,\cdot,t)\big)\Big\} \quad a.e. \ \ t \in [a,b]. \end{aligned} \tag{6.4}$$

Proof. To prove this theorem by the method of discrete approximations, we first construct a sequence of discrete-time problems (P_N) whose optimal solutions $\bar{x}_N$ strongly approximate $\bar{x}(\cdot)$ in the sense of Theorem 2.1. By necessary optimality conditions for $\bar{x}_N$ from Corollary 5.2 we find $\lambda_N \geq 0$, p_j^N, and q_j^N satisfying relationships (5.2)–(5.6) for all $N \in I\!N$.

Without loss of generality we suppose that $\lambda^N \to \lambda$ as $N \to \infty$ for some $\lambda \geq 0$. As usual, the symbols $\bar{x}^N(t)$, $p^N(t)$, $q^N(t-\Delta)$, $P^N(t)$, and $Q^N(t)$ stand for the piecewise linear extensions of the corresponding discrete functions from Theorem 5.1 with their piecewise constant derivatives on the continuous-time interval $[a,b]$.

Considering θ_j from Theorem 5.1, we define $\theta^N(t) := \theta_j^N/h_N$ for $t \in [t_j, t_{j+1})$ as $j = 0,\ldots,k$ and conclude by Theorem 2.1 that

$$\int_a^b |\theta^N(t)|\,dt = \sum_{j=0}^k |\theta_j^N| \leq 2\sum_{j=0}^k \int_{t_j}^{t_{j+1}} \left|\frac{d}{dt}[\bar{x}(t) - A\bar{x}(t-\Delta)] - \bar{v}_j^N\right| dt$$

$$= 2\int_a^b \left|\frac{d}{dt}[\bar{x}(t) - A\bar{x}(t-\Delta)] - \frac{d}{dt}[\bar{x}^N(t) - A\bar{x}^N(t-\Delta)]\right| dt := \nu_N \to 0 \ \text{ as } \ N \to \infty.$$

We may assume without less of generality that

$$\bar{v}^N(t) := \frac{d}{dt}[\bar{x}^N(t) - A\bar{x}^N(t-\Delta)] \to \frac{d}{dt}[\bar{x}(t) - A\bar{x}(t-\Delta)] \quad \text{and} \quad \theta^N(t) \to 0 \ \text{ a.e. } \ t \in [a,b]$$

as $N \to \infty$; such a *pointwise* convergence plays a significant role in what follows.

Let us estimate $(p^N(t), q^N(t-\Delta))$ for large N. Using (5.2) and (5.3), we derive from (5.5) that

$$\Big(\frac{p_{j+1}^N - p_j^N}{h_N} - \lambda^N\vartheta_j^N,\ \frac{q_{j-N+1}^N - q_{j-N}^N}{h_N} - \lambda^N\kappa_{j-N}^N,\ -\frac{\lambda^N\theta_j^N}{h_N} + p_{j+1}^N\Big)$$
$$\in N((\bar{x}_j^N, \bar{x}_{j-N}^N, \bar{v}_j^N); \operatorname{gph} F_j) \ \text{ with some } \ (\vartheta_j^N, \kappa_{j-N}^N) \in \partial f_j(\bar{x}_j^N, \bar{x}_{j-N}^N)$$

for $j = k-N+2,\ldots,k+N+1$. This means, by definition of the coderivative (4.4), that

$$\Big(\frac{p_{j+1}^N - p_j^N}{h_N} - \lambda^N\vartheta_j^N,\ \frac{q_{j-N+1}^N - q_{j-N}^N}{h_N} - \lambda^N\kappa_{j-N}^N\Big) \in D^*F_j(\bar{x}_j^N, \bar{x}_{j-N}^N, \bar{v}_j^N)\Big(\frac{\lambda^N\theta_j^N}{h_N} - p_{j+1}^N\Big)$$

for such j. Thus it follows from Theorem 4.1 that

$$\Big|\Big(\frac{p_{j+1}^N - p_j^N}{h_N} - \lambda^N\vartheta_j^N,\ \frac{q_{j-N+1}^N - q_{j-N}^N}{h_N} - \lambda^N\kappa_{j-N}^N\Big)\Big| \leq \ell_F\Big|\frac{\lambda^N\theta_j^N}{h_N} - p_{j+1}^N\Big|$$

