Proceedings of the
44th IEEE Conference on Decision and Control, and
the European Control Conference 2005
Seville, Spain, December 12-15, 2005

On Stability of the Pontryagin Maximum Principle with respect to Time Discretization

Boris S. Mordukhovich and Ilya Shvartsman

***Abstract*— The paper deals with optimal control problems for dynamic systems governed by a parametric family of discrete approximations of control systems with continuous time, wherein the discretization step tends to zero. Discrete approximations play an important role in both qualitative and numerical aspects of optimal control and occupy an intermediate position between discrete-time and continuous-time control systems. The central result in optimal control of discrete approximations is the Approximate Maximum Principle (AMP), which is justified for smooth control problems with endpoint constraints under certain assumptions without imposing any convexity, in contrast to discrete systems with a fixed step. We show that these assumptions are essential for the validity of the AMP, and that the AMP does not hold in its expected (lower) subdifferential form for nonsmooth problems. Moreover, a new upper subdifferential form of the AMP is established for both ordinary and time-delay control systems. This solves a long-standing question about the possibility to extend the AMP to nonsmooth control problems.**

I. INTRODUCTION AND PRELIMINARIES

This paper is devoted to discrete approximations of continuous-time control systems that, viewed as a *parametric process* with a decreasing discretization step, occupy an *intermediate* position between control systems with discrete and continuous times. As the basic model for our study, we consider discrete approximations of the following Mayer-type optimal control problem governed by ordinary differential equations with endpoint constraints:

$$(P) \qquad \begin{cases} \text{minimize } J(x,u) := \varphi_0(x(t_1)) \\ \text{subject to} \\ \dot{x}(t) = f(t,x(t),u(t)) \ \text{ a.e. } \ t \in [t_0,t_1], \\ x(t_0) = x_0 \in I\!R^n, \\ u(t) \in U \ \text{ a.e. } \ t \in [t_0,t_1], \\ \varphi_i(x(t_1)) \le 0, \quad i = 1,\ldots,m, \\ \varphi_i(x(t_1)) = 0, \quad i = m+1,\ldots,m+r, \end{cases}$$

over measurable controls $u(\cdot)$ and absolutely continuous trajectories $x(\cdot)$ on the fixed time interval $T := [t_0, t_1]$. It is well known that many other control problems (of Lagrange and Bolza types, with integral constraints, on variable time intervals, etc.) can be reduced to the form of (P). The results of this paper can be extended to control problems with non-fixed initial vector $x(t_0)$ as well as to problems with continuously time-dependent control constraint $U = U(t)$.

In our study of the continuous-time problem (P), we use an approach of approximating the derivative $\dot{x}(t)$ by the finite-difference $\dot{x}(t) \approx (x(t+h) - x(t))/h$ as $h \to 0$. Allowing also perturbations of the endpoint constraints (which is very essential for variational stability), problem (P) is replaced in this way by the following family of discrete-time problems (P_N) with discretization step $h_N = (t_1 - t_0)/N$ depending on the natural parameter $N = 1, 2, \ldots$:

(P_N)

$$\begin{cases} \text{minimize } J(x_N,u_N) := \varphi_0(x_N(t_1)) \\ \text{subject to} \\ x_N(t+h_N) = x_N(t) + h_N f(t,x_N(t),u_N(t)), \\ x_N(t_0) = x_0 \in I\!R^n, \\ u_N(t) \in U,\ t \in T_N := \{t_0, t_0+h_N, \ldots, t_1 - h_N\}, \\ \varphi_i(x_N(t_1)) \le \gamma_{iN}, \quad i = 1,\ldots,m, \\ |\varphi_i(x_N(t_1))| \le \delta_{iN}, \quad i = m+1,\ldots,m+r, \\ h_N := \dfrac{t_1 - t_0}{N}, \quad N \in I\!N := \{1,2,\ldots\}, \end{cases}$$

where $\gamma_{iN} \to 0$ and $\delta_{iN} \downarrow 0$ as $N \to \infty$ for all i. For each fixed $N \in I\!N$ problem (P_N) is *finite-dimensional* and seems to be simpler than the continuous-time problem (P). Indeed, applying well-developed methods of finite-dimensional variational analysis, it is possible to derive necessary optimality conditions in problems (P_N) even with nonsmooth data and general dynamic constraints governed by discrete inclusions and then obtain the corresponding results for optimal control of differential inclusions by passing to the limit from discrete approximations; see [4], [6], [10] for detailed proofs and discussions. However, this approach has some limitation regarding necessary optimality conditions of the *maximum principle* type.

