Control and Cybernetics

vol. **32** (2003) No. 3

Optimal control of constrained delay-differential inclusions with multivalued initial conditions [1]

by

Boris S. Mordukhovich and Lianwen Wang

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

Abstract: This paper studies a general optimal control problem for nonconvex delay-differential inclusions with endpoint constraints. In contrast to previous publications on this topic, we incorporate time-dependent set constraints on the initial interval, which are specific for systems with delays and provide an additional source for optimization. Our variational analysis is based on well-posed discrete approximations of constrained delay-differential inclusions by a family of time-delayed systems with discrete dynamics and perturbed constraints. Using convergence results for discrete approximations and advanced tools of nonsmooth variational analysis, we derive necessary optimality conditions for constrained delay-differential inclusions in both Euler-Lagrange and Hamiltonian forms involving nonconvex generalized differential constructions for nonsmooth functions, sets, and set-valued mappings.

Keywords: delay-differential inclusions, discrete approximations, necessary optimality conditions, variational analysis, stability, nonsmooth optimization, generalized differentiation.

1. Introduction

The primary object of this paper is the following *generalized Bolza problem* (P) for delay-differential inclusions with general initial conditions and endpoint constraints:

$$\text{minimize } J[x] := \varphi(x(a), x(b)) + \int_a^b f(x(t), x(t-\Delta), \dot{x}(t), t)\, dt \tag{1}$$

[1]This research was partly supported by the National Science Foundation under grants DMS-0072179 and DMS-0304989.

over arcs $x : [a - \Delta, b] \to I\!R^n$ with $\Delta \geq 0$, that are absolutely continuous on $[a, b]$ and L^∞ functions on $[a - \Delta, a]$, subject to

$$\dot{x}(t) \in F(x(t), x(t - \Delta), t) \quad \text{a.e. } t \in [a, b], \tag{2}$$

$$x(t) \in C(t) \quad \text{a.e. } t \in [a - \Delta, a), \tag{3}$$

$$(x(a), x(b)) \in \Omega \subset I\!R^{2n}. \tag{4}$$

For the nondelayed systems ($\Delta = 0$) this problem was studied in a number of publications mainly devoted to necessary optimality conditions; see Clarke (1983), Ioffe (1997), Loewen and Rockafellar (1997), Mordukhovich (1995), Sussmann (2000), Vinter (2000), Zhu (1996), and the references therein. To the best of our knowledge, there are just a few papers devoted to the study of optimization problems for delay-differential inclusions (mostly with the Mayer-type cost functional); see Clarke and Watkins (1986), Clarke and Wolenski (1996), Minchenko (1999), and Mordukhovich and Trubnik (2001). These papers (except Clarke and Watkins, 1986, for a free-endpoint Mayer problem) concern delay-differential inclusions with the initial condition (3) given by a single-valued mapping $C(t) = \{c(t)\}$ that closely relates delayed systems to their nondelayed counterparts.

The present paper deals with the generalized Bolza problem (P) involving a *set-valued* mapping $C(t)$ in the initial condition (3), which is *specific for delay-differential systems* and essentially distinguishes them from nondelayed ones. A choice of the initial function $x(t)$ from the set $C(t)$ on $[a - \Delta, a)$ provides an additional source for optimizing the cost functional (1) subject to the constraints (2)–(4).

We employ the *method of discrete approximations* for the study of problem (P). This method is based on the finite-difference replacement of the derivative

$$\dot{x}(t) \approx \big[x(t + h) - x(t)\big]/h, \quad h \to 0, \tag{5}$$

in (2) with appropriate approximations of the cost functional and endpoint constraints. The method of discrete approximations is well-developed in the case of ordinary control systems (one of the pioneering work was done by Malanowski, 1979, see also a more recent survey by Dontchev, 1996). In contrast to the vast majority of publications on discrete approximations, we mostly focus not on numerical aspects of this method (particularly involving estimates of convergence rates in various finite-difference schemes) but rather on *qualitative aspects* allowing us to use discrete approximations as a *vehicle* for deriving necessary optimality conditions in continuous-time systems. Such an approach to optimization of nondelayed differential inclusions was developed by Mordukhovich (1988, 1995) (see also the recent book by Smirnov, 2002, and the references therein); related developments for delay-differential problems with single-valued initial conditions were given in Mordukhovich and Trubnik (2001) in an essentially different framework.

The method of discrete approximations applied to the problem (1)–(4) under consideration allows us to build a well-posed sequence of *finite-dimensional* optimization problems for time-delayed discrete inclusions with a *strong convergence* of optimal solutions; see below. The obtained finite-dimensional problems are intrinsically *nonsmooth* (containing in fact an increasing number of set constraints with possibly empty interiors), but they fortunately can be handled by generalized differential tools of modern variational analysis involving *nonconvex-valued* normal cones, subdifferentials, and coderivatives that enjoy full calculi. Using these tools, we first derive necessary optimality conditions in delay-difference counterparts of the original problem (P). Then, by passing to the limit from discrete approximations, we obtain necessary optimality conditions for problem (P) in the extended Euler-Lagrange form, which is equivalent to the enhanced Hamiltonian form under additional assumptions.

In this paper we relax, in the case of *nonautonomous* systems, some assumptions previously made in the method of discrete approximations even for nondelayed differential inclusions. To furnish this, we employ, along with the basic/limiting normal cone, subdifferential, and coderivative as in Mordukhovich (1995), their *extended counterparts* for time-dependent sets, functions, and set-valued mappings discussed in Section 3.

The rest of the paper is organized as follows. In Section 2 we construct well-posed discrete approximations of the original problem (1)–(4), which ensure the required *strong convergence* of optimal solutions under minimal assumptions. Section 3 presents basic constructions and necessary background of generalized differentiation that are needed for the variational analysis of discrete-time and continuous-time systems performed in this paper. In Section 4 we obtain necessary optimality conditions for nonconvex delay-difference inclusions arising in discrete approximations of the above problem (P). The concluding Section 5 contains necessary optimality conditions in the Euler-Lagrange and Hamiltonian forms for problem (P) derived via its discrete approximations under the assumption on relaxation stability discussed in Section 2.

Our notation is basically standard, see Mordukhovich (1995) and Rockafellar and Wets (1998).

2. Well-posed discrete approximations

The main goal of this section is to construct *well-posed* discrete approximations of the original problem (P) that ensure the *strong convergence* of optimal trajectories in the norm topologies of $W^{1,2}[a,b]$ and $L^2[a-\Delta,a]$, respectively. Such a strong convergence plays a crucial role in the study of delay-differential inclusions via discrete approximations.

Let $\bar{x}(\cdot)$ be a feasible trajectory for (2) with the initial condition (3). We impose the following assumptions, where $I\!B$ stands for the closed unit ball in $I\!R^n$.

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

$$F(x,y,t) \subset M_F I\!B \text{ for all } (x,y,t) \in U \times U \times [a,b], \tag{6}$$

$$F(x_1,y_1,t) \subset F(x_2,y_2,t) + L_F(|x_1-x_2| + |y_1-y_2|)I\!B \tag{7}$$

whenever $(x_1,y_1),\ (x_2,y_2) \in U \times U$ and $t \in [a,b]$.

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

(H3) The multifunction $C\colon [a-\Delta, a] \rightrightarrows I\!R^n$ is closed-valued, uniformly bounded, and Hausdorff continuous for a.e. $t \in [a-\Delta, a]$.

Following Dontchev and Farkhi (1989), we consider the so-called *averaged modulus of continuity* for the multifunction $F(x,y,t)$ in $t \in [a,b]$ when $(x,y) \in U \times U$ defined by:

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

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

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

and where $\text{haus}(\cdot,;\cdot)$ stands for the Hausdorff distance between two compact sets. It is proved in the mentioned paper that if $F(x,y,\cdot)$ 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 $h \to 0$. Of course, a simplified version of the above definition applies to the average modulus of continuity $\tau[C;h]$ of the multifunction $C(\cdot)$ on $[a-\Delta, a]$.

Let us construct a discrete approximation of (2) based on the Euler finite-difference replacement of the derivative (5). For any $N \in I\!N := \{1,2,\ldots\}$ we take $t_j := a + jh_N$ for $j = -N,\ldots,0,1,\ldots,k$ and $t_{k+1} := b$, where $h_N := \frac{\Delta}{N}$ and $k \in I\!N$ is defined by $a + kh_N \le b < a + (k+1)h_N$. Note that $t_{-N} = a - \Delta,\ t_0 = a$, and $h_N \to 0$ as $N \to \infty$. Then, the sequence of *delay-difference inclusions* approximating (2) is constructed as follows:

$$\begin{cases} x_N(t_{j+1}) \in x_N(t_j) + h_N F(x_N(t_j), x_N(t_j - \Delta), t_j) & \text{for j = 0,\ldots,k,} \\ x_N(t_j) \in C(t_j) & \text{for j = -N, \ldots,-1.} \end{cases} \tag{8}$$

A collection of vectors $\{x_N(t_j)|\ j = -N,\ldots,k+1\}$ satisfying (8) is called a *discrete trajectory*. The corresponding collection

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

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

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

and the corresponding *extended discrete trajectories* defined by

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

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

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

on the initial interval $[a - \Delta, a)$. Observe that

$$\dot{x}_N(t) = v_N(t) \quad \text{a.e. } t \in [a, b].$$

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

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

The next theorem ensures the *strong approximation* of $\bar{x}(\cdot)$ by feasible trajectories of delay-difference inclusions (8).