for $j = k-N+2,\ldots,k+N+1$. Since $|(\vartheta_j^N, \kappa_{j-N}^N)| \leq \ell_f$ due to the Lipschitz continuity of f with modulus ℓ_f, we derive from the above that

$$|(p_j^N, q_{j-N}^N)| \leq \ell_F|\theta_j^N| + (\ell_F+1)h_N\ell_f + (\ell_F h_N + 1)|(p_{j+1}^N, q_{j-N+1}^N)|$$
$$\leq \ell_F|\theta_j^N| + (\ell_F h_N + 1)\ell_F|\theta_{j-1}^N| + (\ell_F+1)h_N\ell_f + (\ell_F h_N+1)(\ell_F+1)h_N\ell_f$$
$$+ (\ell_F h_N + 1)^2|(p_{j+2}^N, q_{j-N+2}^N)| \leq \ldots$$
$$\leq \exp[\ell_F(b-a)](1 + \ell_f(\ell_F+1)/\ell_F + \ell_F\nu_N), \quad j = k-N+2,\ldots,k+N+1,$$

which implies the uniform boundedness of $\{(p_j^N, q_{j-N}^N)\,|\, j = k-N+2,\ldots,k+N+1\}$ and hence of $(p^N(t), q^N(t-\Delta))$ on $[b-\Delta, b]$.

Next we consider $j = k - 2N + 2, \dots, k + 1$ and derive from (5.5) that

$$\Big|\Big(\frac{p_{j+1}^N - p_j^N}{h_N} - \lambda^N \vartheta_j^N, \frac{q_{j-N+1}^N - q_{j-N}^N}{h_N} - \lambda^N \kappa_{j-N}^N\Big)\Big|$$
$$\le \ell_F \Big|\frac{\lambda^N \theta_j^N}{h_N} - p_{j+1}^N - q_{j+1}^N\Big| + \Big|\Big(\frac{A^* p_{j+N+1}^N - A^* p_{j+N}^N}{h_N}, \frac{A^* q_{j+1}^N - A^* q_j^N}{h_N}\Big)\Big|.$$

This implies due to Theorem 4.1 and the uniform boundedness of p_{j+N}^N and q_j^N by some constant $\alpha > 0$ for such j that

$$\Big|\Big(\frac{p_{j+1}^N - p_j^N}{h_N} - \lambda^N \vartheta_j^N, \frac{q_{j-N+1}^N - q_{j-N}^N}{h_N} - \lambda^N \kappa_{j-N}^N\Big)\Big| \le \ell_F \Big|\frac{\lambda^N \theta_j^N}{h_N} - p_{j+1}^N - q_{j+1}^N\Big| + \frac{\alpha}{h_N}$$

for $j = k - 2N + 2, \dots, k + 1$. Therefore

$$|(p_j^N, q_{j-N}^N)| \le \ell_F |\theta_j^N| + (\ell_F + 1) h_N \ell_f + (\ell_F h_N + 1) |(p_{j+1}^N, q_{j-N+1}^N)| + (\ell_F h_N + 1)\alpha$$
$$\le \ell_F |\theta_j^N| + (\ell_F h_N + 1)\ell_F |\theta_{j+1}^N| + (\ell_F + 1) h_N \ell_f + (\ell_F h_N + 1)(\ell_F + 1) h_N \ell_f$$
$$+ (\ell_F h_N + 1)(\ell_F + 1)\alpha + (\ell_F h_N + 1)^2 |(p_{j+2}^N, q_{j-N+2}^N)| \le \dots$$
$$\le \exp[\ell_F(b-a)](1 + (\ell_f + \alpha)(\ell_F + 1)/\ell_F + \ell_F \nu_N), \quad j = k - 2N + 2, \dots, k + 1.$$

This shows that p_j^N and q_{j-N}^N are uniformly bounded for $j = k - 2N + 2, \dots, k + 1$, and hence the sequence$\{p^N(t), q^N(t-\Delta)\}$ is uniformly bounded on $[b-2\Delta, b-\Delta]$. Repeating the above procedure, we conclude that both sequences $\{p^N(t), q^N(t-\Delta)\}$ and $\{P^N(t), Q^N(t-\Delta)\}$ are uniformly bounded on the whole interval $[a, b]$.