It is well known that the central result of the optimal control theory for continuous-time problems (P), the Pontryagin Maximum Principle (PMP) [8], holds with *no convexity* assumptions on the admissible velocity sets $f(t,x,U)$. This specific result is due to the fact that continuous-type control systems enjoy a certain *hidden convexity*, which is related to the classical Lyapounov theorem on convexity of an integral of a measurable multimap with respect to nonatomic vector measures and eventually leads the to maximum principle form. It is not surprising therefore, that an analogue of the maximum principle for discrete-time control systems *does not generally hold* without a priori convexity assumptions. This may create difficulties for applications of the PMP in numerical calculations of nonconvex continuous-time control systems, which inevitably involve finite-difference approximations via time discretization. To avoid such troubles, it is sufficient to justify not a full analogue of the PMP, with the exact maximum condition, but its *approximate* counterpart, where an error in the maximum condition *tends to zero*, when

B. S. Mordukhovich is the Department of Mathematics, Wayne State University, Detroit, MI 48202, USA boris@math.wayne.edu

I. Shvartsman is with the Department of Mathematical Sciences, University of Bath, Bath BA2 7AY, UK i.shvartsman@maths.bath.ac.uk

the discretization step decreases.

The first result of this type in the absence of convexity assumptions was given by Gabasov and Kirillova [2], [3], under the name of "quasi-maximum principle," for parametric discrete systems with smooth cost and dynamics and with no endpoint constraints. The proof of this result essentially exploited the unconstrained nature of the problem.

The following *Approximate Maximum Principle* (AMP) for the nonconvex constrained problems (P_N) was established by Mordukhovich [4], [5]. The proof in [4], [5] is based on the discovered finite-difference counterpart of the hidden convexity property and the separation theorem. Denote

$$H(t,x,p,u) := \langle p, f(t,x,u)\rangle, \quad p \in I\!R^n, \tag{1}$$

the *Hamilton-Pontryagin function* for the dynamic constraints under consideration.

Theorem 1.1: (APPROXIMATE MAXIMUM PRINCIPLE). Let the pairs $(\bar{x}_N, \bar{u}_N)$ be optimal to (P_N) for all $N \in I\!N$, where U is a compact subset of a metric space with the metric $d(\cdot,\cdot)$, f is continuous with respect to its variables and continuously differentiable with respect to x in a tube containing the optimal trajectories $\bar{x}_N(t)$ for large N, and where each φ_i is continuously differentiable around the limiting points of $\{\bar{x}_N(t_1)\}$. Impose the following assumptions:

(a) The CONSISTENCY CONDITION on the perturbation of the equality constraints and the discretization step meaning that the latter is decreasing faster than the former:

$$\lim_{N\to\infty} \frac{h_N}{\delta_{iN}} = 0 \quad \text{for all} \quad i = m+1,\ldots,m+r. \tag{2}$$

(b) The PROPERNESS of the sequences of optimal controls $\{\bar{u}_N\}$, which means that for every increasing subsequence $\{N\}$ of natural numbers and every sequence of mesh points $\tau_{\theta(N)} \in T_N$ satisfying $\tau_{\theta(N)} = t_0 + \theta(N)h_N$, $\theta(N) = 0,1,\ldots,N-1$, and $\tau_{\theta(N)} \to t \in [t_0,t_1]$ one has

$$\text{either} \qquad d\big(u_N(\tau_{\theta(N)}), u_N(\tau_{\theta(N)+q})\big) \to 0 \quad \text{or}$$
$$d\big(u_N(\tau_{\theta(N)}), u_N(\tau_{\theta(N)-q})\big) \to 0$$

as $N \to \infty$ with any natural constant q.

Then there are numbers $\{\lambda_{iN} |\ i = 0,\ldots,m+r\}$ and a function $\varepsilon(t,h_N) \downarrow 0$ as $N \to \infty$ uniformly in $t \in T_N$ such that

$$\begin{aligned} &H(t,\bar{x}_N(t),p_N(t+h_N),\bar{u}_N(t)) = \\ &\max_{u\in U} H(t,\bar{x}_N(t),p_N(t+h_N),u) - \varepsilon(t,h_N) \end{aligned} \tag{3}$$

for all $t \in T_N$ and that

$$\lambda_{iN}(\varphi_i(\bar{x}_N(t_1)) - \gamma_{iN}) = O(h_N), \quad i = 1,\ldots,m, \tag{4}$$

$$\lambda_{iN} \ge 0, \ \ i = 0,\ldots,m, \quad \text{and} \quad \sum_{i=0}^{m+r} \lambda_{iN}^2 = 1 \tag{5}$$

for all $N \in I\!N$, where $p_N(t)$, $t \in T_N \cup \{t_1\}$, is the corresponding trajectory of the adjoint system

$$p_N(t) = p_N(t+h_N) + h_N \frac{\partial H}{\partial x}(t,\bar{x}_N(t),p_N(t),\bar{u}_N(t)), \quad t \in T_N, \tag{6}$$

with the transversality condition

$$p_N(t_1) = -\sum_{i=0}^{m+r} \lambda_{iN} \nabla\varphi_i(\bar{x}_N(t_1)). \tag{7}$$

Observe that the closer h_N is to zero, the more precise the approximate maximum condition (3) and the approximate complementary slackness condition (4) are. This means that the AMP in (P_N) tends to the PMP in (P) as $N \to \infty$, which actually justifies the *stability* of the Pontryagin Maximum Principle with respect to discrete approximations under the assumptions made.