THEOREM 2.1 *Let $\bar{x}(\cdot)$ be a feasible trajectory to (2) and (3) under assumptions (H1)–(H3). Then there exists a sequence $\{z_N(t_j) |\ j = -N, \ldots, k+1\}$ of solutions to the delay-difference inclusions (8) with*

$$z_N(t_0) := z_N(a) = \bar{x}(a)$$

such that the extended discrete trajectories $z_N(t)$, $a - \Delta \le t \le b$, converge to $\bar{x}(\cdot)$ in the L^2-norm on $[a - \Delta, a]$ and in the $W^{1,2}$-norm on $[a, b]$ as $N \to \infty$.

Proof. Due to (6) and the uniform boundedness of $C(\cdot)$ in (H3), it is sufficient to establish the required convergence in the norm topologies of $L^1[a - \Delta, a]$ and $W^{1,1}[a, b]$. Let $\{w_N(\cdot)\}$, $N \in I\!\!N$, be a sequence of functions on $[a - \Delta, b]$, with $w_N(a) := \bar{x}(a)$, that are constant on the interval $[t_j, t_{j+1})$, $j = -N, \ldots, k$, and converge to $\bar{x}(\cdot)$ on $[a - \Delta, a]$ and to $\dot{\bar{x}}(\cdot)$ on $[a, b]$, respectively, in the norm topology of L^1. Such a sequence always exists because of the density of step-functions in $L^1[a - \Delta, b]$. In the estimates below we use the sequence

$$\xi_N := \int_{a-\Delta}^a |\bar{x}(t) - w_N(t)| dt + \int_a^b |\dot{\bar{x}}(t) - w_N(t)| dt \to 0 \text{ as } N \to \infty. \quad (9)$$

Observe that due to the uniform boundedness assumptions in (H1) and (H3), there is $M > 0$ with

$$|w_N(t)| \le M \quad \text{for all} \quad t \in [a-\Delta, b] \quad \text{and} \quad N \in I\!N.$$

Denote $w_{N_j} := w_N(t_j)$ for $j = -N, \ldots, k+1$ and define the discrete functions $\{u_N(t_j) |\ j = -N, \ldots, k+1\}$ by

$$\begin{cases} u_N(t_j) := w_{N_j} \quad \text{for } j = -N, \ldots, 0, \\ u_N(t_{j+1}) := u_N(t_j) + h_N w_{N_j} \quad \text{for j = 0,\ldots, k.} \end{cases} \tag{10}$$

The extensions of these functions on the continuous intervals $[a - \Delta, a)$ and $[a, b]$, respectively, are given by

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

Let $\mathrm{dist}(w; \Omega)$ be the *Euclidean distance* between the point w and the closed set Ω. Then the Lipschitz condition (7) can be written as

$$\mathrm{dist}(w; F(x_1, t)) \le \mathrm{dist}(w; F(x_2)) + L_F|x_1 - x_2|, w \in I\!R^n, \; x_1, x_2 \in U, \; t \in [a, b],$$

and one obviously has

$$\mathrm{dist}(w; F(x, t_1)) \le \mathrm{dist}(w; F(x, t_2)) + \mathrm{haus}(F(x, t_1); F(x, t_2)), \; w, x \in I\!R^n.$$

Using this and the average modulus of continuity, we get

$$\begin{aligned} \alpha_N : &= h_N \sum_{j=-N}^{-1} \mathrm{dist}(w_{N_j}; C(t_j)) + h_N \sum_{j=0}^{k} \mathrm{dist}(w_{N_j}; F(u_N(t_j), t_j) \\ &= \sum_{j=-N}^{-1} \int_{t_j}^{t_{j+1}} \mathrm{dist}(w_{N_j}; C(t_j))dt + \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} \mathrm{dist}(w_{N_j}; F(u_N(t_j), t_j))dt \\ &\le \sum_{j=-N}^{-1} \int_{t_j}^{t_{j+1}} \mathrm{dist}(w_{N_j}; C(t))dt + \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} \mathrm{dist}(w_{N_j}; F(u_N(t_j), t))dt \\ &+ \tau[C; h_N] + \tau[F; h_N]. \end{aligned}$$

Taking into account the facts that $\big(\xi_N, \tau[C; h_N], \tau[F; h_N]\big) \to 0$ as $N \to \infty$ due to (9) and assumptions (H2) and (H3), that $\bar{x}(t) \in C(t)$ for a.e. $t \in [a - \Delta, a)$, and that

$$\begin{aligned} \mathrm{dist}(w_N(t); F(u_N(t), t)) &\le \mathrm{dist}(w_N(t); F(\bar{x}(t), t)) + L_F|u_N(t) - \bar{x}(t)| \\ &\le |w_N(t) - \dot{\bar{x}}(t)| + L_F \xi_N, \end{aligned}$$

one has the estimate

$$\alpha_N \le (2+L_F)\xi_N + l_F(b-a)(M_F+1)h_N/2 + \tau[C;\triangle_N] + \tau[F;h_N] \to 0\, as\ N \to \infty.$$

Note that the discrete functions defined in (10) may not satisfy (8), since one does not generally have $w_{N_j} \in C(t_j)$ for $j = -N, \dots, -1$ and $w_{N_j} \in F(u_N(t_j), u_N(t_j - \Delta), t_j)$ for $j = 1, \dots, k$. Let us construct the desired trajectories $\{z_N(t_j) |\ j = -N, \dots, k+1\}$ by the following *proximal algorithm*:

$$\begin{cases} z_N(t_j) = v_{N_j} \ \text{ with } \ |v_{N_j} - w_{N_j}| = \text{dist}(w_{N_j}; C(t_j)) \text{ for } \ j = -N, \dots, -1, \\ z_N(t_0) = \bar{x}(a), \quad z_N(t_{j+1}) = z_N(t_j) + h_N v_{N_j} \ \text{ for } j = 0, \dots, k, \\ \text{with } \ v_{N_j} \in F(z_N(t_j), z_N(t_j - \Delta), t_j), \quad |v_{N_j} - w_{N_j}| = \\ \text{dist}(w_{N_j}; F(z_N(t_j), z_N(t_j \Delta), t_j)). \end{cases} \tag{11}$$

One can see that all $z_N(\cdot)$ in (11) are feasible trajectories for (8). Now, following the scheme in the proof of Theorem 2.1 in Mordukhovich (1995) and adapting it to the case of delayed systems with set-valued initial conditions under consideration, we show that the extensions $z_N(t)$, $t \in [a - \Delta, b]$, of the above discrete trajectories converge to $\bar{x}(t)$ in the L^2-norm on $[a - \Delta, a]$ and in the $W^{1,2}$-norm on $[a, b]$. Moreover, we can get efficient estimates of the convergence rate that involve ξ_N in (9), the modulus $\tau[C; h_N]$ and $\tau[F; h_N]$, and the constants defined in (H1). ■

Our next goal is to construct a well-posed discrete approximation of the whole dynamic optimization problem (1)–(4) (not only of the delay-differentialinclusion) such that optimal solutions to discrete approximation problems strongly converge to a given optimal solution $\bar{x}(\cdot)$ to the original problem (P). The following construction *explicitly* involves the optimal solution $\bar{x}(\cdot)$ to problem (P) under consideration.

Given $\bar{x}(t)$, $a - \Delta \le t \le b$, take its approximation $z_N(t)$ from Theorem 2.1 and denote $\eta_N := |z_N(b) - \bar{x}(b)|$. For any $N \in I\!N$ we consider the dynamic optimization problem (P_N) for constrained delay-difference inclusions:

$$\begin{aligned} \text{minimize } J_N[x_N] &:= \varphi(x_N(a), x_N(b)) + |x_N(a) - \bar{x}(a)|^2 \\ &+ \sum_{j=-N}^{-1} \int_{t_j}^{t_{j+1}} |x_N(t_j) - \bar{x}(t)|^2 dt \\ &+ h_N \sum_{j=0}^{k} f\big(x_N(t_j), x_N(t_j - \Delta), \frac{x_N(t_{j+1}) - x_N(t_j)}{h_N}, t_j\big) \\ &+ \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} \Big|\frac{x_N(t_{j+1}) - x_N(t_j)}{h_N} - \dot{\bar{x}}(t)\Big|^2 dt \end{aligned} \tag{12}$$

subject to the constraints

$$x_N(t_{j+1}) \in x_N(t_j) + h_N F(x_N(t_j), x_N(t_j - \Delta), t_j), \quad j = 0, \ldots, k, \tag{13}$$

$$x_N(t_j) \in C(t_j), \quad j = -N, \ldots, -1, \tag{14}$$

$$(x_N(a), x_N(b)) \in \Omega_N := \Omega + \eta_N I\!B, \tag{15}$$

$$|x_N(t_j) - \bar{x}(t_j)| \le \varepsilon, \quad j = 1, \ldots, k+1, \tag{16}$$

where ε is a given positive number. In addition to (H1)–(H3) with some neighborhood U of $\bar{x}(t)$, we impose the following hypotheses on the behavior of φ, f, and Ω around the optimal trajectory:

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

(H5) There exists $\mu > 0$ such that $f(\cdot, \cdot, \cdot, t)$ is continuous on the set

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

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

In what follows we select $\varepsilon > 0$ in (16) such that $\bar{x}(t) + \varepsilon I\!B \subset U$ for all $t \in [a - \Delta, b]$ and take sufficiently large N satisfying $\eta_N < \varepsilon$. Note that problems (P_N) have feasible solutions, since the trajectories z_N from Theorem 2.1 satisfy all the constraints (13)–(16) for large N. Moreover, the sets of feasible solutions to (P_N) are bounded for all N due to (14) and (16). Hence, each (P_N) admits an optimal solution $\bar{x}_N(\cdot)$ by the classical Weierstrass theorem in finite dimensions.