Next we estimate $(\dot{P}^N(t), \dot{Q}^N(t-\Delta))$ on $[a, b]$ using (5.5) and Theorem 4.1. This yields, for $t_j \le t < t_{j+1}$ with $j = 0, \dots, k$, that

$$|(\dot{P}^N(t), \dot{Q}^N(t-\Delta))| = \Big|\Big(\frac{P_{j+1}^N - P_j^N}{h_N}, \frac{Q_{j-N+1}^N - Q_{j-N}^N}{h_N}\Big)\Big|$$
$$\le \ell_F \Big|\frac{\lambda^N \theta_j^N}{h_N} - p_{j+1}^N - q_{j+1}^N\Big| + \ell_f \le \ell_F |\theta_j^N| + \ell_F |p_{j+1}^N| + \ell_F |q_{j+1}^N| + \ell_f.$$

Thus the sequence $\{\dot{P}^N(t), \dot{Q}^N(t-\Delta)\}$ is weakly compact in $L^1[a, b]$. Taking the whole sequence of $N \in \mathbb{N}$ without loss of generality, we find two absolutely continuous functions $P(\cdot)$ and $Q(\cdot - \Delta)$ on $[a, b]$ such that

$$\dot{P}^N(t) \to \dot{P}(t), \quad \dot{Q}^N(t-\Delta) \to \dot{Q}(t-\Delta) \text{ weakly in } L^1[a, b]$$

and $P^N(t) \to P(t)$, $Q^N(t-\Delta) \to Q(t-\Delta)$ uniformly on $[a, b]$ as $N \to \infty$. Since $p^N(t)$ and $q^N(t-\Delta)$ are uniformly bounded on $[a, b+\Delta]$, they surely converge to some functions $p(t)$ and $q(t-\Delta)$ weakly in $L^1[a, b+\Delta]$. Taking into account the above convergence of $P^N(t)$ and $Q^N(t-\Delta)$, we get that $p(\cdot)$ and $q(\cdot)$ satisfy (6.2), that

$$P(t) = p(t) - A^* p(t+\Delta), \quad Q(t-\Delta) = q(t-\Delta) - A^* q(t), \quad t \in [a, b],$$

and that $p(t)$ and $q(t)$ are piecewise continuous on $[a, b+\Delta]$ and $[a-\Delta, b]$, respectively, with possible discontinuity (from the right) at the points $b - i\Delta$ at $i = 0, 1, \ldots$. Conditions (6.1) and (6.3) follow by passing to the limit from (5.6) and (5.4), respectively, taking into account the robustness of the basic subdifferential (4.3) and the normal cone (4.1).

It remains to justify the Euler-Lagrange inclusion (6.4). To furnish this, we rewrite the discrete Euler-Lagrange inclusion (5.5) in the form

$$
\begin{aligned}
(6.5) \qquad & (\dot{P}^N(t), \dot{Q}^N(t-\Delta)) \in \Big\{(u,w)\Big| \Big(u, w, p^N(t_{j+1}) + q^N(t_{j+1}) - \frac{\lambda^N \theta_j^N}{h_N}\Big) \\
& \in \lambda^N\big(\partial f(\bar{x}(t_j), \bar{x}(t_j - \Delta), t_j), 0\big) + \big(N(\bar{x}_j^N, \bar{x}_{j-N}^N, \bar{v}_j^N); \operatorname{gph} F_j\big)\Big\}
\end{aligned}
$$

for $t \in [t_j, t_{j+1}]$ with $j = 0, \ldots, k$. By the classical Mazur theorem there is a sequence of convex combinations of the functions $(\dot{P}^N(t), \dot{Q}^N(t-\Delta))$ that converges to $(\dot{P}(t), \dot{Q}(t-\Delta))$ for a.e. $t \in [a, b]$. Passing the limit in (6.5) with taking into account the pointwise convergence of $\theta^N(t)$ and $\bar{v}^N(t)$ established above, as well as the constructions of the extended normal cone (4.9) and the extended subdifferential (4.10) and their robustness property (4.12) with respect to all variables and parameters, we arrive at (6.4) and complete the proof of the theorem. $\triangle$