It has been shown in [4], [5] that the *consistency* condition in (a) is *essential* for the validity of the AMP in problems with equality constraints. The first goal of the paper is to examine the other two significant assumptions made in Theorem 1.1: the *properness* condition in (b) and the *smoothness* of the initial data. We show in Section 2 that *both of these assumptions are essential for the validity of the AMP*.

Note that the properness of the sequence of optimal controls in (b) is a *finite-difference counterpart* of the piecewise continuity (or, more generally, of *Lebesgue regular points* having full measure) for optimal controls in continuous-time systems. It turns out that the situation when sequences of optimal controls are not proper in discrete approximations is not unusual for systems with nonconvex velocities, and it leads to the violation of the AMP already in the case of smooth problems with inequality constraints.

The impact of nonsmoothness to the validity of the AMP happens to be even more striking: the AMP does *not hold* in the expected conventional subdifferential form already for minimizing *convex* cost functions in discrete approximations of linear systems with no endpoint constraints, as well as for problems with nonsmooth dynamics. It seems that the *AMP is one of very few results on necessary optimality conditions that do not have expected counterparts in nonsmooth settings*.

On the other hand, we derive the AMP in problems (P_N) with nonsmooth functions describing the objective and inequality constraints in a new *upper subdifferential* (or superdifferential) form, which is also new for necessary optimality conditions in continuous-time control systems. The main difference between the conventional subdifferential form, which does not hold for the AMP, but holds for the PMP, and the new one, is that the latter involves upper (not lower) subgradients of nonsmooth functions in transversality conditions. This form applies to a class of *uniformly upper subdifferentiable* functions described in this paper, which particularly contains smooth and concave continuous functions being closed with respect to taking minimum over compact sets. The results obtained solve a long-standing question about the possibility to establish the AMP in nonsmooth control problems. We also derive the upper subdifferential form of the AMP in discrete approximations of control systems with *time delays*, for which no results of this type have been known before.

The rest of the paper is organized as follows. Section 2 contains examples on the *violation of the AMP* in smooth problems (P_N) without the properness condition as well as in problems with nonsmooth cost functions and/or nonsmooth dynamics. In Section 3 we discuss appropriate tools of nonsmooth analysis. In Section 4 we formulate the AMP for the discrete approximation problems (P_N) in the *upper subdifferential form*. In Section 5 we formulate the extension of the AMP to discrete approximations of constrained *time-delay* systems, which is new in both smooth and nonsmooth frameworks.

II. Counterexamples

We start with an example on the violation of the AMP in discrete approximations of linear control systems with linear cost functions and linear endpoint inequality constraints but with *no properness condition*.

Example 2.1: (AMP does not hold in smooth control problems with no properness condition).
Let us consider a linear continuous-time optimal control problem (P) with a two-dimensional state $x = (x_1, x_2) \in I\!R^2$ in the following form:

$$\begin{cases} \text{minimize } \varphi(x(1)) := -x_1(1) \\ \text{subject to} \\ \dot{x}_1 = u, \quad \dot{x}_2 = x_1 - 2t, \quad x_1(0) = x_2(0) = 0, \\ u(t) \in U := \{0,1\}, \quad 0 \le t \le 1, \\ x_2(1) \le -\dfrac{1}{2}. \end{cases} \tag{8}$$

Observe that the only "unpleasant" feature of this problem is that the control set $U = \{0,1\}$ is *nonconvex*, and hence the feasible velocity sets $f(t,x,U)$ are nonconvex as well. It is clear that $\bar{u}(t) \equiv 1$ is the unique optimal solution to problem (8), and that the corresponding optimal trajectory is $\bar{x}_1(t) = t$, $\bar{x}_2(t) = -\frac{1}{2}t^2$. Moreover, the inequality constraint is active, since $\bar{x}_2(1) = -\frac{1}{2}$.

Let us now discretize this problem with the stepsize $h_N := \frac{1}{2N}$, $N \in I\!N$. The discrete approximation problems (P_N) corresponding to (8) are written as:

$$\begin{cases} \text{minimize } \varphi(x_N(1)) = -x_{1N}(1) \\ \text{subject to} \\ x_{1N}(t+h) = x_{1N}(t) + h_N u(t), \quad x_{1N}(0) = 0, \\ x_{2N}(t+h) = x_{2N}(t) + h_N\big(x_{1N}(t) - 2t\big), \quad x_{2N}(0) = 0, \\ u_N(t) \in \big\{0,1\big\}, \quad t \in \big\{0, h_N, \ldots, 1-h_N\big\}, \\ x_{2N}(1) \le -\dfrac{1}{2} + h_N^2, \end{cases}$$

i.e., we put $\gamma_N := h_N^2$ in the constraint perturbation for (P_N). It can be shown that the control

$$u_N(t) := \begin{cases} 1 & t \ne \frac{1}{2}, \\ 0 & t = \frac{1}{2} \end{cases}$$

is optimal and does not satisfy the AMP at the point $t = 1/2$. Observe that the sequence of these controls does *not satisfy the properness property* in the assumption (b) of the AMP formulated in Section 1.