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*. Let us consider, along with (2), the *convexified* delay-differential inclusion

$$\dot{x}(t) \in \operatorname{co} F(x(t), x(t - \Delta), t) \quad \text{a.e. } t \in [a, b], \tag{17}$$

where "co" stands for the convex hull of a set. Further, given the integrand f in (1), we consider its extension

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

with respect to the set-valued mapping F of (2), where $\delta(\cdot; F)$ stands for the indicator function of a set. Denote by $\widehat{f}_F(x, y, v, t)$ the *convexification* of f_F in the v variable and define the *relaxed generalized Bolza problem* (R) as follows:

$$\text{minimize } \widehat{J}\,[x] := \varphi(x(a), x(b)) + \int_a^b \widehat{f}_F(x(t), x(t - \Delta), \dot{x}(t), t)\, dt \tag{18}$$

over functions $x\colon [a-\Delta,b] \to I\!R^n$, which are absolutely continuous on $[a,b]$ and continuous on $[a-\Delta,a)$, subject to (3) and the endpoint constraints (4). It follows from the structure of (18) that the condition $\widehat{J}\,[x(\cdot)] < \infty$ implies that $x(\cdot)$ is a trajectory for the convexified delay-differential inclusion (17) called a *relaxed trajectory* for (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 assumptions, turns out also to be natural for nonconvex continuous-time problems governed by differential and delay-differential inclusions due to the inherent "hidden convexity" of such systems (related to the convexity of integrals for set-valued mappings over nonatomic measures). In particular, the following *fundamental approximation property* holds under the assumed Lipschitz continuity of F in (x,y): Every relaxed trajectory $x(\cdot)$ can be uniformly on $[a,b]$ approximated by original trajectories $x_m(\cdot)$ with the same initial history $x_m(t) = x(t)$ on $[a-\Delta,a]$ and

$$\liminf_{m\to\infty} \int_a^b f(x_m(t), x_m(t-\Delta), \dot{x}_m(t), t)\, dt \le \int_a^b \widehat{f}_F(x(t), x(t-\Delta), \dot{x}(t), t)\, dt.$$

This result, which ensures the relaxation stability of problems (P) for delay-differential inclusions with no endpoint constraints at $t=b$, can be proved similarly to the one for nondelayed differential inclusions; see Aubin and Cellina (1984). The reader can find in Mordukhovich (1995), more discussions and references on the validity of this property for nonconvex constrained systems governed by differential inclusions. Similar results and discussions also hold for the delay-differential inclusions under consideration, see also the book of Warga (1972) for various classes of functional differential control systems.

To be able to establish the desired strong convergence of discrete approximations, we have to impose the following additional assumptions that are specific for delay-differential inclusions with *set-valued* initial conditions as in (3).

(H6) $C(t)$ is convex for a.e. $t \in [a-\Delta,a]$; $F(x,y,t)$ is linear in y for a.e. $t \in [a,a+\Delta]$; $f(x,y,v,t)$ is convex in (y,v) for a.e. $t \in [a,a+\Delta]$.

Now we are ready to establish the strong convergence theorem that makes a bridge between optimization problems for delay-differential and delay-difference inclusions.

THEOREM 2.2 *Let $\bar{x}(\cdot)$ be an optimal solution to problem (P), which is assumed to be stable with respect to relaxation. Assume also that hypotheses (H1)–(H6)* *hold. Then, any sequence* $\{\bar{x}_N(\cdot)\}$, $N \in I\!N$, *of optimal solutions to* (P_N), *extended to the continuous interval* $[a-\Delta, b]$, *converges to* $\bar{x}(\cdot)$ *in the* L^2*-norm on* $[a-\Delta, a]$ *and 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

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

Similarly to the proof of Theorem 3.3 in Mordukhovich (1995) for the case of nondelayed differential inclusions, we can show that

$$J_N[z_N] \to J[\bar{x}] \quad \text{as} \quad N \to \infty.$$

The above two relationships yield

$$\limsup_{N\to\infty} J_N[\bar{x}_N] \le J\ [\bar{x}]. \tag{19}$$

To justify the required convergence $\bar{x}_N(\cdot) \to \bar{x}(\cdot)$, we need to prove that

$$\rho_N := \int_{a-\Delta}^{a} |\bar{x}_N(t) - \bar{x}(t)|^2 dt + |\bar{x}_N(a) - \bar{x}(a)|^2 + \int_a^b |\dot{\bar{x}}_N(t) - \dot{\bar{x}}(t)|^2 dt \to 0 \text{ as } N \to \infty.$$

Suppose it is not true. Then, we can find a constant $c > 0$ and a subsequence $\{N_m\} \subset I\!N$ such that $\rho_{N_m} \to c$. Without loss of generality assume that $\rho_N \to c$ as $N \to \infty$. Since the sequences $\{\bar{x}_N(\cdot)\}$ and $\{\dot{\bar{x}}_N(\cdot)\}$ are uniformly bounded on $[a-\Delta, a]$ and $[a, b]$, respectively, under the assumptions made there is a function $\widetilde{x}\colon [a-\Delta, b] \to I\!R^n$ belonging to L^2 on $[a-\Delta, a]$ and to $W^{1,2}$ on $[a, b]$ such that

$$\begin{array}{rll} \bar{x}_N(\cdot) \to \widetilde{x}(\cdot) & \text{weakly in} & L^2[a-\Delta, a] \\ \text{and } \dot{\bar{x}}_N(\cdot) \to \dot{\widetilde{x}} & \text{weakly in} & L^2[a, b] \end{array} \tag{20}$$

along a subsequence of $N \in I\!N$, which is supposed to be equal to the whole $I\!N$. Invoking now the classical Mazur theorem, we conclude that there are *convex combinations* of the sequences in (20) converging *pointwisely* to $\widetilde{x}(t)$ and to $\dot{\widetilde{x}}(t)$ for a.e. $t \in [a-\Delta, a]$ and $[a, b]$, respectively.

It follows from (H3) and the convexity of $C(t)$ that $\widetilde{x}(t) \in C(t)$ for a.e. $t \in [a-\Delta, a]$. Taking into account the assumptions on F in (H2) and (H6) and that $\bar{x}_N(t) \to \widetilde{x}(t)$ uniformly on $[a, b]$, we arrive at the convexified inclusion

$$\dot{\widetilde{x}}(t) \in \operatorname{co} F(\widetilde{x}(t), \widetilde{x}(t-\Delta), t) \text{ a.e. } t \in [a, b].$$

Due to the corresponding assumptions on f and by

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

one has the inequality

$$\int_a^b \widehat{f}_F(\widetilde{x}(t), \widetilde{x}(t-\Delta), \dot{\widetilde{x}}(t), t) dt$$
$$\le \liminf_{N\to\infty} h_N \sum_{j=0}^{k} f(\bar{x}_N(t_j), \bar{x}_N(t_j - \Delta), \frac{\bar{x}_N(t_{j+1}) - \bar{x}_N(t_j)}{h_N}, t_j),$$

since $\widehat{f}_F \le f_F$. Observe further that the integral functionals

$$I_1[v] := \int_{a-\Delta}^{a} |v(t) - \bar{x}(t)|^2 dt \;\text{ and }\; I_2[v] := \int_a^b |v(t) - \dot{\bar{x}}(t)|^2 dt$$

are lower semicontinuous in the weak topology of $L^2[a-\Delta, a]$ and $L^2[a,b]$, respectively, due to the *convexity* of the integrands in v. Since

$$\sum_{j=-N}^{-1} \int_{t_j}^{t_{j+1}} |\bar{x}_N(t_j) - \bar{x}(t)|^2 dt = \int_{a-\Delta}^{a} |\bar{x}_N(t) - \bar{x}(t)|^2 dt \;\text{ and}$$
$$\sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} \Big| \frac{\bar{x}_N(t_{j+1}) - \bar{x}_N(t_j)}{h_N} - \dot{\bar{x}}(t) \Big|^2 dt = \int_a^b |\dot{\bar{x}}_N(t) - \dot{\bar{x}}(t)|^2 dt,$$

the latter implies that

$$\int_{a-\Delta}^{a} |\widetilde{x}(t) - \bar{x}(t)|^2 dt \le \liminf_{N\to\infty} \sum_{j=-N}^{-1} \int_{t_j}^{t_{j+1}} |\bar{x}_N(t_j) - \bar{x}(t)|^2 dt \;\text{ and}$$
$$\int_a^b |\dot{\widetilde{x}}(t) - \dot{\bar{x}}(t)|^2 dt \le \liminf_{N\to\infty} \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} \Big| \frac{\bar{x}_N(t_{j+1}) - \bar{x}_N(t_j)}{h_N} - \dot{\bar{x}}(t) \Big|^2 dt.$$