Observe that for the *Mayer problem* (P_M), which is (1.1)–(1.4) with $f = 0$, the generalized Euler-Lagrange inclusions (6.4) is equivalently expressed in terms of the extended coderivative (4.11) with respect to the first two variables of $F = F(x, y, t)$, i.e., in the form

$$
\begin{aligned}
(6.6) \qquad & \Big(\frac{d}{dt}[p(t) - A^* p(t+\Delta)], \frac{d}{dt}[q(t-\Delta) - A^* q(t)]\Big) \\
& \in \operatorname{co} \widetilde{D}^*_{x,y} F\big(\bar{x}(t), \bar{x}(t-\Delta), \frac{d}{dt}[\bar{x}(t) - A\bar{x}(t-\Delta)], t\big)\big(-p(t) - q(t)\big) \quad \text{a.e. } t \in [a, b].
\end{aligned}
$$

It turns out that the extended Euler-Lagrange inclusion obtained above implies, under the *relaxation stability* of the original problems, two other principal optimality conditions expressed in terms of the Hamiltonian function built upon the mapping F in (1.2). The first condition called the extended *Hamiltonian inclusion* is given below in terms of a *partial convexification* of the basic subdifferential (4.3) for the Hamiltonian function. The second one is an analog of the classical *Weierstrass-Pontryagin maximum condition* (maximum principle) for neutral functional-differential inclusions. Recall that an analog of the maximum principle does *not generally hold* even in the case of optimal control problems governed by smooth functional-differential equations of neutral type.

The following relationships between the extended Euler-Lagrange and Hamiltonian inclusions are based on Rockafellar's dualization theorem [25] that concerns subgradients of abstract Lagrangian and Hamiltonian associated with set-valued mappings regardless the dynamics in (1.2). For simplicity we consider the case of the Mayer problem (P_M) for autonomous functional-differential inclusions of neutral type. Then the *Hamiltonian* function for F in (1.2) is defined by

$$
(6.7) \qquad H(x, y, p) := \sup\big\{\langle p, v\rangle \big|\, v \in F(x, y)\big\}.
$$

Corollary 6.2 *Let $\bar{x}(\cdot)$ be an optimal solution to the Mayer problem (P_M) for the autonomous neutral functional-differential inclusion* (1.2) *under the assumptions of Theorem 6.1. Then there exist a number $\lambda \geq 0$ and piecewise continuous functions $p\colon [a, b+\Delta] \to I\!R^n$ and $q\colon [a-\Delta, b] \to I\!R^n$ such that $p(t) - A^*p(t+\Delta)$ and $q(t-\Delta) - A^*q(t)$ are absolutely continuous on $[a, b]$ and, besides* (6.1)–(6.4), *one has the extended Hamiltonian inclusion*

$$\begin{aligned}(6.8)\quad &\Big(\frac{d}{dt}[p(t) - A^*p(t+\Delta)],\ \frac{d}{dt}[q(t-\Delta) - A^*q(t)]\Big)\\ &\in \text{co}\Big\{(u, w)\ \Big|\ \Big(-u, -w, \frac{d}{dt}[\bar{x}(t) - A\bar{x}(t-\Delta)]\Big) \in \partial H\big(\bar{x}(t), \bar{x}(t-\Delta), p(t)+q(t)\big)\Big\}\end{aligned}$$

and the maximum condition

$$(6.9)\qquad \Big\langle p(t)+q(t), \frac{d}{dt}[\bar{x}(t) - A\bar{x}(t-\Delta)]\Big\rangle = H\big(\bar{x}(t), \bar{x}(t-\Delta), p(t)+q(t)\big)$$

for almost all $t \in [a, b]$. If moreover F is convex-valued around $(\bar{x}(t), \bar{x}(t-\Delta))$, then (6.8) *is equivalent to the Euler-Lagrange inclusion*

$$\begin{aligned}(6.10)\quad &\Big(\frac{d}{dt}[p(t) - A^*p(t+\Delta)],\ \frac{d}{dt}[q(t-\Delta) - A^*q(t)]\Big)\\ &\in \text{co}\, D^*F\Big(\bar{x}(t), \bar{x}(t-\Delta), \frac{d}{dt}[\bar{x}(t) - A\bar{x}(t-\Delta)]\Big)\big(-p(t) - q(t)\big)\quad \textit{a.e.}\quad t \in [a, b],\end{aligned}$$