Many examples of this type can be constructed based on the above idea, which essentially means the following. Take a continuous-time problem with active inequality constraints and *nonconvex* admissible velocity sets $f(t,x,U)$. It often happens that after the discretization the "former" optimal control becomes not feasible in discrete approximations, and the "new" optimal control in the sequence of discrete-time problems has a singular point of switch (thus making the sequence of optimal controls not proper), where the approximate maximum condition does not hold.

The next example demonstrates that the AMP may be violated in problems of minimizing *nonsmooth cost* functions in linear systems with no endpoint constraint.

Example 2.2: (AMP does not hold for linear systems with nonsmooth and convex cost functions).
Consider the following sequence of one-dimensional optimal control problems (P_N), $N \in I\!N$, for discrete-time systems:

$$\begin{cases} \text{minimize } \varphi(x_N(1)) := |x_N(1) - 2/3| \\ \text{subject to} \\ x_N(t+h_N) = x_N(t) + h_N u_N(t), \\ x_N(0) = 0, \quad u_N(t) \in U := \big\{0,1\big\}, \\ t \in T_N := \big\{0, h_N, \ldots, 1-h_N\big\}, \end{cases} \tag{9}$$

where $h_N := 10^{-N}$, which is a subsequence of $h_N = N^{-1}$. The dynamics in (9) is a discretization of the simplest ODE control system $\dot{x} = u$. It is easy to see that in this case the set of all reachable points at $t = 1$ is the set of rational numbers between 0 and 1 with exactly N digits in the fractional part of their decimal representations. In particular, for $N = 3$ this set is $\{0, 0.001, 0.002, ..., 0.999, 1\}$. Therefore, the closest point to $x = 2/3$ from the reachable set has N digits in the fractional part and is equal to 0.77...7, and such point must be the endpoint of the optimal trajectory $\bar{x}_N(1)$.

Let us show that in this case the approximate maximum condition does *not* hold at points $t \in T_N$ for which $\bar{u}_N(t) = 1$. (Such points exist, since the optimal control is not identically equal to 0). Indeed, it can be verified that

$$H(\bar{x}_N(t), p_N(t), u) = p_N(t+h_N)u \quad \text{and} \quad p_N(t) \equiv -1$$

for the Hamilton-Pontryagin function and the adjoint trajectory in (6) and (7). Thus

$$\begin{aligned} &\max\nolimits_{u \in U} H(\bar{x}_N(t), p_N(t+h_N), u) = 0 \ \text{ for all } \ t \in T_N, \\ &\text{while } \ H(\bar{x}_N(s), p_N(s+h_N), \bar{u}_N(s)) = -1 \end{aligned}$$

at the points $s \in T_N$, where $\bar{u}_N(s) = 1$ regardless of h_N.

Example 2.2 contradicts the AMP with the transversality condition in the *conventional subdifferential form*, which is

$$-p_N(t_1) \in \partial\varphi(\bar{x}_N(t_1))$$

for problems with no endpoint constraints. In our example the function $\varphi(x) = |x - 2/3|$ is *convex*, and hence the subdifferential ∂ is understood in the sense of convex analysis. Note that in this case the subdifferential agrees with the gradient:

$$\partial\varphi(\bar{x}_N(1)) = \big\{\nabla\varphi(\bar{x}_N(1))\big\} = \big\{1\big\} \ \text{ for all } \ N \in I\!N$$

along the optimal trajectories in (9). Since any reasonable (lower) subdifferential for nonsmooth convex functions must reduce to the convex subdifferential, Example 2.2 proves that there is *no hope for an extension of the AMP in the conventional subdifferential form to problems with nonsmooth costs*.

There are examples, which we omit here, which show that the AMP fails even for problems with *differentiable* but *not continuously differentiable* cost functions and that the AMP is not valid when the *dynamics* of the system is nonsmooth.

III. Uniformly Upper Subdifferentiable Functions

In this section we present some tools of nonsmooth analysis needed for the formulation of the main *positive* results of the paper: the Approximate Maximum Principle for ordinary and time-delay systems in the new *upper subdifferential* form. Results in this form are definitely non-traditional in optimization, since they deal with *minimization* problems for which *lower* subdifferential constructions are usually employed. However, we saw in the preceding section that results of the conventional lower type simply do not hold for the AMP. In Sections 4 and 5 we are going to use *upper* subdifferential constructions for nonsmooth minimization problems of optimal control, which happen to work for a special class of *uniformly upper subdifferentiable* functions described in this section.