Now passing to the limit in (12), we arrive at

$$\widehat{J}[\widetilde{x}] + c \le \liminf_{N\to\infty} J_N[\bar{x}_N],$$

where $\widetilde{x}$ is a feasible trajectory to the relaxed problem (R) due to the above discussion. So if

$$c = \lim_{N\to\infty} \Big[\int_{a-\Delta}^{a} |\bar{x}_N(t) - \bar{x}(t)|^2 dt$$
$$+ |\bar{x}_N(a) - \bar{x}(a)|^2 + \int_a^b |\dot{\bar{x}}_N(t) - \dot{\bar{x}}(t)|^2 dt \Big] > 0,$$

we get $\widehat{J}[\tilde{x}] < J[\bar{x}] = \widehat{J}[\bar{x}]$ due to (19), which contradicts the optimality of $\bar{x}(\cdot)$ in the relaxed problem and completes the proof of theorem. ■

3. Tools of generalized differentiation

The results of the previous section allow us to make a bridge between the original infinite-dimensional optimization problem (P) for delay-differential inclusions and the sequence of finite-dimensional dynamic optimization problems (P_N) for delay-difference inclusions. Our strategy is to obtain first the necessary optimality conditions for each of 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 unconstrained delay-difference inclusions. The main source of nonsmoothness comes from the (increasing number of) geometric constraints (13) and (14) as $N \to \infty$, which may have empty interiors. To deal with such problems, we use appropriate tools of generalized differentiation introduced by Mordukhovich (1976, 1988) and then developed and applied in many publications; see, in particular, the book of Rockafellar and Wets (1998) for detailed treatments and the extensive bibliography.

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

$$N(\bar{x};\Omega) := \mathop{\rm Lim\,sup}_{x \xrightarrow{\Omega} \bar{x}} \widehat{N}(x;\Omega), \tag{21}$$

where $x \xrightarrow{\Omega} \bar{x}$ means that $x \to \bar{x}$ with $x \in \Omega$, where "Lim sup" stands for the the Painlevé-Kuratowski upper (outer) limit

$$\mathop{\rm Lim\,sup}_{x \to \bar{x}} F(x) := \big\{ y \in Y \big|\ \exists x_k \to \bar{x},\ \exists y_k \to y \ \text{ with }\ y_k \in F(x_k),\ k \in I\!N \big\}$$

for a multifunction $F\colon X \rightrightarrows Y$, and where

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

is the cone of Fréchet (or regular) normals to Ω at $\bar{x}$. Note that for *convex* sets Ω one has

$$N(\bar{x};\Omega) = \widehat{N}(\bar{x};\Omega) = \big\{ x^* \in I\!R^n \big|\ \langle x^*, x - \bar{x}\rangle \le 0 \ \text{ for all }\ x \in \Omega \big\}. \tag{22}$$

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

$$\partial\varphi(\bar{x}) := \big\{ x^* \in I\!R^n \big|\ (x^*, -1) \in N((\bar{x}, \varphi(\bar{x})); \operatorname{epi}\varphi) \big\}, \tag{23}$$

where $\operatorname{epi}\varphi := \{(x,\mu) \in I\!R^{n+1} \mid \mu \geq \varphi(x)\}$. Then, the *coderivative* $D^*F(\bar{x},\bar{y})\colon I\!R^m \rightrightarrows I\!R^n$ of a set-valued mapping $F\colon I\!R^n \rightrightarrows I\!R^m$ at a point $(\bar{x},\bar{y}) \in \operatorname{gph} F$ is defined by

$$D^*F(\bar{x},\bar{y})(y^*) := \big\{x^* \in I\!R^n \big|\ (x^*,-y^*) \in N((\bar{x},\bar{y});\operatorname{gph} F)\big\}. \tag{24}$$

Note the useful relationships

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

between the subdifferential and coderivative, where $E_\varphi(x) := \{\mu \in I\!R \mid \mu \geq \varphi(x)\}$ is the epigraphical multifunction associated with $\varphi\colon I\!R^n \to \overline{I\!R}$ and where $\langle y^*,f\rangle(x) := \langle y^*,f(x)\rangle$ is the scalarized function associated with a locally Lipschitz mapping $f\colon I\!R^n \to I\!R^m$.

For applications in this paper we need to consider proper extensions of the basic constructions (21), (23), and (24) to the case of sets, functions, and set-valued mappings depending on parameters. The following extended constructions fit our requirements.

Definition 3.1 *Let T be a topological space.*

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

$$\widetilde{N}(\bar{x};\Omega(\bar{t})) := \operatorname*{Lim\,sup}_{(t,x)\overset{\operatorname{gph}\Omega}{\to}(\bar{t},\bar{x})} \widehat{N}(x;\Omega(t)). \tag{25}$$

$\Omega(\cdot)$ is said to be **normally semicontinuous** *at $(\bar{x},\bar{t})$ if $\widetilde{N}(\bar{x};\Omega(\bar{t})) = N(\bar{x};\Omega(\bar{t}))$.*

(ii) *Given $\varphi\colon I\!R^n \times T \to \overline{I\!R}$ finite at $(\bar{x},\bar{t})$, the* **extended subdifferential** *of φ at $(\bar{x},\bar{t})$ with respect to x is*

$$\widetilde{\partial}_x\varphi(\bar{x},\bar{t}) := \big\{x^* \in I\!R^n \big|\ (x^*,-1) \in \widetilde{N}((\bar{x},\varphi(\bar{x},\bar{t}));\operatorname{epi}\varphi(\cdot,\bar{t}))\big\}. \tag{26}$$

The function φ is **subdifferentially semicontinuous** *at $(\bar{x},\bar{t})$ with respect to t if*

$$\widetilde{\partial}_x\varphi(\bar{x},\bar{t}) = \partial_x\varphi(\bar{x},\bar{t}),$$

where $\partial_x\varphi(\bar{x},\bar{t})$ stands for the subdifferential (23) of $\varphi(\cdot,\bar{t})$ at $\bar{x}$, i.e., for the partial subdifferential of φ with respect to x. **(iii)** *Given $F\colon I\!R^n \times T \rightrightarrows I\!R^m$ and $\bar{y} \in F(\bar{x},\bar{t})$, we define the* **extended coderivative** *of F at $(\bar{x},\bar{y},\bar{t}) \in \operatorname{gph} F$ with respect to x by*

$$\widetilde{D}^*_xF(\bar{x},\bar{y},\bar{t})(y^*) := \big\{x^* \in I\!R^n \big|\ (x^*,-y^*) \in \widetilde{N}((\bar{x},\bar{y});\operatorname{gph} F(\cdot,\bar{t}))\big\}, \quad y^* \in I\!R^m. \tag{27}$$

The mapping F is **coderivatively semicontinuous** *at $(\bar{x},\bar{y},\bar{t})$ with respect to t if*

$$\widetilde{D}^*_xF(\bar{x},\bar{y},\bar{t})(y^*) = D^*_xF(\bar{x},\bar{y},\bar{t})(y^*) \ \text{ for all } \ y^* \in R^m,$$

where $D^*_xF(\bar{x},\bar{y},\bar{t})$ stands for the coderivative (24) of $F(\cdot,\bar{t})$ at $(\bar{x},\bar{y})$.

It follows from (21) that the extended normal cone does not change if $\widehat{N}(x;\Omega(t))$ is replaced by $N(x;\Omega(t)$ in the limiting procedure (25). Thus the normal semicontinuity of $\Omega(\cdot)$ from Definition 3.1 agrees with the N-(normal) semicontinuity introduced by Mordukhovich (1984, 1988) in connection with covering/metric regularity results, see also Zhu (2000), where this property was used under the name of "regularity." The subdifferential and coderivative semicontinuity assumptions were directly imposed on functions and mappings in Mordukhovich (1995) and Mordukhovich and Trubnik (2001) for deriving necessary optimality conditions in optimal control problems governed by differential and delay-differential inclusions. The extended normal cone (25) was recently used by Bellaassali and Jourani (2002) and by Mordukhovich, Treiman and Zhu (2002) in applications to multiobjective optimization problems without any normal semicontinuity assumptions.

Let us discuss some conditions ensuring the fulfilment of the normal semicontinuity for moving sets; they automatically generate the corresponding conditions for the subdifferential semicontinuity of extended-real-valued functions and for the coderivative semicontinuity of set-valued mappings due to the above definitions.

First, observe that these properties always hold for sets, functions, and set-valued mappings not depending on the parameter t, which corresponds to optimal control problems for *autonomous* systems. Also, it is easily implied by the definitions that $\Omega(\cdot)$ is normally semicontinuous at $(\bar{x},\bar{t})$ if $\Omega(t)-f(t)$ is a constant set near $\bar{t}$ for some single-valued continuous mapping f. The next useful sufficient conditions for the normal semicontinuity of moving sets were given in Mordukhovich (1984, 1988); see also Proposition 4.4 in Mordukhovich, Treiman and Zhu (2002).