which automatically implies the maximum condition (6.9) *in this case.*

Proof. Since (P_M) is stable with respect to relaxation, $\bar{x}(\cdot)$ is an optimal solution to the relaxed problem (R_M) whose only difference from (P_M) is that the neutral functional-differential inclusion (1.2) is replaced by its convexification (3.4). By Theorem 6.1 the optimal solution $\bar{x}(\cdot)$ satisfies conditions (6.1)–(6.4) and the relaxed counterpart of (6.6), which is the same as (6.10) in this case with F replaced by $\text{co}\, F$. According to [25, Theorem 3.3] one has

$$\text{co}\Big\{(u, v)\Big|\ (u, w, p) \in N((x, y, v); \text{gph}(\text{co}\, F)\Big\} = \text{co}\Big\{(u, w)\Big|\ (-u, -w, v) \in \partial H_R(x, y, p)\Big\},$$

where H_R stands for the Hamiltonian (6.7) of the relaxed system, i.e., with F replaced by $\text{co}\, F$. It is easy to check that $H_R = H$. Thus the extended Euler-Lagrange inclusion for the relaxed system implies the extended Hamiltonian inclusion (6.8), which surely yields the maximum condition (6.9). When F is *convex-valued*, (6.8) and (6.10) are *equivalent* due to the mentioned result of [25]. This completes the proof of the corollary. $\triangle$

Remark 6.3 Let us emphasize that the necessary optimality conditions of Theorem 6.1 and Corollary 6.2 are derived for neutral functional-differential systems with *nonzero delays* $\Delta \neq 0$, which is essentially used in the construction of discrete approximations in Section 2. However, by passing to the limit as $\Delta \downarrow 0$, one can get necessary optimality conditions for *implicit differential inclusions* given in the form

$$(6.11)\qquad \frac{d}{dt}[Ax(t)] \in F(x(t), t)\quad \text{a.e.}\quad t \in [a, b],$$

where the $n \times n$ matrix A may be *singular*, $x(\cdot)$ is continuous while $Ax(\cdot)$ is absolutely continuous on $[a, b]$. The corresponding analogs of the Euler-Lagrange and Hamiltonian inclusions of Corollary 6.2 for the constrained Mayer problem (P_0) over the autonomous inclusion (6.11) are the following:

$$\frac{d}{dt}[A^*p(t)] \in \text{co}\, D^*F\big(\bar{x}(t), \frac{d}{dt}[A\bar{x}(t)]\big)\big(-p(t)\big) \quad \text{a.e. } t \in [a, b],$$

$$\frac{d}{dt}[A^*p(t)] \in \text{co}\,\big\{u \in I\!R^n \big|\, \big(-u, \frac{d}{dt}[A\bar{x}(t)]\big) \in \partial H\big(\bar{x}(t), p(t)\big)\big\} \quad \text{a.e. } t \in [a, b],$$

where the adjoint arc $p(\cdot)$ is continuous while $A^*p(\cdot)$ is absolutely continuous on $[a, b]$ satisfying the transversality condition

$$(p(a), -p(b)) \in \lambda\partial\varphi(\bar{x}(a), \bar{x}(b)) + N((\bar{x}(a), \bar{x}(b)); \Omega)$$

and the maximum condition

$$\langle p(t), \frac{d}{dt}[A\bar{x}(t)]\rangle = H\big(\bar{x}(t), p(t)\big) \quad \text{a.e. } t \in [a, b]$$

with $H(x, p) := \sup\{\langle p, v\rangle \mid v \in F(x)\}$. To establish these conditions, we approximate problem (P_0) by perturbed problems (P_Δ) for neutral functional-differential inclusions with $\Delta > 0$ and pass to the limit as $\Delta \downarrow 0$ similarly to the proof of Theorem 6.1 under the relaxation stability of (P_0).

References

[1] T. S. ANGELL, *On conrollability for nonlinear hereditary systems: a fixed point approach*, Nonlinear Anal., 4 (1980), pp. 93–107.

[2] J.-P. AUBIN and A. CELLINA, *Differential Inclusions*, Springer-Verlag, Berlin. 1984.