Given an extended-real-valued function $\varphi\colon I\!\!R^n \to \overline{I\!\!R} := [-\infty, \infty]$ finite at $\bar{x}$, we first define its *Fréchet upper subdifferential* (or *Fréchet superdifferential*) by

$$\widehat{\partial}^+\varphi(\bar{x}) := \Big\{ x^* \in I\!\!R^n \big| \limsup_{x\to\bar{x}} \frac{\varphi(x)-\varphi(\bar{x})-\langle x^*, x-\bar{x}\rangle}{\|x-\bar{x}\|} \le 0 \Big\}. \tag{10}$$

The set (10) is symmetric to the (lower) Fréchet subdifferential: $\widehat{\partial}^+\varphi(\bar{x}) = -\widehat{\partial}(-\varphi)(\bar{x})$, which is widely used in variational analysis under the name of "regular" or "strict" subdifferential; see, e.g., [9] and [11]. The upper subdifferential (10) is our primary generalized differential construction in this paper. This set is closed and convex but may be empty for many functions useful in minimization. In fact, both $\widehat{\partial}^+\varphi(\bar{x})$ and $\widehat{\partial}\varphi(\bar{x})$ are nonempty simultaneously if and only if φ is Fréchet differentiable at $\bar{x}$, in which case $\widehat{\partial}^+\varphi(\bar{x}) = \widehat{\partial}\varphi(\bar{x}) = \{\nabla\varphi(\bar{x})\}$.

Following [5], we define the *basic upper subdifferential* of φ at $\bar{x}$ by

$$\partial^+\varphi(\bar{x}) := \Big\{ x^* \in I\!\!R^n \big| \; \exists\, x_k \to \bar{x} \;\text{ with }\; \varphi(x_k)\to\varphi(\bar{x}) \;\text{ and }\; \exists\, x_k^* \in \widehat{\partial}^+\varphi(x_k) \;\text{ with }\; x_k^* \to x^* \Big\}$$

and call φ to be *upper regular* at $\bar{x}$ if $\partial^+\varphi(\bar{x}) = \widehat{\partial}^+\varphi(\bar{x})$. This class includes, in particular, all strictly differentiable functions as well as proper concave functions. In the concave case $\widehat{\partial}^+\varphi(\bar{x})$ reduces to the upper subdifferential of convex analysis. It is interesting to observe that, for Lipschitzian upper regular functions, the Fréchet upper subdifferential (10) agrees with Clarke's generalized gradient $\overline{\partial}\varphi(\bar{x})$ of [1].

Let us now define a class of functions for which we obtain an extension of the AMP to nonsmooth control problems in the next section.

Definition 3.1: A function $\varphi\colon I\!\!R^n \to \overline{I\!\!R}$ is *uniformly upper subdifferentiable* around a point $\bar{x}$ where it is finite, if there is a neighborhood V of $\bar{x}$ such that for every $x \in V$ there exists $x^* \in I\!\!R^n$ with the following property: given any $\varepsilon > 0$, there is $\eta > 0$ for which

$$\varphi(v) - \varphi(x) - \langle x^*, v-x\rangle \le \varepsilon\|v-x\| \tag{11}$$

whenever $v \in V$ with $\|v-x\| \le \eta$.

It is easy to check that the class of uniformly upper subdifferentiable functions includes all continuously differentiable functions, concave continuous functions, and also it is closed with respect to taking minimum over compact sets. Note that if φ is Lipschitz continuous and differentiable at some point, it may not be uniformly upper subdifferentiable around it. Example: $\varphi(x) = x^2\sin(1/x)$ for $x \neq 0$ with $\varphi(0) = 0$.

IV. AMP in Upper Subdifferential Form

This section of the paper collects the main positive results on the fulfillment of the AMP in the upper subdifferential form for the discrete approximation problems (P_N). We state two closely related versions of the AMP in somewhat different settings of (P_N). The first version applies to problems with *no endpoint constraints* and establishes the upper subdifferential form of the AMP with *no properness* requirement on the sequence of optimal controls and with an error estimate as $\varepsilon(t, h_N) = O(h_N)$ in the approximate maximum condition. The second result is the major version of the AMP for the *constrained nonsmooth* problems (P_N), which extends that formulated in Theorem 1.1.

Let us start with the upper subdifferential form of the AMP for problems with no endpoint constraints. Throughout this section impose the following *standing assumptions* on the mapping f and the control set U:

(H1) $f = f(t, x, u)$ is continuous with respect to all its variables and continuously differentiable with respect to the state variable x in some tube containing optimal trajectories for all u from the compact set U in a metric space and for all $t \in T_N$ uniformly in $N \in I\!\!N$.