PROPOSITION 3.1 *$\Omega(\cdot)$ is normally semicontinuous at $(\bar{x},\bar{t})$ if it is convex-valued near $\bar{t}$ and inner/lower semicontinuous at this point, i.e.,*

$$\Omega(\bar{t})\subset\big\{x\in I\!\!R^n\big|\ \forall t_k\to\bar{t}\ \exists x_k\to\bar{x}\ \ \text{with}\ \ x_k\in\Omega(t_k),\ k\in I\!\!N\big\}.$$

Recently Lionel Thibault (personal communication) obtained more general sufficient conditions for the normal semicontinuity of moving sets. In particular, he proved this property for inner semicontinuous $\Omega(\cdot)$ whose images are *uniformly prox-regular* near reference points in the sense of Poliquin, Rockafellar and Thibault (2000).

Observe that the extended normal subdifferential, and coderivative constructions from Definition 3.1 satisfy the the following *robustness* property important for performing limiting procedures. For brevity, we present this property only in the case of moving sets.

PROPOSITION 3.2 *Let $\Omega\colon T\rightrightarrows I\!\!R^n$ with $\bar{x}\in\Omega(\bar{t})$. Then one has*

$$\widetilde{N}(\bar{x};\Omega(\bar{t}))=\mathop{\rm Lim\,sup}_{(t,x)\overset{\text{gph}\Omega}{\to}(\bar{t},\bar{x})}\widetilde{N}(x;\Omega(t)).\tag{28}$$

Proof. It is sufficient to prove the inclusion "$\supset$" in (28), since the opposite one is obvious. Taking x^* from the right-hand side of (28), we find sequences (t_k, x_k, x_k^*) satisfying

$$t_k \to \bar{t},\ x_k \to \bar{x},\ x_k^* \to x^*$$
$$\text{as } k \to \infty \text{ with } x_k \in \Omega(t_k) \text{ and } x_k^* \in \widetilde{N}(x_k; \Omega(t_k)),\ k \in I\!N.$$

By the construction of $\widetilde{N}$ in (25), for each $k \in I\!N$ there are sequences $(t_{km}, x_{km}, x_{km}^*)$ such that

$$t_{km} \to t_k,\ x_{km} \to x_k,\ x_{km}^* \to x_k^*$$
$$\text{as } m \to \infty \text{ with } x_{km} \in \Omega(t_{km}) \text{ and } x_{km}^* \in \widehat{N}(x_{km}; \Omega(t_{km}))$$

for all $m \in I\!N$. Employing the diagonal process, we construct sequences (t_m, x_m, x_m^*) satisfying

$$t_m \to \bar{t},\ x_m \to \bar{x},\ x_m^* \to x^*$$
$$\text{as } m \to \infty \text{ with } x_m \in \Omega(t_m) \text{ and } x_m^* \in \widehat{N}(x_m; \Omega(t_m)),\ m \in I\!N.$$

The latter yields $x^* \in \widetilde{N}(\bar{x}; \Omega(\bar{t}))$ and completes the proof of the proposition. ■

Note that the extended constructions from Definition 3.1 enjoy a full generalized differential calculus similar to the basic constructions (21), (23), and (24). This can be derived similarly to the latter ones, e.g., by using a fuzzy calculus for Fréchet-like preliminary objects. We are not going to use such a calculus in the present paper.

4. Necessary optimality conditions for delay-difference inclusions

The objective of this section is to obtain necessary conditions for optimal solutions to discrete approximations problems (P_N) governed by delay-difference inclusions. We derive new necessary optimality conditions in the extended Euler-Lagrange form employing the basic generalized differential constructions (21), (23), and (24). The results obtained do not require any restrictive assumptions on the initial data. In particular, we do not impose either convexity assumptions like in (H6) or the Lipschitz continuity of F like in (H1).

Our approach is based on reducing the dynamic optimization problems (P_N) for each $N \in I\!N$ to a static mathematical programming problem (MP) with *many geometric constraints* given by sets with possibly empty interiors, the number of which tends to infinity together with the approximating parameter $N \to \infty$. This makes problems (MP) to be intrinsically *nonsmooth*, even in the case of smooth cost functions and endpoint constraints.

The general structure of problems (MP) is as follows:

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

where $\phi_j : I\!R^d \to I\!R$, $g_j : I\!R^d \to I\!R^n$, and $\Lambda_j \subset I\!R^d$. For our applications in this paper we need the following version of the *generalized Lagrange multiplier rule* taken from Corollary 7.5.1 in Mordukhovich (1988):

PROPOSITION 4.1 *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 $\{\mu_j \mid j = 0, \dots, r\}$ as well as vectors $\{\psi_j \in I\!R^n \mid j = 0, \dots, m\}$ and $\{z_j^* \in I\!R^d \mid j = 0, \dots, l\}$, not all zero, such that*

$$\mu_j \ge 0 \quad for \quad j = 0, \dots, r, \tag{29}$$

$$\mu_j \phi_j(\bar{z}) = 0 \quad for \quad j = 1, \dots, r, \tag{30}$$

$$z_j^* \in N(\bar{z}; \Lambda_j) \quad for \quad j = 0, \dots, l, \tag{31}$$

$$-\sum_{j=0}^{l} z_j^* \in \partial\Big(\sum_{j=0}^{r} \mu_j \phi_j\Big)(\bar{z}) + \sum_{j=0}^{m} \nabla g_j(\bar{z})^* \psi_j. \tag{32}$$

Now we employ Proposition 4.1 and calculus rules for the generalized differential constructions used therein to derive the necessary optimality conditions for discrete approximation problems (P_N). Fix $N \in I\!N$ and consider the "long" vector z defined by

$$\begin{aligned} z &= (x_{-N}^N, \dots, x_{k+1}^N, y_0^N, \dots, y_k^N) \\ &:= (x^N(t_{-N}), \dots, x^N(t_{k+1}), y^N(t_0), \dots, y^N(t_k)). \end{aligned}$$

Then the discrete approximation problem (P_N) can be reduced to the above problem (MP) with

$$\begin{aligned} \phi_0(z) := \;& \varphi(x_0^N, x_{k+1}^N) + |x_0^N - \bar{x}(a)|^2 + \sum_{j=-N}^{-1} \int_{t_j}^{t_{j+1}} \left| x_j^N - \bar{x}(t) \right|^2 dt \\ & + h_N \sum_{j=0}^{k} f(x_j^N, x_{j-N}^N, y_j^N, t_j) + \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} \left| y_j^N - \dot{\bar{x}}(t) \right|^2 dt, \end{aligned} \tag{33}$$

$$\phi_j(z) := |x_j^N - \bar{x}(t_j)| - \varepsilon, \quad j = 1, \dots, k+1, \tag{34}$$

$$g_j(z) := x_{j+1}^N - x_j^N - h_N y_j^N, \quad j = 0, \dots, k, \tag{35}$$

$$\Lambda_j := \big\{(x^N_{-N},\ldots,y^N_k)\big|\ x^N_j \in C(t_j)\big\},\quad j=-N,\ldots,-1, \tag{36}$$

$$\Lambda_j := \big\{(x^N_{-N},\ldots,y^N_k)\big|\ y^N_j \in F(x^N_j, x^N_{j-N}, t_j)\big\},\quad j=0,\ldots,k, \tag{37}$$

$$\Lambda_{k+1} := \big\{(x^N_{-N},\ldots,y^N_k)\big|\ (x^N_0, x^N_{k+1}) \in \Omega_N\big\}. \tag{38}$$

Let $\bar z^N = (\bar x^N_{-N},\ldots,\bar x^N_{k+1},\bar y^N_0,\ldots,\bar y^N_k)$ be an optimal solution to problem (MP) with the data (33)–(38), for each fixed $N \in I\!\!N$. Employing Proposition 4.1, we find real numbers $\{\mu^N_j|\ j=0,\ldots,k+1\}$ as well as vectors $\{\psi^N_j \in I\!\!R^n|\ j=0,\ldots,k\}$ and $\{z^*_j \in I\!\!R^{n(2k+N+3)}|\ j=-N,\ldots,k+1\}$, not all zero, such that conditions (29)–(32) are satisfied.
Taking $z^*_j = (x^*_{-N,j},\ldots,x^*_{k+1,j},y^*_{0,j},\ldots,y^*_{k,j}) \in N(\bar z^N;\Lambda_j)$ for $j=-N,\ldots,-1$, we observe from the structure of Λ_j that all but one components of z^*_j are zero with the remaining one satisfying

$$x^*_{j,j} \in N(\bar x^N_j; C(t_j)),\quad j=-N,\ldots,-1. \tag{39}$$

Similarly, the conditions $z^*_j \in N(\bar z^N;\Lambda_j)$ for $j=0,\ldots,k$ and $z^*_{k+1} \in N(\bar z^N;\Lambda_{k+1})$ are equivalent, respectively, to

$$(x^*_{j,j},\ x^*_{j-N,j},\ y^*_{j,j}) \in N((\bar x^N_j, \bar x^N_{j-N}, \bar y^N_j); \mathrm{gph}\, F_j),\ j=0,\ldots,k,\ \text{ and} \tag{40}$$

$$(x^*_{0,k+1},\ x^*_{k+1,k+1}) \in N((\bar x^N_0, \bar x^N_{k+1}); \Omega_N) \tag{41}$$

with all the other components of z^*_j, $j=0,\ldots,k+1$, equal to zero.