[3] H. T. BANKS and A. MANITIUS, *Applications of abstract variational theory to hereditary systems-a survey*, IEEE Trans. Automat. Control. 19 (1974), pp. 524–533.

[4] R. BELLMAN and K. L. COOKE, *Differential-Difference Equations*, Academic Press, New York, 1963.

[5] F. H. CLARKE, YU. S. LEDYAEV, R. J. STERN and P. R. WOLENSKI, *Nonsmooth Analysis and Control Theory*, Spinger-Verlag, New York, 1998.

[6] F. H. CLARKE and G. G. WATKINS, *Necessary conditions, controllability and the value function for diffferential-difference inclusions*, Nonlinear Anal. 10 (1986), pp. 1155–1179.

[7] F. H. CLARKE and P. R. WOLENSKI, *Necessary conditions for functional differential inclusions*, Appl. Math. Optim., 34 (1996), pp. 34–51.

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

[9] R. GABASOV and F. M. KIRILLOVA, *Maximum Principle in Theory of Optimal Control*, Nauka i Technika, Minsk, 1974.

[10] J. HALE, *Theory of Functional Differential Equations*, Springer-Verlag, New York, 1977.

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

[12] M. Q. JACOBS and C. E. LANGENHOP, *Criteria for function space controllability of linear neutral systems*, SIAM J. Control Optim., 14 (1986), pp. 1009–1948.

[13] G. A. KENT, *A maximum principle for optimal control problems with neutral functional differential systems*, Bull. Amer. Math. Soc., 77 (1971).

[14] M. KISIELEWICZ, *Differential Inclusions and Optimal Control*, Kluwer, Dordrecht, The Netherlands, 1991.

[15] V. B. KOLMANOVSKII and L. E. SHAIKHET, *Control of Systems with Aftereffect*, Acedemic Press, New York, 1996.

[16] P. D. LOEWEN and R. T. ROCKAFELLAR, *Bolza problems with general time constraints*, SIAM J. Control Optim., 35 (1997), pp. 2050–2069.

[17] L. I. MINCHENKO. *Necessary optimality conditions for differential–difference inclusions*, Nonlinear Anal., 35 (1999), pp. 307–322.

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

[19] B. S. MORDUKHOVICH. *Approximation Methods in Problems of Optimization and Control*, Nauka, Moscow, 1988; 2nd edition to appear in Wiley, New York.

[20] B. S. MORDUKHOVICH. *Complete characterization of openness, metric regularity, and Lipschitzian properties of multifunctions*. Trans. Amer. Math. Soc., 340 (1993), pp. 1–35.

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

[22] B. S. MORDUKHOVICH. J. S. TREIMAN and Q. J. ZHU, *An extended extremal principle with applications to multiobjective optimization*. preprint. 2002.

[23] B. S. MORDUKHOVICH and R. TRUBNIK, *Stability of discrete approximation and necessary optimality conditions for delay-differential inclusions*, Ann. Oper. Res., 101 (2001). pp. 149-170.

[24] B. S. MORDUKHOVICH and L. WANG. *Optimal control of constrained delay-differential inclusions with multivalued initial conditions*. Control and Cybernetics, to appear.

[25] R. T. ROCKAELLAR. *Equivalent subgradient versions of Hamiltonian and Euler–Lagrange conditions in variational analysis*. SIAM J. Control Optim., 34 (1996), pp. 1300–1314.

[26] R. T. ROCKAFELLAR and R. J.-B. WETS. *Variational Analysis*, Springer-Verlag. Berlin. 1998.

[27] G. V. SMIRNOV, *Introduction to the Theory of Differential Inclusions*, American Mathematical Society. Providence, RI, 2002.

[28] H. J. SUSSMANN, *New theories of set-valued differentials and new versions of the maximum principle in optimal control theory*, in Nonlinear Control in the Year 2000, A. Isidori et al., eds., Springer-Verlag, Berlin, 2000, pp. 487–472.

[29] R. B. VINTER, *Optimal Control*, Birkhäuser, Boston, 2000.

[30] J. WARGA, *Optimal Control of Differential and Functional Equations*, Academic Press, Newc York, 1972.

[31] Q. J. ZHU, *Hamiltonian necessary conditions for a multiobjective optimal control problem with endpoint constraints*, SIAM J. Control Optim., 39 (2000), pp. 97–112.