Theorem 4.1: (AMP for problems with no endpoint constraints). Let the pairs $(\bar{x}_N, \bar{u}_N)$ be optimal to problems (P_N) with $\varphi_i = 0$ for all $i = 1, \ldots, m+r$. Assume in addition to (H1) that φ_0 is uniformly upper subdifferentiable around the limiting points of the sequence $\{\bar{x}_N(t_1)\}$, $N \in I\!\!N$. Then for every sequence of upper subgradients $x_N^* \in \widehat{\partial}^+\varphi_0(\bar{x}_N(t_1))$ there is $\varepsilon(t, h_N) \to 0$ as $N \to \infty$ uniformly in $t \in T_N$ such that the approximate maximum condition (3) holds for all $t \in T_N$, where each $p_N(t)$ satisfies the adjoint system (6) with the transversality condition

$$p_N(t_1) = -x_N^* \;\text{ for all }\; N \in I\!\!N. \tag{12}$$

Moreover, $\varepsilon(t, h_N) = O(h_N)$ in (3) if φ_0 is locally concave around $\bar{x}_N(t_1)$ uniformly in N and $\partial f(\cdot, u, t)/\partial x$ is locally Lipschitz around $\bar{x}_N(t)$ with a constant uniform in $u \in U$, $t \in T_N$, $N \in I\!\!N$.

Remark. *(Upper versus lower subdifferential forms of transversality conditions).* The main difference between the conventional (lower) subdifferential form, which is *not* actually fulfilled in the case of AMP, and the upper subdifferential form of Theorem 4.1 is that the transversality condition (12) holds for *every* upper subgradient $x_N^* \in \widehat{\partial}^+\varphi_0(\bar{x}_N(t_1))$ instead of just *some* lower subgradient in the conventional transversality conditions for continuous-time and discrete-time (with a fixed step) systems. In particular, for discrete-time systems with convex velocity sets both lower and upper subdifferential forms of the (exact) discrete maximum principle hold; see [7], where the upper subdifferential/superdifferential form of the discrete maximum principle has been established under milder assumptions on φ_0 in comparison with Theorem 4.1. If φ_0 is Lipschitz continuous and upper regular and hence $\widehat{\partial}^+\varphi_0(\bar{x}) = \overline{\partial}\varphi_0(\bar{x})$, there is indeed a *dramatic* difference between the upper subdifferential form of transversality conditions and a well-recognized form in terms of the Clarke subdifferential: instead of the fulfillment transversality just for some element of $\overline{\partial}\varphi_0(\bar{x}(t_1))$ we establish its fulfillment for the *whole set*! Similar situation takes place for continuous-time systems, where the upper subdifferential form of transversality in the maximum principle can be proved for problems with no endpoint constraints in the line of arguments of Theorem 4.1. Observe, however, that there is a more subtle lower subdifferential form of transversality conditions for continuous-time and discrete-time (of a fixed step) systems that involves basic/limiting subgradients rather than those of Clarke; see [5], [11]. Note that the major drawback of the upper subdifferential form is that it applies to a restrictive class of functions. But, as we saw in Section 2, there is *no alternative* to this form for the Approximate Maximum Principle.

Next let us consider a sequence of the discrete approximation problems (P_N) with *endpoint constraints* of the inequality and equality types. We are going to derive an extension of the AMP formulated in Section 1 to these problems involving *nonsmooth* functions that describe the cost and inequality constraints. The following upper subdifferential version of the AMP for constrained problems require the *uniform upper subdifferentiability* property on the cost and the inequality constraint functions, the *properness* of the sequence of optimal controls, and the *consistency* condition on the perturbations of the equality constraints. As we saw in Section 2, all the three requirements are essential.

Theorem 4.2: (AMP for problems with endpoint constraints). Let the pairs $(\bar{x}_N, \bar{u}_N)$ be optimal to problems (P_N). In addition to (H1) assume

(H2) the sequence of optimal controls $\{\bar{u}_N\}$ is proper,

(H3) φ_i are uniformly upper subdifferentiable around the limiting points of $\{\bar{x}_N(t_1)\}$ for $i = 0, \ldots, m$ and continuously differentiable around them for $i = m+1, \ldots, m+r$,

(H4) the consistency condition (2) holds for the perturbations δ_{iN} of the equality constraints.

Then for any sequences of upper subgradients $x_{iN}^* \in \widehat{\partial}^+\varphi_i(\bar{x}_N(t_1))$, $i = 0, \ldots, l$, there are numbers $\{\lambda_{iN} |\ i = 0, \ldots, m+r\}$ and a function $\varepsilon(t, h_N) \downarrow 0$ as $N \to \infty$ uniformly in $t \in T_N$ such that the approximate maximum condition (3) is fulfilled with the adjoint trajectory $p_N(t)$ to (6) satisfying the transversality condition

$$p_N(t_1) = -\sum_{i=0}^{m} \lambda_{iN} x_{iN}^* - \sum_{i=m+1}^{m+r} \lambda_{iN} \nabla\varphi_i(\bar{x}_N(t_1)) \quad (13)$$

along with

$$\lambda_{iN}(\varphi_i(\bar{x}_N(t_1)) - \gamma_{iN}) = O(h_N) \quad \text{for} \quad i = 1, \ldots, m, \quad (14)$$

$$\lambda_{iN} \geq 0 \ \text{ for } \ i = 0, \ldots, m, \quad \text{and} \quad \sum_{i=0}^{m+r} \lambda_{iN}^2 = 1. \quad (15)$$

V. AMP for Discrete Approximations of Delay Systems

This section is devoted to the extension of the AMP in the upper subdifferential form to finite-difference approximations of *time-delay* control systems. Actually we are not familiar with any previous results on the AMP for optimal control problems with delays, so the results obtained below seem to be new even for smooth delay problems.