By Theorem 2.2 on the convergence of discrete approximations, we conclude that $\phi_j(\bar z^N) < 0$ for $j=1,\ldots,k+1$ when N is sufficiently large. Thus, $\mu^N_j = 0$ for these indexes j due to the complementary slackness conditions (30). Denote by $\lambda^N \ge 0$ the remaining multiplier μ^N_0 from Proposition 4.1. Further, employing the subdifferential sum rule for ϕ_0 in (33), one obtains

$$\begin{aligned}\partial\phi_0(\bar z^N) \quad &\subset \partial\varphi(\bar x^N_0, \bar x^N_{k+1}) + 2(\bar x^N_0 - \bar x(a)) + \sum_{j=-N}^{-1}\int_{t_j}^{t_{j+1}} 2\big(\bar x^N_j - \bar x(t)\big)dt\\ &+h_N\sum_{j=0}^{k}\partial f(\bar x^N_j, \bar x^N_{j-N}, \bar y^N_j, t_j) + \sum_{j=0}^{k}\int_{t_j}^{t_{j+1}} 2\big(\bar y^N_j - \dot{\bar x}(t)\big)dt,\end{aligned} \tag{42}$$

where ∂f stands in (42) and in what follows for the basic subdifferential of f with respect to the first three variables.

One can easily see that

$$\begin{aligned}&\sum_{j=0}^{k}\nabla g_j(\bar z^N)^*\psi_j \\ &= (0,\ldots 0, -\psi^N_0, \psi^N_0 - \psi^N_1, \ldots, \psi^N_{k-1} - \psi^N_k, \psi^N_k, -h_N\psi^N_0, \ldots, -h_N\psi^N_k).\end{aligned} \tag{43}$$

Combining (32) with (39)–(43), we have the following relationships:

$$
\begin{aligned}
-x^*_{j,j} - x^*_{j,j+N} &= \lambda^N h_N \kappa^N_j + \lambda^N \sigma^N_j,\ j = -N,\ldots,-1, && (44)\\
-x^*_{j,j} - x^*_{j,j+N} &= \lambda^N h_N \kappa^N_j + \lambda^N h_N v^N_j + \psi^N_{j-1} - \psi^N_j, && (45)\\
& \quad j = 1,\ldots,k-N,\\
-x^*_{j,j} &= \lambda^N h_N v^N_j + \psi^N_{j-1} - \psi^N_j,\ j = k-N+1,\ldots,k, && (46)\\
-y^*_{j,j} &= \lambda^N h_N \omega^N_j + \lambda^N \theta^N_j - h_N \psi^N_j,\ j = 0,\ldots,k, && (47)\\
-x^*_{k+1,k+1} &= \lambda^N u^N_{k+1} + \psi^N_k, && (48)\\
-x^*_{0,0} - x^*_{0,k+1} &= \lambda^N u^N_0 + \lambda^N h_N \kappa^N_0 + 2\lambda^N(\bar{x}^N_0 - \bar{x}(a)) && (49)\\
& \quad + \lambda^N h_N v^N_0 - \psi_0
\end{aligned}
$$

with the notation

$$
(u^N_0, u^N_{k+1}) \in \partial\varphi(\bar{x}^N_0, \bar{x}^N_{k+1}),\quad (v^N_j, \kappa^N_{j-N}, \omega^N_j) \in \partial f(\bar{x}^N_j, \bar{x}^N_{j-N}, \bar{y}^N_j, t_j),
$$

$$
\theta^N_j := 2\int_{t_j}^{t_{j+1}} \big(\bar{y}^N_j - \dot{\bar{x}}(t)\big)dt,\quad \sigma^N_j := 2\int_{t_j}^{t_{j+1}} \big(\bar{x}^N_j - \bar{x}(t)\big)dt.
$$

The next theorem gives necessary optimality conditions for discrete approximation problems (P_N) governed by constrained delay-difference inclusions.

THEOREM 4.1 *Let $\bar{z}^N$ be an optimal solution to problem (P_N), where $F_j := F(\cdot,\cdot,t_j)$. Assume that the sets $\operatorname{gph} F_j$ are closed and the functions φ and f_j are Lipschitz continuous around $(\bar{x}^N_0, \bar{x}^N_{k+1})$ and $(\bar{x}^N_j, \bar{x}^N_{j-N}, \bar{y}^N_j)$, respectively, for all $j = 0,\ldots,k$. Then there exist $\lambda^N \ge 0$, p^N_j $(j = 0,\ldots,k+1)$, and q^N_j $(j = -N,\ldots,k+1)$, not all zero, such that*

$$
\begin{aligned}
&\Big(\frac{p^N_{j+1} - p^N_j}{h_N}, \frac{q^N_{j-N+1} - q^N_{j-N}}{h_N}, -\frac{\lambda^N \theta^N_j}{h_N} + p^N_{j+1} + q^N_{j+1}\Big)\\
&\in \lambda^N \partial f(\bar{x}^N_j, \bar{x}^N_{j-N}, \bar{y}^N_j, t_j) + N((\bar{x}^N_j, \bar{x}^N_{j-N}, \bar{y}^N_j); \operatorname{gph} F_j),\ j = 0,\ldots,k,
\end{aligned} \tag{50}
$$

$$
-\frac{q^N_{j+1} - q^N_j}{h_N} - \lambda^N \frac{\sigma^N_j}{h_N} \in N(\bar{x}^N_j; C(t_j)),\ j = -N,\ldots,-1, \tag{51}
$$

$$
q^N_j = 0,\ j = k-N+1,\ldots,k+1, \tag{52}
$$

$$
(p^N_0 + q^N_0, -p^N_{k+1}) \in \lambda^N \partial\varphi(\bar{x}^N_0, \bar{x}^N_{k+1}) + N((\bar{x}^N_0, \bar{x}^N_{k+1}); \Omega_N). \tag{53}
$$

Proof. Consider first $\widetilde{p}^N_j := \psi^N_{j-1}$ for $j = 1,\ldots,k+1$, $\widetilde{q}^N_j := \lambda^N \kappa^N_j + x^*_{j,j+N}/h_N$ for $j = -N,\ldots,k-N$, and $\widetilde{q}^N_j := 0$ for $j = k-N+1,\ldots,k+1$. Then let $q^N_{k+1} := 0$ and define $q^N_j := q^N_{k+1} - \tilde{q}^N_j h_N$ for $j = -N,\ldots,k+1$. It is easy to check that $q^N_j = 0$ for $j = k-N+1,\ldots,k+1$. Finally, we define

$$
p^N_0 := \lambda^N u^N_0 + x^*_{0,k+1} - q^N_0 \quad\text{and}\quad p^N_j := \tilde{p}^N_j - q^N_j h_N \ \text{ for } j = 1,\ldots,k+1.
$$

Then (50) follows from (46)–(47), (51) comes from (44), and (53) follows from (48) and (50). This completes the proof of the theorem. ■

Corollary 4.1 *In addition to the assumptions of Theorem 4.1, 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 (50)–(53) and $\lambda^N \geq 0$ hold with $(\lambda^N, p_{k+1}^N) \neq 0$, i.e., one can set*

$$(\lambda^N)^2 + |p_{k+1}^N|^2 = 1. \tag{54}$$

Proof. If $\lambda^N = 0$, then (50) implies that

$$\begin{aligned}&\Big(\frac{p_{j+1}^N - p_j^N}{h_N}, \frac{q_{j-N+1}^N - q_{j-N}^N}{h_N}\Big)\\ &\in D^*F_j(\bar{x}_j^N, \bar{x}_{j-N}^N, \bar{y}_j^N)(-p_{j+1}^N - q_{j+1}^N),\ j = 0, \ldots, k,\end{aligned} \tag{55}$$

by the coderivative definition (24). Set $j = k$. If we assume that $p_{k+1}^N = 0$, then (52) and (55) give the inclusion

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

The latter yields $p_k^N = q_{k-N}^N = 0$ due to the coderivative characterization of the local Lipschitzian property from Theorem 5.11 in Mordukhovich (1993). By repeating the above procedure along (55), we conclude that $p_j^N = 0$ for all $j = 0, \ldots, k+1$ and $q_j^N = 0$ for all $j = -N, \ldots, k+1$. This contradicts the non-triviality assertion of Theorem 4.1 and completes the proof of this corollary. ■

5. Optimality conditions for delay-differential inclusions

Now we come back to the original Bolza problem (P) for delay-differential inclusions and establish the necessary optimality conditions for (P) in the extended Euler-Lagrange form involving the generalized differential constructions of Section 3. Let us keep the assumptions (H1)–(H3), but instead of (H4) and (H5) we impose their following modifications:

(H4') φ is Lipschitz continuous on $U \times U$, $f(x, y, v, \cdot)$ is continuous for a.e. $t \in [a, b]$ and bounded uniformly in $(x, y, v) \in U \times U \times m_F I\!B$, and Ω is locally closed around $(\bar{x}(a), \bar{x}(b))$.

(H5') There are positive numbers μ and l_f such that $f(\cdot, \cdot, \cdot, t)$ is Lipschitz continuous on the set $A_\mu(t)$ from (H5) with the constant l_f.

In the results of this section the subdifferential, normal, and coderivative symbols are used with respect to all variables but t.

THEOREM 5.1 *Let $\bar{x}(\cdot)$ be an optimal solution to the Bolza problem (P) under hypotheses (H1)–(H3), (H4'), (H5'), and (H6). Assume also that problem (P) is stable with respect to relaxation. Then there exist a number $\lambda \geq 0$ as well as absolutely continuous functions $p\colon [a,b] \to I\!R^n$ and $q\colon [a-\Delta, b] \to I\!R^n$ satisfying the conditions:*

$$\begin{aligned}\big(\dot{p}(t), \dot{q}(t-\Delta)\big) \quad &\in \mathrm{co}\Big\{(u,w)\Big|\ \big(u,w,p(t)+q(t)\big)\\ &\in \lambda\widetilde{\partial} f(\bar{x}(t), \bar{x}(t-\Delta), \dot{\bar{x}}(t), t)\\ &+\widetilde{N}((\bar{x}(t), \bar{x}(t-\Delta), \dot{\bar{x}}(t)); \mathrm{gph}\, F(\cdot,\cdot,t))\Big\} \quad a.e. \quad t \in [a,b],\end{aligned} \tag{56}$$

$$\big\langle \dot{q}(t), \bar{x}(t)\big\rangle = \min_{c \in C(t)} \big\langle \dot{q}(t), c\big\rangle \qquad a.e. \quad t \in [a-\Delta, a), \tag{57}$$

$$(p(a)+q(a), -p(b)) \in \lambda\partial\varphi(\bar{x}(a), \bar{x}(b)) + N((\bar{x}(a), \bar{x}(b)); \Omega), \tag{58}$$

$$q(t) = 0, \quad t \in [b-\Delta, b], \tag{59}$$

$$\lambda + |p(b)| > 0. \tag{60}$$

Proof. We are going to prove this theorem by the method of discrete approximations and first construct a sequence of finite-dimensional problems (P_N) whose solutions $\bar{x}^N = (\bar{x}^N_{-N}, \ldots, \bar{x}^N_{k+1})$ strongly approximate $\bar{x}(\cdot)$ in the sense of Theorem 2.1. By employing Corollary 4.1 to $\bar{x}^N$, we find $\lambda^N \geq 0$ and p^N_j, q^N_j satisfying relationships (50)–(54).

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

Define $\theta^N(t) := \theta^N_j/h_N$ for $t \in [t_j, t_{j+1})$, $j = 0, \ldots, k$ and conclude by Theorem 2.1 that

$$\begin{aligned}&\int_a^b |\theta^N(t)|dt = \sum_{j=0}^{k} |\theta^N_j|\\ &\leq 2\sum_{j=0}^{k}\int_{t_j}^{t_{j+1}} \Big|\dot{\bar{x}}(t) - \frac{\bar{x}^N_{j+1} - \bar{x}^N_j}{h_N}\Big|dt = 2\int_a^b |\dot{\bar{x}}(t) - \dot{\bar{x}}^N(t)|dt \to 0.\end{aligned}$$

Similarly, by letting $\sigma^N(t) := \sigma^N_j/h_N$ for $t \in [t_j, t_{j+1})$, $j = -N, \ldots, -1$, one obtains

$$\begin{aligned}&\int_{a-\Delta}^a |\sigma^N(t)|dt = \sum_{j=-N}^{-1} |\sigma^N_j|\\ &\leq 2\sum_{j=-N}^{-1}\int_{t_j}^{t_{j+1}} \big|\bar{x}(t) - \bar{x}^N_j\big|dt = 2\int_{a-\Delta}^a |\bar{x}(t) - \bar{x}^N(t)|dt \to 0.\end{aligned}$$

Since the convergence in $L^1(T)$ of a sequence of functions defined on some interval T implies the convergence of its subsequence almost everywhere on T, we suppose with no restriction that $\dot{\bar{x}}^N(t) \to \dot{\bar{x}}(t)$, $\theta^N(t) \to 0$, and $\sigma^N(t) \to 0$ as $N \to \infty$ a.e. on the corresponding intervals. Such a *pointwise* convergence is important in what follows.

Let us estimate $(p^N(t), q^N(t-\Delta))$ for large N. It follows from (50) that for all $j = 0, \dots, k$ one has the inclusions

$$\Big(\frac{p^N_{j+1} - p^N_j}{h_N} - \lambda^N v^N_j, \frac{q^N_{j-N+1} - q^N_{j-N}}{h_N} - \lambda^N \kappa^N_{j-N},$$
$$-\frac{\lambda^N \theta^N_j}{h_N} + p^N_{j+1} + q^N_{j+1} - \lambda^N \omega^N_j\Big) \in N((\bar{x}^N_j, \bar{x}^N_{j-N}, \bar{y}^N_j); \operatorname{gph} F_j)$$
with some $(v^N_j, \kappa^N_{j-N}, \omega^N_j) \in \partial f(\bar{x}^N_j, \bar{x}^N_{j-N}, \bar{y}^N_j, t_j)$.

This implies by (24) that

$$\Big(\frac{p^N_{j+1} - p^N_j}{h_N} - \lambda^N v^N_j, \frac{q^N_{j-N+1} - q^N_{j-N}}{h_N} - \lambda^N \kappa^N_{j-N}\Big)$$
$$\in D^* F_j(\bar{x}^N_j, \bar{x}^N_{j-N}, \bar{y}^N_j)\Big(\lambda^N \omega^N_j + \frac{\lambda^N \theta^N_j}{h_N} - p^N_{j+1} - q^N_{j+1}\Big).$$

Using again Theorem 5.1 from Mordukhovich (1993) providing coderivative characterizations of the Lipschitz continuity for F_j, we get the estimate

$$\Big|\Big(\frac{p^N_{j+1} - p^N_j}{h_N} - \lambda^N v^N_j, \frac{q^N_{j-N+1} - q^N_{j-N}}{h_N} - \lambda^N \kappa^N_{j-N}\Big)\Big|$$
$$\le L_F \Big|\lambda^N \omega^N_j + \frac{\lambda^N \theta^N_j}{h_N} - p^N_{j+1} - q^N_{j+1}\Big|.$$

Since $|(v^N_j, \kappa^N_{j-N}, \omega^N_j)| \le l_f$ due to the Lipschitz continuity of f with modulus l_f, one derives from the above that

$$|(p^N_j, q^N_{j-N})| \le L_F |\lambda^N \theta^N_j| + L_F \lambda^N h_N |\omega^N_j|$$
$$+ L_F h_N |p^N_{j+1} + q^N_{j+1}| + |(p^N_{j+1}, q^N_{j-N+1})| + \lambda^N h_N |(v^N_j, \kappa^N_{j-N})| \le L_F |\theta^N_j|$$
$$+ (L_F h_N + h_N) l_f + L_F h_N |p^N_{j+1} + q^N_{j+1}| + |(p^N_{j+1}, q^N_{j-N+1})|, \ j = 0, \dots, k,$$

and taking (52) into account, that

$$|(p^N_j, q^N_{j-N})| \le L_F |\theta^N_j| + (L_F + 1) h_N l_f + (L_F h_N + 1) |(p^N_{j+1}, q^N_{j-N+1})|$$
$$\le L_F |\theta^N_j| + (L_F h_N + 1) L_F |\theta^N_{j+1}| + (L_F + 1) h_N l_f$$
$$+ (L_F h_N + 1)(L_F + 1) h_N l_f + (L_F h_N + 1)^2 |(p^N_{j+2}, q^N_{j-N+2})| \le \cdots$$
$$\le \exp[L_F(b-a)](1 + l_f(L_F + 1)/L_F + L_F \nu_N), \quad j = 0, \dots, k+1,$$

where $\nu_N := \int_a^b |\dot{\bar{x}}(t) - \dot{\bar{x}}^N(t)|dt \to 0$ as $N \to \infty$. This implies the uniform boundedness of $\{p_j^N, q_{j-N}^N\}$ and hence of $\{(p^N(t), q^N(t-\Delta))\}$ on $[a,b]$.

To estimate $(\dot{p}^N(t), \dot{q}^N(t-\Delta))$, we have

$$\begin{aligned}
&|(\dot{p}^N(t), \dot{q}^N(t-\Delta))| = \Big|\Big(\frac{p_{j+1}^N - p_j^N}{h_N}, \frac{q_{j-N+1}^N - q_{j-N}^N}{h_N}\Big)\Big| \\
&\le L_F\Big|\lambda^N\omega_j^N + \frac{\lambda^N\theta_j^N}{h_N} - p_{j+1}^N - q_{j+1}^N\Big| + \lambda^N|(v_j^N, v_{j-N}^N)| \\
&\le L_F l_f + l_f + L_F(|\theta^N(t)| + |p_{j+1}^N| + |q_{j+1}^N|), \quad t \in [t_j, t_{j+1}),
\end{aligned}$$

which implies the (essential) uniform boundedness of $\{\dot{p}^N(t), \dot{q}^N(t-\Delta)\}$ on $[a,b]$. By standard compactness results of real analysis we find absolutely continuous functions $p(\cdot)$ and $q(\cdot - \Delta)$ with

$$\begin{aligned}
&p^N(t) \to p(t), \quad q^N(t-\Delta) \to q(t-\Delta) \quad \text{uniformly on } [a,b], \\
&\dot{p}^N(t) \to \dot{p}(t), \quad \dot{q}^N(t-\Delta) \to \dot{q}(t-\Delta) \quad \text{weakly in } L^2[a,b] \text{ as } N \to \infty.
\end{aligned}$$

It is easy to observe that the discrete Euler-Lagrange inclusion (50) can be rewritten as

$$\begin{aligned}
&(\dot{p}^N(t), \dot{q}^N(t-\Delta)) \in \Big\{(u,v)\Big|\ \Big(u, v, p^N(t_{j+1}) + q^N(t_{j+1}) - \frac{\lambda^N\theta_j^N}{h_N}\Big) \\
&\in \lambda^N\partial f(\bar{x}(t_j), \bar{x}(t_j-\Delta), \dot{\bar{x}}^N(t_j), t_j) + N((\bar{x}^N(t_j), \bar{x}^N(t_j-\Delta), \dot{\bar{x}}^N(t_j)); \\
&\operatorname{gph} F(\cdot,\cdot,t_j)\Big\}
\end{aligned} \tag{61}$$

for all $t \in [t_j, t_{j+1})$ and all $j = 0, \ldots, k$.