We pay the main attention to discrete approximations of the following time-delay problem with no endpoint constraints:

$$(D) \qquad \begin{cases} \text{minimize } J(x,u) := \varphi(x(t_1)) \\ \text{subject to} \\ \dot{x}(t) = f(t, x(t), x(t-\theta), u(t)) \ \text{ a.e. } \ t \in [t_0, t_1], \\ x(t) = c(t), \quad t \in [t_0 - \theta, t_0], \\ u(t) \in U \ \text{ a.e. } \ t \in [t_0, t_1] \end{cases}$$

over measurable controls and absolute continuous trajectories, where $\theta > 0$ is the constant time-delay, and where $c\colon [t_0 - \theta, t_0] \to I\!R^n$ is a given function defining the initial "tail" condition that is necessary to start the delay system. Based on the above constructions for non-delayed systems, one can derive similar results for delay systems with endpoint constraints. We may also extend the results obtained to more complicated delay systems involving variable delays, set-valued tail conditions, etc. On the other hand, there are examples, which we omit here, that show, that the AMP does *not hold* for discrete approximations of even smooth functional-differential systems of *neutral type* that contain time-delays not only in state variables but in velocity variables as well.

Let us build discrete approximations of the time-delay problem (D) based on the Euler finite-difference replacement of the derivative. In the case of time-delay systems we need to ensure that the point $t - \theta$ belongs to the discrete grid whenever t does. It can be achieved by defining the discretization step as $h_N := \theta/N$ in contrast to $h_N = (t_1 - t_0)/N$ for the non-delayed problems (P_N). In such a scheme the length of the time interval $t_1 - t_0$ is generally *no longer commensurable* with the discretization step h_N.

To this end we consider the following sequences of discrete approximations of the delay problem (D) with the grid on the main interval $[t_0, t_1]$ given by

$$T_N := \{t_0, t_0 + h_N, \dots, t_1 - \widetilde{h}_N - h_N\}, \quad h_N := \frac{\theta}{N},$$
$$\widetilde{h}_N := t_1 - t_0 - h_N\Big[\frac{t_1 - t_0}{h_N}\Big],$$

(here $[a]$ denotes the greatest integer less than or equal to the real number a) and also involving the grid T_{0N} on the initial interval
$[t_0 - \theta, t_0]$:

(D_N)

$$\begin{cases} \text{minimize } J(x_N, u_N) := \varphi(x_N(t_1)) \\ \text{subject to} \\ x_N(t + h_N) = x_N(t) + \\ \qquad h_N f(t, x_N(t), x_N(t - Nh_N), u_N(t)),\ t \in T_N, \\ x_N(t_1) = x_N(t_1 - \widetilde{h}_N) + \\ \qquad \widetilde{h}_N f(t_1 - \widetilde{h}_N, x_N(t_1 - \widetilde{h}_N), u_N(t_1 - \widetilde{h}_N)), \\ x_N(t) = c(t),\ t \in T_{0N} := \{t_0 - \theta, t_0 - \theta + h_N, \dots, t_0\}, \\ u_N(t) \in U, \quad t \in T_N, \end{cases}$$

where $[a]$ stands, as usual, for the greatest integer less than or equal to the real number a.

Our assumptions on the initial data of (P) are similar to those in Section 4 for non-delay systems. A counterpart of (H1) is formulated as:

(H) $f = f(t, x, y, u)$ is continuous with respect to all its variables and continuously differentiable with respect to (x, y) in some tube containing optimal trajectories for all u from the compact set U in a metric space and for all $t \in \widetilde{T}_N := T_N \cup \{t_1 - \widetilde{h}_N\}$ uniformly in $N \in I\!N$.

For convenience we introduce the following notation:

$$\begin{aligned} &\xi_N(t) := (x_N(t), x_N(t - \theta)), \\ &\bar{\xi}_N(t) := (\bar{x}_N(t), \bar{x}_N(t - \theta)), \\ &f(t, \xi_N, u_N) := f(t, x_N(t), x_N(t - \theta), u_N(t)), \\ &f(t, \bar{\xi}_N, u_N) := f(t, \bar{x}_N(t), \bar{x}_N(t - \theta), u_N(t)), \end{aligned}$$

and write the *adjoint system* to (D_N) as

$$p_N(t) = p_N(t + h_N) + h_N \frac{\partial f^*}{\partial x}(t, \bar{\xi}_N, \bar{u}_N) p_N(t + h_N)$$
$$+ h_N \frac{\partial f^*}{\partial y}(t + \theta, \bar{\xi}_N, \bar{u}_N) p_N(t + \theta + h_N) \quad \text{for } t \in T_N,$$