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 to the limit in (61) as $N \to \infty$ and taking into account the construction of the extended normal cone and subdifferential in Definition 3.1 as well as their robustness property from Proposition 3.2, we arrive at the *extended Euler-Lagrange inclusion* (56). To justify the *tail condition* (57), we pass to the limit in (51) with the use of the specific form of the normal cone to convex sets (22) as well as Proposition 3.1 whose assumptions are satisfied for $C(\cdot)$ due to (H3) and (H6). Finally, conditions (58)–(60) follow directly from (52)–(54), which completes the proof of the theorem. ■

For the *Mayer problem* (P_M), that is, (1)–(4) with $f = 0$, the extended Euler-Lagrange condition (56) is equivalently expressed via the extended coderivative (27) with respect to the first two variables of the multifunction $F = F(x,y,t)$, i.e.,

$$\big(\dot{p}(t), \dot{q}(t-\Delta)\big) \in \operatorname{co} \widetilde{D}^*_{x,y}F\big(\bar{x}(t), \bar{x}(t-\Delta), \dot{\bar{x}}(t), t\big)\big(-p(t) - q(t)\big) \tag{62}$$

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

One can replace $\widetilde{D}^*_{x,y}$ by the basic coderivative (24) with respect to (x,y) if F is coderivatively semicontinuous at $(\bar{x}(t), \bar{x}(t-\Delta), \dot{\bar{x}}(t), t)$ with respect to t almost everywhere on $[a,b]$.

It happens that the extended Euler-Lagrange condition obtained above implies, under the relaxation stability of the original problems, two principal optimality conditions expressed in terms of the classical Hamiltonian. In the following corollary we consider for simplicity the case of the Mayer problem (P_M) for autonomous delay-differential inclusions. Then, the *Hamiltonian* function for (2) is given by

$$H(x,y,p) := \sup\big\{\langle p, v\rangle \big|\ v \in F(x,y)\big\}. \tag{63}$$

COROLLARY 5.1 *Let $\bar{x}(\cdot)$ be an optimal solution to the Mayer problem (P_M) for the autonomous delay-differential inclusion (2) under assumptions (H1), (H3), (H4'), and (H6). Suppose that the problem (P_M) is stable with respect to relaxation. Then there exist a number $\lambda \geq 0$ and the absolutely continuous functions $p\colon [a,b] \to I\!R^n$ and $q\colon [a-\Delta, b] \to I\!R^n$ satisfying conditions (57)–(60) as well as the Hamiltonian inclusion*

$$\begin{aligned}&\big(\dot{p}(t), \dot{q}(t-\Delta)\big)\\ &\in \operatorname{co}\big\{(u,w)\big|\ \big(-u, -w, \dot{\bar{x}}(t)\big) \in \partial H\big(\bar{x}(t), \bar{x}(t-\Delta), p(t)+q(t)\big)\big\}\end{aligned} \tag{64}$$

and the maximum condition

$$\langle p(t)+q(t)\rangle, \dot{\bar{x}}(t)\rangle = \max\big\{\langle p(t)+q(t), v\rangle \big|\ v \in F(\bar{x}(t), \bar{x}(t-\Delta))\big\} \tag{65}$$

for almost all $t \in [a,b]$. If, moreover, F is convex-valued around $(\bar{x}(t), \bar{x}(t-\Delta))$, then (64) is equivalent to the Euler-Lagrange inclusion

$$\begin{aligned}&\big(\dot{p}(t), \dot{q}(t-\Delta)\big) \in \operatorname{co} D^*F\big(\bar{x}(t), \bar{x}(t-\Delta), \dot{\bar{x}}(t)\big)\big(-p(t)-q(t)\big)\\ &\quad \text{a.e. } t \in [a,b],\end{aligned} \tag{66}$$

which automatically implies the maximum condition (65) 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 with respect to (P_M) is that the delay-differential inclusion (2) is replaced by its convexification (17). Due to Theorem 5.1 the optimal solution $\bar{x}(\cdot)$ satisfies conditions (57)–(60) and the relaxed counterpart of (62), that is the same as (66) in this case, with F replaced by $\operatorname{co} F$. According to Theorem 3.3 in Rockafellar (1996), one has

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

where H_R stands for the Hamiltonian (63) of the relaxed system, i.e., with F replaced by $\operatorname{co} F$. It is easy to see that $H_R = H$. Thus the Euler-Lagrange inclusion for the relaxed system implies the Hamiltonian inclusion (64), which surely yields the maximum condition (65). When F is convex-valued, (64) and (66) are equivalent due to the mentioned result of Rockafellar (1996). This completes the proof of the corollary. ■

Note that the Hamiltonian inclusion (64) obviously improves over following one

$$\big(-\dot{p}(t), -\dot{q}(t-\Delta), \dot{\bar{x}}(t)\big) \in \operatorname{co} \partial H\big(\bar{x}(t), \bar{x}(t-\Delta), p(t)+q(t)\big) \quad \text{a.e.} \quad t \in [a,b] \tag{67}$$

obtained by Clarke and Watkins (1986) for the Mayer problem (P_M) with convex velocities in (2) and no endpoint constraints (4) ,but without imposing the convexity hypothesis (H6).

References

Aubin, J.-P. and Cellina, A. (1984) *Differential Inclusions*. Springer, Berlin.

Bellaassali, S. and Jourani, A. (2002) Necessary optimality conditions in multiobjective dynamic optimization. To appear in *SIAM J. Control Optim.*

Clarke, F.H. (1983) *Optimization and Nonsmooth Analysis*. Wiley, New York.

Clarke, F.H. and Watkins, G.G. (1986) Necessary conditions, controllability and the value function for differential–difference inclusions. *Nonlinear Anal.*, **10**, 1155–1179.

Clarke, F.H. and Wolenski, P.R. (1996) Necessary conditions for functional differential inclusions. *Appl. Math. Optim.*, **34**, 34–51.

Dontchev, A.L. (1996) Discrete approximations in optimal control. In: B. S. Mordukhovich and H. J. Sussmann, eds., Nonsmooth Analysis and Geometric Methods in Deterministic Optimal Control. *IMA Volumes in Mathematics and its Applications*, **78**, 59–80. Springer, New York.

Dontchev, A.L. and Farkhi, E.M. (1989) Error estimates for discretized differential inclusions. *Computing*, **41**, 349–358.

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

Loewen, P.D. and Rockafellar, R.T. (1997) Bolza problems with general time constraints. *SIAM J. Control Optim.*, **35**, 2050–2069.

Malanowski, K. (1979) On convergence of finite-difference approximations to control and state constrained convex optimal control problems. *Arch. Autom. Telemech.*, **24**, 319–337.

Minchenko, L.I. (1999) Necessary optimality conditions for differential–difference inclusions. *Nonlinear Anal.*, **35**, 307–322.

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

Mordukhovich, B.S. (1984) Nonsmooth analysis with nonconvex generalized differentials and adjoint mappings. *Doklady Akad Nauk BSSR*, **28**, 976–979.

Mordukhovich, B.S. (1988) *Approximation Methods in Problems of Optimization and Control.* Nauka, Moscow.

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

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

Mordukhovich, B.S. and Trubnik, R. (2001) Stability of discrete approximation and necessary optimality conditions for delay–differential inclusions. *Annals Oper. Res.*, **101**, 149–170.

Mordukhovich, B.S., Treiman, J.S. and Zhu, Q.J. (2002) An extended extremal principle with applications to multiobjective optimization. To appear in *SIAM J. Control Optim.*

Poliquin, R.A., Rockafellar, R.T. and Thibault, L. (2000) Local differentiability of distance functions. *Trans. Amer. Math. Soc.*, **332**, 5231–5249.

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

Rockafellar, R.T. and Wets, R. J.-B. (1998) *Variational Analysis.* Sprin--ger, Berlin.

Smirnov, G.V. (2002) *Introduction to the Theory of Differential Inclusions.* American Mathematical Society, Providence, R.I.

Sussmann, H.J. (2000) New theories of set-valued differentials and new versions of the maximum principle in optimal control theory. In: A. Isidori et al., eds. *Nonlinear Control in the Year 2000.* Springer, Berlin, 487–472.

Vinter, R. (2000) *Optimal Control.* Birkhäuser, Boston.

Warga, J. (1972) *Optimal Control of Differential and Functional Equations.* Academic Press, New York.

Zhu, Q.J. (1996) Necessary optimality conditions for nonconvex differential inclusions with endpoint constraints. *J. Diff. Equ.*, **124**, 186–204.

Zhu, Q.J. (2002) Hamiltonian necessary conditions for a multiobjective optimal control problem with endpoint constraints. *SIAM J. Control Optim.*, **39**, 97–112.