$$p_N(t_1 - \widetilde{h}_N) = p_N(t_1) + \widetilde{h}_N \frac{\partial f^*}{\partial x}(t_1 - \widetilde{h}_N, \bar{\xi}_N, \bar{u}_N) p_N(t_1)$$

along the optimal processes $(\bar{x}_N, \bar{u}_N)$ to the delay problems for each $N \in I\!N$. Introducing the corresponding *Hamilton-Pontryagin function*

$$H(t, x_N, y_N, p_N, u) := \begin{cases} \langle p_N(t + h_N), f(t, x_N, y_N, u) \rangle & \text{if } t \in T_N, \\ \langle p_N(t), f(t - \widetilde{h}_N, x_N, y_N, u) \rangle & \text{if } t = t_1 - \widetilde{h}_N, \end{cases} \tag{16}$$

with $\bar{y}_N(t) = \bar{x}_N(t - \theta)$, we rewrite the adjoint system as

$$\begin{cases} p_N(t) = p_N(t + h_N) + h_N\Big[\frac{\partial H}{\partial x}(t, \bar{\xi}_N, p_N, \bar{u}_N) + \\ \qquad \frac{\partial H}{\partial y}(t + \theta, \bar{\xi}_N, p_N, \bar{u}_N)\Big],\ t \in T_N, \\ p_N(t_1 - \widetilde{h}_N) = p_N(t_1) + \widetilde{h}_N \frac{\partial H}{\partial x}(t_1 - \widetilde{h}_N, \bar{\xi}_N, p_N, \bar{u}_N). \end{cases} \tag{17}$$

Theorem 5.1: (AMP for delay systems). Let the pairs $(\bar{x}_N, \bar{u}_N)$ be optimal to problems (D_N). Assume in addition to (H) that φ is uniformly upper subdifferentiable around the limiting points of the sequence $\{\bar{x}_N(t_1)\}$, $N \in I\!N$. Then for every sequence of upper subgradients $x_N^* \in \widehat{\partial}^+ \varphi(\bar{x}_N(t_1))$ the approximate maximum condition

$$H(t, \bar{\xi}_N, p_N, \bar{u}_N) = \max_{u \in U} H(t, \bar{\xi}_N, p_N, u) - \varepsilon(t, h_N),\ t \in \widetilde{T}_N,$$

holds with the Hamilton-Pontryagin function (16) and with some $\varepsilon(t, h_N) \downarrow 0$ as $h_N \to 0$ uniformly in $t \in \widetilde{T}_N$, where the adjoint trajectory p_N satisfies (17) and the transversality relations

$$p_N(t_1) = -x_N^*, \qquad p_N(t) = 0 \ \text{ if } \ t > t_1. \tag{18}$$

Note that in the case of continuously differentiable cost functions φ around $\bar{x}_N(t_1)$ uniformly in N, the transversality relations (18) reduce to

$$p_N(t_1) = -\nabla\varphi(\bar{x}_N(t_1)), \qquad p_N(t) = 0 \ \text{ if } \ t > t_1.$$

Similarly to the proof of Theorem 5.1 we can deduce from Theorem 4.2 its delay counterpart for discrete approximation problems with *endpoint constraints*. In this result we add assumptions (H2)-(H4) to those in (H) and replace the transversality relations (18) in Theorem 5.1 by the conditions (13)–(15) with $p_N(t) = 0$ if $t > t_1$.

REFERENCES

[1] F. H. Clarke, *Optimization and Nonsmooth Analysis*, Wiley, New York, 1983.
[2] R. Gabasov and F. M. Kirillova, *On the extension of the maximum principle by L.S. Pontrygin to discrete systems,* Autom. Telem., 27 (1966), pp. 46-61.
[3] R. Gabasov and F. M. Kirillova, *Qualitative Theory of Optimal Processes,* Marcel Dekker, New York, 1976.
[4] B. S. Mordukhovich, *Approximate maximum principle for finite-difference control systems,* U.S.S.R. Comput. Maths. Math. Phys., 28 (1988), pp. 106–114.
[5] B. S. Mordukhovich, *Approximation Methods in Problems of Optimization and Control*, Nauka, Moscow, 1988.
[6] B. S. Mordukhovich, *Discrete approximations and refined Euler–Lagrange conditions for nonconvex differential inclusions*, SIAM J. Control Optim., 33 (1995), pp. 882–915.
[7] B. S. Mordukhovich and I. Shvartsman, *Discrete maximum principle for nonsmooth optimal control problems with delays*, Cybern. Syst. Anal., 2002, No. 2, pp. 123–133.
[8] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidza and E. F. Mischenko, *The Mathematical Theory of Optimal Processes*, Wiley, New York, 1962.
[9] R. T. Rockafelar and R. J.-B. Wets, *Variational Analysis*, Springer,
[10] G. V. Smirnov, *Introduction to the Theory of Differential Inclusions*, American Mathematical Society, Providence, RI, 2002.
[11] R. B. Vinter, *Optimal Control*, Birkhäuser, Boston, 2000.